

Research Goals in Epidemiology

LEARNING OBJECTIVES

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

- Use the research goal as the organizing principle of a study design.
- Distinguish between general research goals.
- Explain the requirements for causality.
- Realize the practical limitations of a research goal.

CHAPTER OUTLINE

- I. Introduction
- II. Types and Examples of Research Goals
- III. Practical Limitations of a Research Goal
- IV. Conclusion
- V. Vocabulary
- VI. Study Questions and Exercises

What is the meaning of life?

INTRODUCTION

Well, that question is probably better suited for a philosopher than an epidemiologist. To determine appropriate research goals in epidemiology, significance, scientific rigor, and feasibility are key considerations. Questioning the meaning of life is certainly significant in terms of importance to humankind, but designing a study to answer this question is decidedly not feasible in terms of scope, time, and cost; and a rigorous scientific design, given the scope of this question, would be impossible.

Starting with a question that is not feasible or even possible to answer illustrates how the research goal is the organizing principle for every element of the study. In this chapter, we will review the feasible epidemiologic research goals as well as their practical limitations.

TYPES AND EXAMPLES OF RESEARCH GOALS

We will start with the least ambitious and most limited type of goal and move to the most ambitious and least limited research goal. Within the context of the overall substantive goals of epidemiology as a field, a limited and less ambitious goal is not a compromise of practical significance. Description is usually the first step in the quest for causality.

Descriptive Research Goal

A descriptive research goal is intended to “describe” a health phenomenon in terms of its distribution across person, place, and time. The goal is not to test a hypothesis about the causes or even correlates of the phenomenon. Rather, the value of a descriptive study is to gain information necessary to formulate a hypothesis. This goal is appropriate for emerging or rare diseases or health problems about which little is known. It is also useful for monitoring and tracking diseases over time.

Incidence and Prevalence

The health phenomenon under study can be anything that compromises human physical and mental health, from cancer to infectious diseases, and from drug abuse to traffic injuries. The health problem is typically measured as incidence and prevalence. Briefly, incidence is the number of new cases of disease or death during a specific period of time out of the total population at risk for the disease or death. The population at risk can be either the number of subjects without

the disease at the onset of the study (cumulative incidence) or the total number without the disease multiplied by their time at risk during the study period (incidence density or incidence rate). Prevalence is the number of existing cases of disease or deaths out of the total population of interest. Calculating incidence or prevalence is a basic, but very important component of a descriptive study.*

Person, Place, and Time

Beyond incidence and prevalence, a descriptive study could focus on the “who, when, and where” (person, time, and place) of the health phenomenon of interest. Characteristics of the person with the health problem include sociodemographic characteristics, such as age and gender, as well as other characteristics, such as general health status, or risk factors (like cigarette smoking) that may be relevant to the disorder. Characteristics of time include the hour, day, week, and so on of an outbreak, as well as collective measures of time like historical periods, cycles, and trends. Characteristics of the place where the health problem occurs include geographic, geologic, climactic, and similar criteria. In the case of outbreaks, specific buildings or similar localized areas would be the focus. The goal of descriptive research is to gather “clues” about the concentration of health phenomena among specific populations, during particular times, and localized in key areas. Ideally, these clues suggest potential causes and, subsequently, hypotheses to be tested and interventions to be evaluated.

Example: Obesity Research

Technically, childhood obesity in the United States is an epidemic (the prevalence and incidence is higher than what would be expected historically) with its beginning increase in the late 1970s (Ogden, Carroll, Curtin, et al., 2010). The increase in prevalence since 1971 is shown in **Figure 2–1**. These results are an example of a descriptive study with a focus on the timing of the disease or condition. The unexpected increase in prevalence among children and adolescents, and even among adults in the United States was alarming enough to stimulate additional descriptive and analytic research on obesity.

Focus on the person characteristics of childhood obesity shows that Hispanic boys and non-Hispanic black girls are at increased risk for obesity (Ogden et al., 2010). More focused person-centered research on adolescent obesity shows that

*For more detailed information about incidence and prevalence, see *Essential Epidemiology: Principles and Applications* by William Oleckno (2002) and *Epidemiology for Public Health Practice* by Robert Friis and Thomas Sellers (2009).

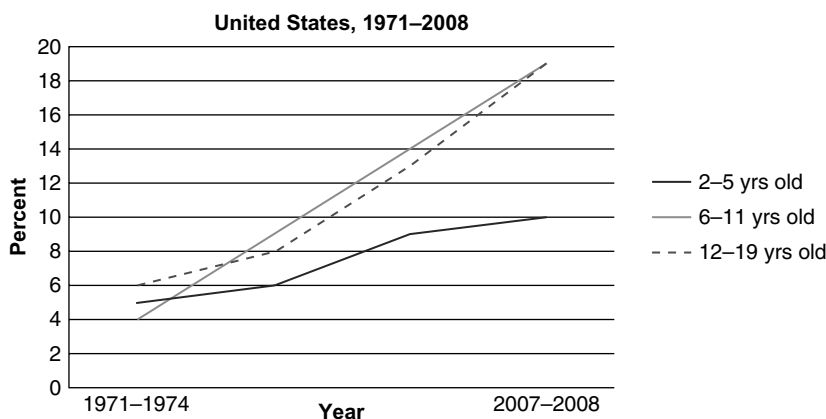


FIGURE 2–1 Increase in Childhood Obesity

Source: Data from Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., & Flegal, K. M. (2010). Prevalence of high body mass index in US children and adolescents, 2007–2008. *Journal of the American Medical Association*, 303(3), 242–249; AMA Publishing Group © 2010.

adolescents with autism and Down syndrome have an exceptionally high prevalence of obesity (Murray & Ryan-Krause, 2010; Rimmer, Yamaki, Lowry, et al., 2010). These descriptive results stimulate a focus on potential cultural, genetic, and biological causes of obesity in these subgroups of adolescents, as well as identify these groups as appropriate target populations for prevention and intervention.

We move to Canada for an example of a descriptive study of obesity that focuses on clustering and comparisons of prevalence by place—urban and rural. The Canadian Heart Health Surveys Research Group (Reeder, Chen, Macdonald, et al., 1997) compared obesity prevalence in nine Canadian provinces from 1986 to 1992. They found no urban–rural differences in obesity for men and women in all provinces except those in western Canada. In western Canada, rural men and women are more likely to be obese compared to their urban counterparts. These results stimulate research to pinpoint the reasons why there are rural–urban differences in western Canada and indicate a need for tailored intervention programs in this region of Canada.

Association Research Goal

A next important research goal in epidemiology is to determine factors that are associated with or related to morbidity and mortality. This type of study typically moves beyond person, place, and time comparisons to comparisons based on potential risk factors for disease and death.

Risk Factors

Risk factors are exposures that are associated with a disease (Friis & Sellers, 2009). Exposures can be a vast array of experiences such as contact with an infectious agent, behaviors including high-fat diet and excessive alcohol use, contact with environmental poisons, and conditions of life such as crowded housing, low income, and stress. Risk factors are especially relevant for epidemiologic research because the etiology of most diseases is a complex combination of causes and potential causes. For many practical and empirical reasons, the relationships between risk factors and disease onset cannot always be shown to be causal. For example, as reviewed later in this chapter, coronary heart disease has been associated with a number of risk factors, including elevated serum cholesterol, high blood pressure, obesity, and cigarette smoking (Kagan, Kannel, Dawber, et al., 1963).

According to Friis and Sellers (2009), an exposure that is associated with a disease or other health problem can be considered a risk factor if it meets the following criteria:

1. There is a dose-response relationship—the higher the level or intensity of the exposure, the higher the probability or severity of the disease.
2. Temporality of the exposure and disease are appropriate—the exposure precedes the onset of the disease in time.
3. The observed relationship between exposure and the disease is not due to some source of error in the design or conduct of the study.

These criteria are a subset of the criteria needed to establish causality, as discussed later in this chapter.

But like most things in life, the identification of risk factors is more complicated than simply demonstrating a relationship between exposure and disease onset. Robust results are those that consider potential sources of error in relationships, as well as other factors that may influence the significance, strength, or direction of relationships between an exposure and disease.

Measures of Association or Effect

In epidemiologic studies, measures of association test relationships between exposures and outcomes. These measures are used in studies focusing on associations and those testing causal relationships (effects). Their interpretation depends on whether certain criteria for causality are met (discussed shortly). The measures are best understood in the context of contingency or two-by-two tables with the exposure reported in the rows and the outcome in the columns. **Table 2–1** shows an example contingency table, as well as formulas for the most commonly used measures.

Table 2-1 Two-by-Two Table Showing the Relationship Between an Exposure and an Outcome

		Outcome		
		Has outcome	Does not have outcome	Total
Exposure	Exposed	a	b	a + b
	Not exposed	c	d	c + d
	Total	a + c	a + d	a + b + c + d

Risk Ratio = $\frac{[a / (a + b)]}{[c / (c + d)]}$ = [prevalence or incidence of outcome for the exposed group] divided by [prevalence or incidence of the outcome for the unexposed group]

Risk Difference = $[a / (a + b)] - [c / (c + d)]$ = [prevalence or incidence of outcome for the exposed group] minus [prevalence or incidence of the outcome for the unexposed group]

Odds Ratio = $\frac{(a / c)}{(b / d)}$ = (exposed divided by unexposed for those with the outcome, or cases) divided by (exposed divided by unexposed for those without the outcome, or controls)

Attributable Risk = $\frac{[a / (a + b)] - [c / (c + d)]}{[a / (a + b)]}$ = Proportion of the prevalence or incidence of the outcome for the exposed group that can be attributed to the exposure

Population Attributable Risk = $\frac{[(a + c) / (a + b + c + d)] - [c / (c + d)]}{[(a + c) / (a + b + c + d)]}$ = Proportion of the total prevalence or incidence that can be attributed to the exposure

The purpose of all of these calculations is to measure how strongly the outcome is associated with or attributable to the exposure. They can be calculated by comparing incidences or comparing prevalences of the outcome. Risk ratios and risk differences (called rate ratios and rate differences for comparisons of incidence densities) are used in cross-section and cohort study designs, and odds ratios are used in case-control studies. Attributable risk (AR) and population attributable risk (PAR) are descriptive measures showing the proportion of the incidence or prevalence of the outcome that is attributed to (or the result of, or explained by, or associated with) the exposure.*

Confounding

Confounding is present when a third factor or variable distorts the relationship between an exposure and disease (Rothman, 1986). The distortion can take the form of erroneously inflating or deflating the strength of the association between the exposure and disease. Consequently, confounding is often a form of systematic error

*For more detailed information about measure of association, see *Essential Epidemiology: Principles and Applications* by William Oleckno (2002) and *Epidemiology for Public Health Practice* by Robert Friis and Thomas Sellers (2009).

or bias. Reporting an association between exposure and disease without testing for obvious or even potential confounding in the association is a biased study and result.

Testing for confounding involves comparing the magnitude and even direction of the simple association between exposure and outcome (crude association) with the adjusted association when the potential confounder is included in the analysis. Confounding is present when the following criteria are met:

1. The confounding variable has an effect, either as a risk factor or cause, on the disease.
2. The confounding variable has an effect on the exposure, independent of the disease.
3. The confounding variable does not intervene in time and the causal chain between the exposure and disease. In other words, the exposure does not influence the confounding variable.

Figure 2–2a illustrates these criteria of associations.

Although not epidemiologic, the following example provides a clear illustration of confounding. A city planning intern noticed that there is a dose-response relationship between the number of fire trucks present at a fire and the amount of property damage at the site of the fire. Specifically, the greater the number of fire trucks, the greater the amount of property damage, as illustrated in **Figure 2–2b**.

The intern proposes to his preceptor that efforts should be made to limit the number of fire trucks called to a fire in order to limit the amount of property damage. The preceptor suggests that the intern study this relationship by observing several fires and measuring any feasible relevant factors associated with property damage. **Figure 2–2c** illustrates the results of the study.

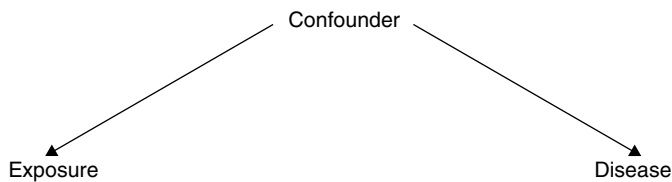


FIGURE 2–2a Illustration of the Criteria for Confounding

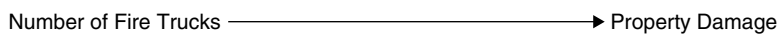


FIGURE 2–2b Relationship Between the Number of Fire Trucks and the Severity of Property Damage

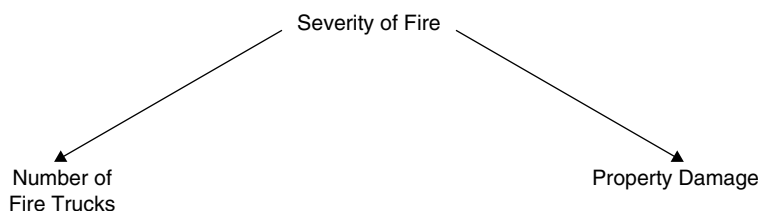


FIGURE 2–2c Severity of the Fire Confounds the Relationship Between the Number of Fire Trucks and Amount of Property Damage

The intern was able to show that the association between the number of fire trucks and property damage was very weak and not statistically significant when the severity of the fire is considered by examining the associations between trucks and damage for mild fires, for moderate fires, and for severe fires. In addition, he found that the severity of the fire influences both the number of trucks called to the fire and the amount of property damage due to the fire. Subsequently, the intern withdrew his original proposal.

Confounding can be controlled or prevented in all stages of the research study. In the design or planning stage, potential confounders can be included in the measures planned for data collection. In the sampling stage, subjects can be randomly selected or randomly assigned to groups in order to minimize systematic differences between groups and between those selected and not selected for the study. Other preventative sampling measures include matching characteristics between study groups and restriction (limiting or excluding) based on the confounder. Finally, in the analysis stage, analyses can be conducted within stratification or grouping based on the confounder, or multivariate models including the confounder can be tested.

Effect Modification (Interaction)

While an unrecognized confounding relationship can be a bias in a study, effect modification, also called interaction, is a meaningful result. Effect modification is present if the strength or direction of the relationship between exposure and disease differs for subgroups of a third factor, such as gender or age group (Hennekens, Buring, & Mayrent, 1987). The demonstration of effect modification adds additional information about the association between exposure and disease onset.

For example, an investigator determines, in a hypothetical study, that a measure of association between lead exposure and brain damage has a value of 5.0 (a quite strong relationship). Because she also hypothesizes that the risk of lead exposure for older children with brain damage would be greater than that for infants with brain damage, she examines the relationship between lead exposure and brain damage separately for infants (younger than 1 year of age), preschool children (ages 1 to 4 years), and young school children (ages 5 to 9 years). She finds that the measure of association is 2.0 for infants, 4.5 for preschool children, and 6.0 for young school children. She concludes that the risk of lead exposure increases by age among children with brain damage. In other words, age modifies the effect of lead exposure on brain damage. Again, this is a hypothetical result.

The presence of effect modification can be tested in two general ways. As was done in the hypothetical study, the data can be stratified or grouped according to categories of the potential modifier and the association between exposure and disease determined within the groups, then compared across the groups. Second, an interaction term, essentially multiplying the exposure by the modifier, can be tested in multivariate models.

Example: The Framingham Heart Study

This, then, is what has been called the coronary profile, a picture of the individual most prone to develop coronary heart disease. With some of these features still subject to confirmation, he may be described as a mesomorphic, obese, middle-aged male, with high serum cholesterol, high blood pressure, low vital capacity, and an abnormal electrocardiogram. He eats too much of too rich foods, smokes cigarettes to excess, and is physically inactive both in occupation and in recreation. He is ambitious, aggressive, and subject to frequent deadlines and other emotional stresses. The closer an individual comes to fitting this pattern, the greater should be the efforts of his physician to alter, where practicable, these characteristics of the patient and his environment (Kagan et al., 1963, p. 893).

Much of what we know currently about the risk factors for cardiovascular disease (CVD) has been demonstrated in a long and ongoing history of studies collectively known as the Framingham Heart Study. The study began in 1948 in response to the shift in the 20th century from infectious diseases to chronic disorders as the leading causes of death in the United States. At this time, CVD emerged as the leading cause of death.

More than 5,000 adults aged 30 to 62 years and free of CVD were recruited for the study in Framingham, Massachusetts. Framingham was chosen as the study site because its residents numbered high enough to provide the desired quota of subjects, were relatively heterogeneous in ethnicity and socioeconomic status, and were stable in terms of limited out-migration (Feinleib, 1983). Subjects returned to the study every 2 years for medical examinations and testing. In 1971, their adult children were enrolled in the study. In 2002, their adult grandchildren were recruited to participate. In addition, another group was enrolled in 1994 to reflect the growing ethnic and socioeconomic diversity of Framingham (Framingham Heart Study, 2012). As the study has progressed, the research goals have also expanded to include examinations of health phenomena beyond CVD.

In addition to the risk factors for CVD summarized in the quotation on the previous page (Kagan et al., 1963), Framingham studies have found associations in the form of risk factors between menopause and CVD (Gordon, Kannel, Hjortland, et al., 1978), use of oral contraceptives and both hypertension and thromboembolism (Kannel, 1979), chronic atrial fibrillation and stroke (Wolf, Dawber, Thomas, et al., 1978), cigarette smoking and reduced levels of high-density lipoproteins (HDL) (Garrison, Kannel, Feinleib, et al., 1978), diabetes and CVD (Kannel & McGee, 1979), combined use of oral contraceptives and smoking and thrombosis among women 35 years and older (Castelli, 1999), and chronic cough and myocardial infarction (Haider, Larson, O'Donnell, et al., 1999). Identified protective factors against CVD include higher levels of HDL (Gordon, Castelli, Hjortland, et al., 1977) and moderate alcohol use (Kannel & Ellison, 1996).

Confounding and effect modification have also been demonstrated in the Framingham studies. A perplexing negative relationship between low body weight and elevated mortality in men was found to be confounded by cigarette smoking (Garrison, Feinleib, Castelli, et al., 1983). Cigarette smoking leads to both low body weight and elevated mortality among men. However, more recent studies of the relationship between low body weight and mortality have demonstrated a relationship independent of cigarette smoking (Woods, Iuliano-Burns, & Walker, 2011). Many cases of effect modification have also been shown, usually modification by gender and age. The relationship between higher serum cholesterol and CVD among men is modified by age in that the increase in risk is sevenfold among men younger than 50 years of age, but only two-and-one-half-fold for men age 50 years and older (Kagan et al., 1963). There is an association between gout and CVD for men, but not for women (Abbott, Brand, Kannel, et al., 1988). Among men, Raynaud's phenomenon (a circulatory disorder) is associated with age and with smoking, but among women, it is associated with

marital status and alcohol use (Fraenkel et al., 1999). These results may be due to differences in the prevalence of behavioral factors between genders rather than true causal mechanisms. Among men and women younger than 65 years, Type A behavior and emotional lability are associated with CVD for women, but worries about aging are associated with CVD for men (Haynes, Feinleib, Levine, et al., 1978). As the Framingham tradition continues, these and many more proven associations have led to hypotheses, both tested and untested, about the causes of CVD and other health phenomena.

Causal Research Goal

The ultimate goal of epidemiologic research is to determine the causes of morbidity and mortality. Confidence that a factor is truly a cause depends on a rigorous study design, repetition of studies showing the same result, and adherence to strict criteria developed through centuries of research by scientific thinkers, including Robert Koch and Jakob Henle (Evans, 1976), Sir Austin Bradford Hill (1965), and Mervyn Susser (1977).

Validity and Reliability

Although a valid and reliable study is the ideal for all research goals (descriptive, association, causal, evaluation), accuracy and precision are vital to make inferences about causality. Confounding bias in a relationship between exposure and disease would compromise the status of an exposure as a true cause. A sample that does not adequately represent the target population to which results are generalized (lack of external validity) compromises the inference that the exposure is a cause in the presumed context. A potential cause that is not accurately measured cannot represent a true cause no matter the strength of the relationship between exposure and disease. Bias and error in study design and conduct compromise the ability to conclude that an exposure is an actual cause of morbidity or mortality.

The Scientific Method

A discussion of the scientific method is particularly informative in the context of appreciating what is required to infer a causal relationship. The method has its roots in ancient Egypt and Greece with substantial refinements and expansions made by Muslim scientist al-Haytham in the 11th century, Francis Bacon and René Descartes in the 17th century, and John Stuart Mill in the 19th century.

Figure 2–3 illustrates a modern conceptualization of the necessary steps in the scientific method.

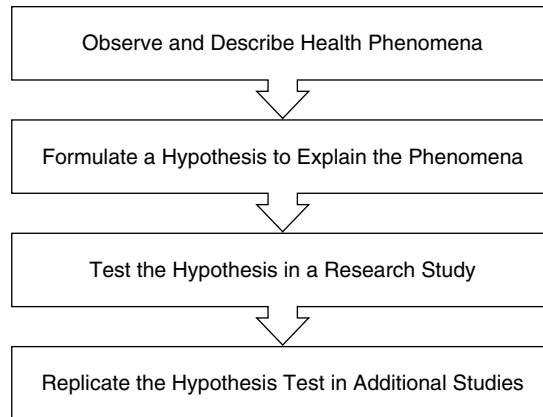


FIGURE 2–3 Steps in the Scientific Method

Note that the first step in the scientific method, to observe and describe health phenomena, is the goal of descriptive epidemiology described earlier in this chapter. The requirements for a causal research goal include the remaining steps of formulating, testing, and retesting hypotheses. Suppose we observe that sickle cell anemia is more common among African Americans compared to other racial or ethnic groups. Naturally, we begin to hypothesize about the reasons for this observation. Our hypothesis might focus on genetic factors unique to African Americans. Perhaps we would focus on environmental factors or cultural factors or some combination of these. Next, we would test our hypothesis by designing a study to obtain a sample that is representative of African Americans and to measure with optimal validity and precision candidate genes, prevalence of potentially relevant factors (such as exposure to malaria) in the historical environment, and whatever else we hypothesize to be relevant. Not only do we strive to design a valid and precise study, but also one that is replicable by other investigators and in other contexts. Every aspect of our study must be empirical or observable and measurable. The greater the number of subsequent studies that replicate our results, especially in different contexts such as multiple study designs, alternative but still precise measurements, and various samples of African Americans, the stronger our case for a causal relationship.

Criteria for Causality

In 1882, Robert Koch demonstrated the “germ theory” of disease by specifying disease organisms as the causes of specific diseases. With later refinement of the

theory by Jakob Henle in 1887, the Henle-Koch postulates were developed in an effort to prove the pathogenesis of disease (Evans, 1976). They specified that the infectious agent must be:

1. Present in every case of the disease;
2. Isolated and grown in pure culture;
3. The cause of the disease when introduced into a healthy host;
4. Recoverable and grown again in pure culture; and
5. The cause of no other disease.

These criteria are relevant for explaining infectious diseases, but infectious diseases are only part of the focus of epidemiology. They do little to guide the discovery of the causes of chronic diseases, which are the predominant causes of death in contemporary developed countries.

An expanded set of criteria for determining causality was outlined in medical statistician Sir Austin Bradford Hill's President's Address to the Section of Occupational Medicine of the Royal Society of Medicine in 1965 (Hill, 1965). This seminal address identified nine criteria that are applicable to causality in epidemiologic research:

1. Strength of the association between exposure and disease. A strong association is less likely to be due to bias or random error.
2. Consistency in observing the association in multiple investigations. Ideally, the additional studies should be conducted by multiple investigators examining the potential association in various contexts of places, circumstances, and times.
3. The association is specific to particular persons, places, times, and/or health phenomena.
4. The exposure precedes in time the development of the disease. If this particular criterion is not met, there is no point in further investigating the possibility of causality.
5. Dose-response or biological gradient in linear relationships. The greater the exposure, the greater the risk of disease.
6. Plausibility. It is biologically possible, at least within the current knowledge of biology, that the exposure can cause the disease.
7. Coherence. The association makes sense given what is known about the biology and natural history of the particular disease.
8. Experiment. Causality is more plausible when the investigator is able to manipulate the exposure and observe the results.

9. *Analogy.* Causality is more plausible if an association between a similar exposure and/or a similar disease has already been established.

Some of these criteria are illustrated in the following landmark epidemiologic case study.

Example: Cigarette Smoking and Lung Cancer

Together with Richard Doll, Sir Austin Bradford Hill demonstrated that cigarette smoking causes lung cancer by finding evidence that meets many of Hill's criteria for causality (Doll & Hill, 1950, 1952, 1954, 1956, 1964). This series of studies grew out of clinical observations in the 1920s that cigarette smoking might cause lung cancer. It was not until the 1964 Surgeon General's Report on Smoking (U.S. Public Health Service, 1964) that enough evidence was collected to begin convincing the general public that cigarette smoking is strongly associated with morbidity and mortality. What follows is a brief summary of research evidence in the context of criteria for causality.

1. *Strength of the association between exposure and disease.* **Figure 2–4** presents measures of association between smoking and lung cancer. These measures indicate that smokers were nearly 10 times more likely than nonsmokers to have lung cancer and 18 times more likely to die from lung cancer. More than 90% of lung cancer deaths among smokers were attributable to smoking. These associations are undeniably strong.
2. *Consistency in observing the association in multiple investigations.* In the 1950s, Doll and Hill demonstrated a strong relationship between smoking and lung cancer in two large-scale and long-term studies (case-control and prospective cohort) of British physicians. A search of the medical literature in 2011 produced 13,559 articles about cigarette smoking and lung cancer. Results of a few of the most recent studies showed that exposure to secondhand smoke increases cotinine (an alkaloid found in tobacco) levels among nonsmokers (Baltar et al., 2011); cigarette smoking explains more than 50% of the difference in life expectancy at 50 years between U.S. immigrants (nonsmokers) and U.S. citizens (smokers) (Blue & Fenelon, 2011); and smokers have significantly elevated white blood cell counts compared to nonsmokers (Frost-Pineda et al., 2011).
3. *Specificity.* A recent case-control study of lung cancer showed that occupational exposure to sulfuric acid (a known carcinogen) in mist form

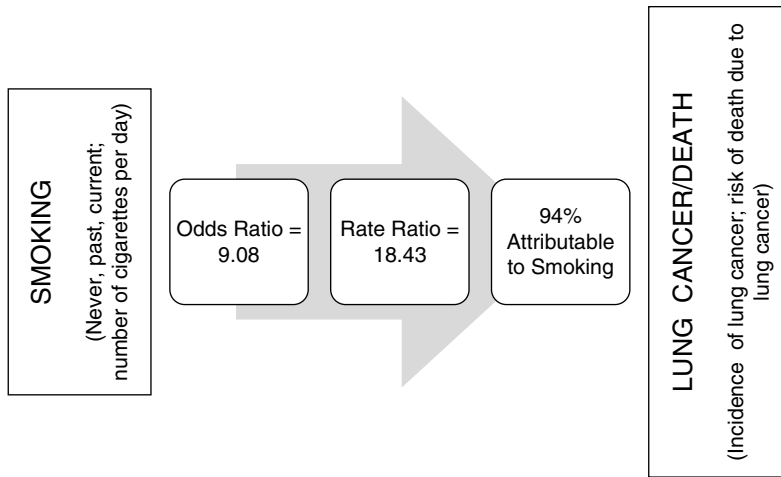


FIGURE 2–4 Strength of the Association Between Cigarette Smoking and Lung Cancer

Source: Data from Doll, R., & Hill, A. B. A study of the aetiology of carcinoma of the lung. *British Medical Journal*, 2(4797), 1271–1286, BMJ Publishing Group, © 1952; Data from Doll, R., & Hill, A. B. Mortality in relation to smoking: Ten years' observation on male British doctors. *British Medical Journal*, 1(5395), 1399–1410, BMJ Publishing Group, © 1964.

is associated with damage to the larynx, but not with lung cancer. The only demonstrated risk factor for lung cancer was cigarette smoking (Soskolne et al., 2011). These results offer some support that lung cancer is specific to cigarette exposure, but not to a similar inhaled carcinogen.

4. *The exposure precedes the development of the disease.* Doll and Hill's studies measured current or recent lung cancer and current or former cigarette use. The prospective cohort study helped to establish that a history of smoking precedes the onset of cancer. However, the temporal order is clouded by current smoking, and the long latency period between exposure and the onset of lung cancer can allow other potential causes to intervene.
5. *Dose-response.* Doll and Hill were able to provide very strong evidence for a dose-response relationship between cigarette smoking and lung cancer. **Figure 2–5** presents this evidence. These data show that the risk for lung cancer mortality increases linearly with the number of cigarettes smoked per day. Clearly, risk increases with exposure.

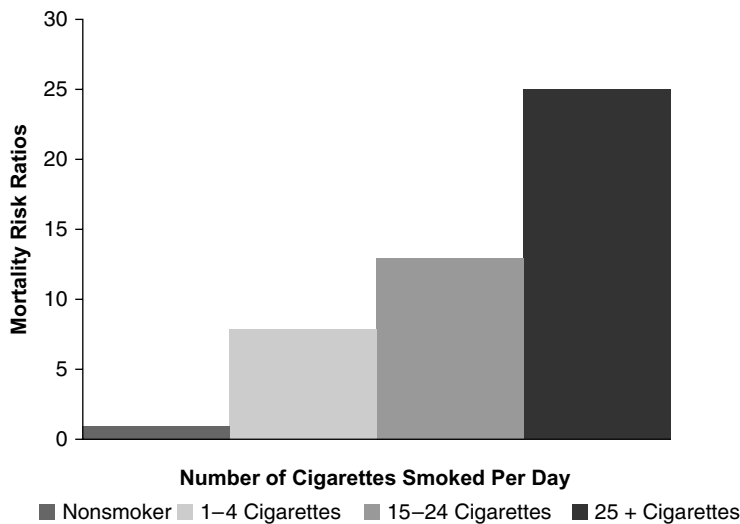


FIGURE 2–5 Relationship Between Number of Cigarettes Smoked and the Risk of Lung Cancer Deaths

Source: Data from Doll, R., & Peto R. Mortality in relation to smoking: 20 years of observation on male British doctors. *British Medical Journal*, 2(6051), 1525–1536. BMJ Publishing Group, © 1976.

Evaluation Research Goal

The most important goal of epidemiology as a field is to discover ways to prevent morbidity and mortality. Once a cause is identified, prevention efforts can be developed to eliminate or minimize the cause, thereby preventing or minimizing the health phenomena. However, even with a known and confirmed cause, the prevention program will not be effective if it does not affect the cause. The purpose of an evaluation study is to determine whether or not the program is efficacious (effective change for those receiving the treatment and not for those not receiving the treatment) in preventing the health outcome.

Ideally, the evaluation will use the traditional experimental research design with a pretest measuring the health phenomena before the intervention, a concurrent measure of the fidelity in implementing the intervention, then a posttest to measure any change in the health phenomena subsequent to the intervention. A strong design would include a control group that does not receive the intervention and random assignment of subjects into the experimental and control groups to minimize bias between groups. A strong design would also include a relatively long follow-up period to determine whether or not an effect is enduring. As seen

in the following case study, studies showing that an intervention is not effective are valuable in indicating that interventions other than the one evaluated should be implemented.

Example: Evaluation of Drug Abuse Resistance Education (DARE)

D.A.R.E. was founded in 1983 in Los Angeles and has proven so successful that it is now being implemented in 75 percent of our nation's school districts and in more than 43 countries around the world (DARE, 2012).

This school-based drug abuse prevention program was founded in 1983, and evaluation studies began in the mid-1980s. This review will focus on evaluations that used the essential features of the experimental design—control or comparison group, pretest–posttest measures, or posttest only with random assignment. These features help to minimize study bias (by randomly assigning students to the experimental or control group) and isolate effects that are due to the intervention (by comparing outcomes for those who did and did not participate in the program and measuring changes in the outcome from before to after the program).

Evaluations of the effect of DARE participation on students' drug use conducted in Illinois (Ennett, Rosenbaum, Flewelling, et al., 1994), Kentucky (Faine & Bohlander, 1989), North Carolina (Ringwalt, Ennett, & Holt, 1991), and South Carolina (Harmon, 1993) showed that DARE had negligible effects on participants' drug use. More recent studies indicated that a revised DARE had no impact on drug use among elementary school students in urban schools (Vincus, Ringwalt, Harris, et al., 2010), but a study of Tennessee schools did show that the program had an effect on preventing the initiation of cigarette use (Ahmed, Ahmed, Bennett, et al., 2002).

With such a relatively long tradition of DARE evaluation studies, it is possible to “study the studies” or “analyze the analysis” through the use of meta-analysis. Meta-analysis is a process used to summarize and even analyze the results of several evaluation studies (Sutton, Jones, Abrams, et al., 1999). Measures of intervention effects on the outcome of interest from individual studies are combined and weighted to calculate one mean effect size across all studies. In a meta-analysis of eight evaluation studies, Ennett, Tobler, Ringwalt, and colleagues (1994) calculated a mean effect size of only 0.06. Another study (West & O'Neal, 2004) analyzed 10 studies and reported an even smaller effect size of 0.01. Pan and Bai (2009) analyzed 20 studies and found a very small (within the context of relevant study factors such as sample sizes) overall effect of DARE on adolescent drug use.

DARE may be an effective program in some ways, but these evaluations show that it is likely not effective in preventing drug use among young people.

PRACTICAL LIMITATIONS OF A RESEARCH GOAL

One mistake even seasoned researchers sometimes make is to draw implications from their study results that go beyond their original research goal. A descriptive study can suggest, but not actually test the association between exposure and disease. An association study does not demonstrate causality. An evaluation study without the essential elements of an experimental design does not demonstrate the efficacy of an intervention. As a researcher and a consumer of research, particular attention should be paid to the connection of the research goal outlined in the introduction of the research article, thesis, capstone, and so on, and the implications of results drawn in the discussion section. Generalizations of results beyond the scope of the research goal have no value.

CONCLUSION

We still do not know much about the indisputable meaning of life, but we should have a better understanding of the different levels of research questions and the conclusions that can be drawn from them. Descriptive goals describe health phenomena and incidence and prevalence of morbidity and mortality in terms of the distribution of the disease among particular groups of people, in specific areas, and in focused periods of time. Association goals move to the next step of discovering additional factors that are related to the disease. Care should be taken to show that the exposure is truly associated with the outcome and the relationship is not confounded by other factors or exposures. The determination of true risk factors suggests hypotheses to be tested in causal studies to support or refute their role as the cause of disease. Ideally, causal studies identify factors that should be manipulated to prevent or intervene in the course of the disease. This is the ideal because many intervention programs are designed and implemented without knowledge of whether or not the focus of the program is even related to the outcome of interest. They are not “evidence-based.” Finally, evaluation studies indicate whether or not the intervention changes the cause, thereby preventing the disease.

Studies that address these goals should be designed specifically for the goal. They should be free, to the extent possible and practical, of bias and random error.

Finally, the implications drawn from study results must be limited to the scope of the research goal. A study designed and implemented according these criteria may someday tell us the meaning of life.

VOCABULARY

Analytic goal	Exposure	Reliability
Attributable risk	Evaluation	Risk difference
Causal research goal	Evidence-based	Risk factor
Confounding	Incidence	Risk ratio
Criteria for causality	Odds ratio	Scientific method
Descriptive goal	Outcome	Temporality
Dose-response	Person, place, time	Validity
Effect modification	Population attributable risk	
Efficacy	Prevalence	

STUDY QUESTIONS AND EXERCISES

1. A businessman returns home to San Diego from a trip to China. On the way home, he had a layover in Hawaii where he visited relatives. While in Hawaii, he started to feel ill. His symptoms included respiratory problems, fever, fatigue, and vertigo. When he returned home, he suffered a seizure and died. In 1 week's time, 4 of his family members in both Hawaii and San Diego were dead, as were more than 10 people in China. As an epidemiologist at the Centers for Disease Control and Prevention (CDC), you are assigned to get to the bottom of this outbreak—quickly.
 - a. What research focus should you use to begin your investigation?
 - b. What factors should you examine and enumerate?
2. Continuing the scenario from problem 1, you discover that this new illness is spread easily by airborne and hard-surface contacts. From two of the victims, laboratory personnel on your team successfully isolate and replicate a virus similar to the H1N1 (bird flu) virus.
 - a. What research focus should you use in your next study?
 - b. What factors should you examine and enumerate?
3. Continuing the scenario from problems 1 and 2:
 - a. How would you demonstrate, scientifically, that this new virus causes this new illness?
 - b. How would you address the key criteria for causality?

4. Continuing this example from problems 1–3, your team has developed a promising vaccine to prevent this new illness.
 - a. What research focus should you use to demonstrate the efficacy of the vaccine?
 - b. What factors should you examine and enumerate?
5. If you are preparing to conduct your own research study, what research focus is most important for your research idea? Why is this focus appropriate for your study? What practical problems do you anticipate with this focus at this point in your study?

REFERENCES

- Abbott, R. D., Brand, F. N., Kannel, W. B., & Castelli, W. P. (1988). Gout and coronary heart disease: The Framingham study. *Journal of Clinical Epidemiology*, 41(3), 237–242.
- Ahmed, N. U., Ahmed, N. S., Bennett, C. R., & Hinds, J. E. (2002). Impact of a drug abuse resistance education (D.A.R.E) program in preventing the initiation of cigarette smoking in fifth- and sixth-grade students. *Journal of the National Medical Association*, 94(4), 249–256.
- Baltar, V. T., Xun, W. W., Chuang, S. C., Relton, C., Ueland, P. M., Vollset, S. E., et al. (2011). Smoking, secondhand smoke, and cotinine levels in a subset of EPIC cohort. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 20(5), 869–875.
- Blue, L., & Fenelon, A. (2011). Explaining low mortality among US immigrants relative to native-born Americans: The role of smoking. *International Journal of Epidemiology*, 40(3), 786–793.
- Castelli, W. P. (1999). Cardiovascular disease: Pathogenesis, epidemiology, and risk among users of oral contraceptives who smoke. *American Journal of Obstetrics and Gynecology*, 180(6 Pt 2), S349–356.
- Doll, R., & Hill, A. B. (1950). Smoking and carcinoma of the lung; preliminary report. *British Medical Journal*, 2(4682), 739–748.
- Doll, R., & Hill, A. B. (1952). A study of the aetiology of carcinoma of the lung. *British Medical Journal*, 2(4797), 1271–1286.
- Doll, R., & Hill, A. B. (1954). The mortality of doctors in relation to their smoking habits: A preliminary report. *British Medical Journal*, 1(4877), 1451–1455.
- Doll, R., & Hill, A. B. (1956). Lung cancer and other causes of death in relation to smoking: A second report on the mortality of British doctors. *British Medical Journal*, 2(5001), 1071–1081.
- Doll, R., & Hill, A. B. (1964). Mortality in relation to smoking: Ten years' observations of British doctors. *British Medical Journal*, 1(5395), 1399–1410.
- Drug Abuse Resistance Education (DARE). (2012). *About DARE*. Retrieved February 10, 2012, from http://www.dare.com/home/about_dare.asp.

- Ennett, S. T., Rosenbaum, D. P., Flewelling, R. L., Bieler, G. S., Ringwalt, C. L., & Bailey, S. L. (1994). Long-term evaluation of drug abuse resistance education. *Addictive Behaviors, 19*(2), 113–125.
- Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). How effective is drug abuse resistance education? A meta-analysis of project DARE outcome evaluations. *American Journal of Public Health, 84*(9), 1394–1401.
- Evans, A. S. (1976). Causation and disease: The Henle-Koch postulates revisited. *The Yale Journal of Biology and Medicine, 49*(2), 175–195.
- Faine, J. R., & Bohlander, E. (1989). *DARE in Kentucky schools 1988–1989* (unpublished report). Bowling Green, KY: Western Kentucky University, Social Research Laboratory.
- Feinleib, M. (1983). The Framingham Study: Sample selection, follow-up and methods of analyses. *National Cancer Institute Monograph, 67*, 59–64.
- Fraenkel, L., Zhang, Y., Chaisson, C. E., Maricq, H. R., Evans, S. R., Brand, F., et al. (1999). Different factors influencing the expression of Raynaud's phenomenon in men and women. *Arthritis and Rheumatism, 42*(2), 306–310.
- Framingham Heart Study. (2012). *Omni cohorts*. Retrieved February 12, 2012, from www.framinghamheartstudy.org/participants/omni.html
- Friis, R. H., & Sellers, T. A. (2009). *Epidemiology for public health practice* (4th ed.). Sudbury, MA: Jones and Bartlett Publishers.
- Frost-Pineda, K., Liang, Q., Liu, J., Rimmer, L., Jin, Y., Feng, S., et al. (2011). Biomarkers of potential harm among adult smokers and nonsmokers in the total exposure study. *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco, 13*(3), 182–193.
- Garrison, R. J., Feinleib, M., Castelli, W. P., & McNamara, P. M. (1983). Cigarette smoking as a confounder of the relationship between relative weight and long-term mortality. The Framingham heart study. *The Journal of the American Medical Association, 249*(16), 2199–2203.
- Garrison, R. J., Kannel, W. B., Feinleib, M., Castelli, W. P., McNamara, P. M., & Padgett, S. J. (1978). Cigarette smoking and HDL cholesterol: The Framingham off-spring study. *Atherosclerosis, 30*(1), 17–25.
- Gordon, T., Castelli, W. P., Hjortland, M. C., Kannel, W. B., & Dawber, T. R. (1977). High density lipoprotein as a protective factor against coronary heart disease. The Framingham study. *The American Journal of Medicine, 62*(5), 707–714.
- Gordon, T., Kannel, W. B., Hjortland, M. C., & McNamara, P. M. (1978). Menopause and coronary heart disease. The Framingham study. *Annals of Internal Medicine, 89*(2), 157–161.
- Haider, A. W., Larson, M. G., O'Donnell, C. J., Evans, J. C., Wilson, P. W., & Levy, D. (1999). The association of chronic cough with the risk of myocardial infarction: The Framingham heart study. *The American Journal of Medicine, 106*(3), 279–284.
- Harmon, M. A. (1993). Reducing the risk of drug involvement among early adolescents: An evaluation in drug abuse resistance education (DARE). *Evaluation Review, 17*, 221.
- Haynes, S. G., Feinleib, M., Levine, S., Scotch, N., & Kannel, W. B. (1978). The relationship of psychosocial factors to coronary heart disease in the Framingham study. II. Prevalence of coronary heart disease. *American Journal of Epidemiology, 107*(5), 384–402.
- Hennekens, C. H., Buring, J. E., & Mayrent, S. L. (1987). *Epidemiology in medicine* (1st ed.). Boston: Little, Brown.

- Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58, 295–300.
- Kagan, A., Kannel, W. B., Dawber, T. R., & Revotskie, N. (1963). The coronary profile. *Annals of the New York Academy of Sciences*, 97, 883–894.
- Kannel, W. B. (1979). Possible hazards of oral contraceptive use. *Circulation*, 60(3), 490–491.
- Kannel, W. B., & Ellison, R. C. (1996). Alcohol and coronary heart disease: The evidence for a protective effect. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, 246(1–2), 59–76.
- Kannel, W. B., & McGee, D. L. (1979). Diabetes and cardiovascular disease. The Framingham study. *The Journal of the American Medical Association*, 241(19), 2035–2038.
- Murray, J., & Ryan-Krause, P. (2010). Obesity in children with Down syndrome: Background and recommendations for management. *Pediatric Nursing*, 36(6), 314–319.
- Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., & Flegal, K. M. (2010). Prevalence of high body mass index in US children and adolescents, 2007–2008. *The Journal of the American Medical Association*, 303(3), 242–249.
- Oleckno, W. A. (2002). *Essential epidemiology: Principles and applications*. Prospect Heights, IL: Waveland.
- Pan, W., & Bai, H. (2009). A multivariate approach to a meta-analytic review of the effectiveness of the D.A.R.E. program. *International Journal of Environmental Research and Public Health*, 6(1), 267–277.
- Reeder, B. A., Chen, Y., Macdonald, S. M., Angel, A., & Sweet, L. (1997). Regional and rural-urban differences in obesity in Canada. Canadian heart health surveys research group. *Canadian Medical Association Journal (Journal De l'Association Medicale Canadienne)*, 157(Suppl 1), S10–16.
- Rimmer, J. H., Yamaki, K., Lowry, B. M., Wang, E., & Vogel, L. C. (2010). Obesity and obesity-related secondary conditions in adolescents with intellectual/developmental disabilities. *Journal of Intellectual Disability Research*, 54(9), 787–794.
- Ringwalt, C., Ennett, S. T., & Holt, K. D. (1991). An outcome evaluation of Project DARE (Drug Abuse Resistance Education). *Health Education Research*, 6, 327–337.
- Rothman, K. J. (1986). *Modern epidemiology*. Boston: Little, Brown.
- Soskolne, C. L., Jhangri, G. S., Scott, H. M., Brenner, D. R., Siemiatycki, J., Lakhani, R., et al. (2011). A population-based case-control study of occupational exposure to acids and the risk of lung cancer: Evidence for specificity of association. *International Journal of Occupational and Environmental Health*, 17(1), 1–8.
- Susser, M. (1977). Judgement and causal inference: Criteria in epidemiologic studies. *American Journal of Epidemiology*, 105(1), 1–15.
- Sutton, A. J., Jones, D. R., Abrams, K. R., Sheldon, T. A., & Song, F. (1999). Systematic reviews and meta-analysis: A structured review of the methodological literature. *Journal of Health Services Research & Policy*, 4(1), 49–55.
- U.S. Public Health Service. Surgeon General's Advisory Committee on Smoking and Health. (1964). *Smoking and health; report of the advisory committee to the surgeon general of the public health service*. Washington, DC: Government Printing Office
- Vincus, A. A., Ringwalt, C., Harris, M. S., & Shamblen, S. R. (2010). A short-term, quasi-experimental evaluation of D.A.R.E.'s revised elementary school curriculum. *Journal of Drug Education*, 40(1), 37–49.

- West, S. L., & O'Neal, K. K. (2004). Project D.A.R.E. outcome effectiveness revisited. *American Journal of Public Health, 94*(6), 1027–1029.
- Wolf, P. A., Dawber, T. R., Thomas, H. E., Jr., & Kannel, W. B. (1978). Epidemiologic assessment of chronic atrial fibrillation and risk of stroke: The Framingham study. *Neurology, 28*(10), 973–977.
- Woods, J. L., Iuliano-Burns, S., & Walker, K. Z. (2011). Weight loss in elderly women in low-level care and its association with transfer to high-level care and mortality. *Clinical Interventions in Aging, 6*, 311–317.

