2

CHRONIC DISEASE EPIDEMIOLOGY

Robert Redwood, MD, MPH, Patrick L. Remington, MD, MPH, and Ross C. Brownson, PhD

Introduction

As noted in Chapter 1, chronic diseases are distinguished from other health problems by their multiple and interrelated causes often rooted in early life and related to multiple health factors, including individual behaviors, health care, social and economic factors, and the environment. As public health advances have reduced the burden from infectious diseases, chronic diseases have become the leading causes of death and disability in the United States. In addition, many people are surviving with chronic diseases, as most are controlled through health care or lifestyle interventions, rather than being cured or rapidly fatal.

The field of epidemiology has been invaluable in understanding the causes of chronic diseases and measuring the effects of interventions to prevent chronic diseases. Each of the chapters in this book, which describe chronic diseases, conditions, and risk factors, is organized to address these two basic questions:

1. What is the epidemiology of this problem? That is, what do we know about the burden of chronic diseases, their distribution in populations, and their causes?
2. What can be done to control this problem in the population? That is, what are effective strategies to prevent, detect, and treat this problem in populations?

The purpose of this chapter is to provide a brief overview of the methods that can be used to answer this first question, by summarizing the methods used to understand the epidemiology of chronic diseases, conditions, and risk factors.

Breast cancer and tobacco use will be used to illustrate the various methods used in chronic disease epidemiology and control. The next chapter (Chapter 3, Chronic Disease Surveillance) provides additional information about how public health surveillance can be used to measure and monitor chronic diseases and their risk factors in populations. And the following two chapters describe what can be done to prevent or control chronic diseases in communities (Chapter 4, Community-Based Interventions) and health care systems (Chapter 5, The Role of Health Care Systems in Chronic Disease Prevention and Control).

Definition of Epidemiology

Epidemiology is considered the "science of public health." The word comes from the Greek terms *epi* (upon), *demos* (people), and *logos* (study); thus, it is literally the study upon people. Traditionally, epidemiology was defined as the "study of the distribution and determinants of disease in populations." Over time, this definition broadened to include the study of other health states, such as injury, disability, risk factors, and health-related quality of life.

More than 40 years ago, William Foege coined the term "consequential epidemiology" to further broaden the definition of the term to include the use of epidemiology in disease control (Koplan and Thacker 2001). He stated that "the reason for collecting, analyzing, and disseminating information on a disease is to control that disease. Collection and analysis should not be allowed to consume resources if action does not follow" (Foege et al. 1976).

The Centers for Disease Control and Prevention (CDC) currently defines epidemiology as "the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems." The key words include the following (CDC 2016):

- Study: Epidemiology is a quantitative discipline based on principles of statistics and research methods.
- Distribution: Epidemiologists study the distribution of health events within groups in a population, characterizing health events in terms of person, place, and time. This type of epidemiology is referred to as "descriptive epidemiology."
- Determinants: Epidemiologists also search for determinants (i.e., causes or factors) that are associated with increased risk or probability of disease. This type of epidemiology, where we move from questions of

who, what, where, and when and start trying to answer how and why, is referred to as "analytical epidemiology."

- Health-related states: Whereas infectious diseases were clearly the focus of much of the early epidemiological work, the field is no longer limited in this way. Epidemiology as it is practiced today is applied to the whole spectrum of health-related events, which includes chronic diseases, conditions, and risk factors. Although not addressed in this chapter, some forward-thinking epidemiologists are further expanding the concept of health-related states, applying epidemiologic methods to broader social ills (e.g., drug addiction, firearm violence, homelessness).
- Populations: One of the most important distinguishing characteristics of epidemiology is that it deals with groups of people rather than with individual patients.
- Control: Finally, although epidemiology can be used simply as an analytical tool for studying diseases and their determinants, it can also play a more active role. Epidemiological data and methods steer public health decision-making and aid in developing and evaluating interventions to control and prevent health problems. This is the primary function of applied, field, or consequential epidemiology.

Brownson and Petitti (2006) defined "applied epidemiology" according to the intended purpose. In their view, the field can be defined on the basis of five core purposes: (1) the synthesis of the results of etiologic studies as input to practice-oriented policies; (2) the description of disease and risk factor patterns as information to set priorities; (3) the evaluation of public health programs, laws, and policies; (4) the measurement of the patterns and outcomes of health care; and (5) the communication of epidemiological findings effectively to health professionals and the public.

The Chronic Disease Continuum

For diseases of known infectious origin, such as AIDS, measles, and influenza, the presence of a single, known, necessary cause (e.g., the microorganism) helps to focus epidemiological research and intervention strategies. For injuries, the cause is often acute and leads to immediate health consequences, such as drinking and driving or a child drowning.

In contrast, the wide variety of chronic diseases lacks such unifying causal agents and often develops over the life course. Research has demonstrated that many chronic diseases have their origins early in the life course (Felitti et al. 1998). These early life experiences and exposures to social and economic factors increase the risk of unhealthy behaviors later in life. Unhealthy behaviors eventually lead to conditions such as hypertension, obesity, or high blood cholesterol. Finally, individuals with these conditions are at increased risk for developing chronic diseases, such as heart disease, cancer, or diabetes.

This model of a chronic disease continuum is shown in Figure 2-1 and has been used to organize this text, with different parts for chronic disease risk factors, chronic disease conditions, and chronic diseases. The chronic disease continuum also requires special methods for chronic disease epidemiology and control. Because the causal process is prolonged and typically complex, many modest influences, rather than a single predominant cause, contribute to the probability of developing disease. The prolonged duration of these diseases, which often includes a presymptomatic phase and subsequent development of chronic disease conditions, provides numerous overlapping opportunities for intervention. For most chronic diseases, the large number of modest risk factors and the diverse opportunities for

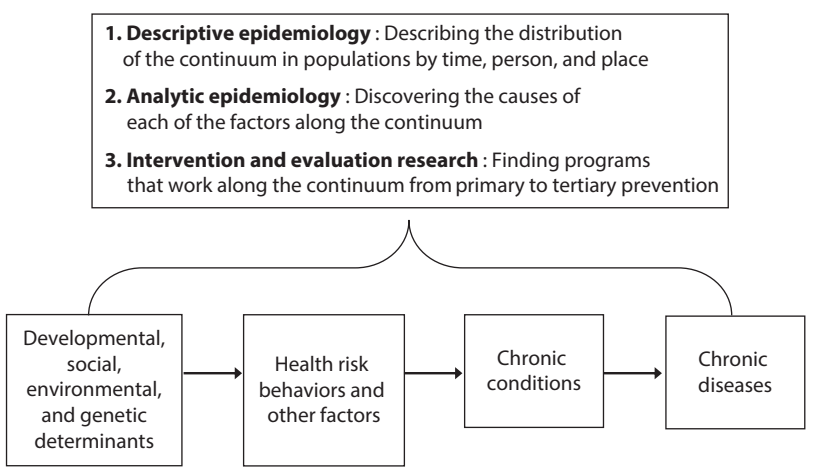


Figure 2-1. Methods Used in Chronic Disease Epidemiology, Prevention, and Control, Applied along the Chronic Disease Continuum

intervention make their control difficult and typically require a multifaceted approach.

Because disease is the end result of a continuum, it is sometimes difficult to determine whether "disease" is even present at all. For example, the large increase in mammography over the late 1980s and 1990s has led to a large increase in the incidence of a pathological lesion known as ductal carcinoma in situ (DCIS). Although DCIS is not invasive breast cancer and many women with DCIS will not develop invasive breast cancer, DCIS is often treated in the same way as invasive cancer—that is, by surgical excision. Many women with DCIS consider themselves to have "breast cancer." Similarly, many men with "prostate cancer" actually have a pathological lesion that would never have progressed to clinical disease. In general, as our ability to detect earlier and earlier stages of the disease process improves, the point at which disease truly begins becomes increasingly unclear (Fryback et al. 2006).

Another important consideration for chronic disease is the meaning of "control" or what it is we are trying to prevent. Chronic disease often affects the quality of life long before it affects the duration of life, if it affects duration at all. People with diabetes, for example, do have higher mortality, but only many years after diagnosis. In the interim, however, they may experience blindness, kidney failure, painful feet ulcers, leg amputation, or premature heart disease, all as a direct result of having diabetes. "Controlling" diabetes implies not only prolonging life, but also, as importantly, improving the quality of life while one is living with diabetes. Both disease-specific and general measures of quality of life have been developed and validated (McDowell and Newell 2006), and these measures are being used increasingly to describe the course ("natural history") of disease as well as the effect of various methods of control.

There is a need to pay careful attention to methodological principles in many areas of the study and control of chronic disease. The need for clarity of terminology and thought is enhanced by the advancements that have been made and the expansion of options for tackling chronic disease through public health measures. In this chapter, we provide some basic terms and concepts common in epidemiological literature on the etiology (i.e., causes), consequences, and control of chronic diseases. We also describe how epidemiology has been used to better understand the factors that increase (or decrease) a person's risk of progressing along the chronic disease continuum shown in Figure 2-1.

Building on this continuum, each chapter in Parts II through IV of this text has the following sections:

1. Significance: The first section of each chapter provides an overview of the "significance" of the disease, condition, or risk factor in populations. This section measures the burden to public health by estimating the number of people affected, rates, and economic costs.
2. Pathophysiology: The second section of each chapter describes the biology and pathophysiology of the disease, condition, or risk factor. This information is important to understand when one is interpreting epidemiological information or designing interventions.
3. Descriptive epidemiology: The third section of each chapter provides an overview of the "distribution" of the disease, condition, or risk factor in populations. This section uses "descriptive epidemiology" to identify high-risk groups (person), geographic variation (place), and secular trends (time).
4. Causes: The fourth section of each chapter provides an overview of the "causes" of the disease, condition, or risk factor in populations. This section uses "analytic epidemiology" to identify those factors that increase the risk of an individual developing a chronic disease, condition, or risk factor. Special techniques, described in the "Analytic Epidemiology" section, are used to go beyond finding associations to finding factors that are causally related to the outcome of interest.
5. Evidence-based interventions: The final section of each chapter provides an overview of the interventions that may be used to prevent or control the chronic disease, condition, or risk factor. These methods go beyond the use of epidemiology to determine the causes to finding programs that affect the disease course, along the continuum from primary prevention to secondary and tertiary prevention. The final section of this chapter describes methods that can be used to move from descriptive and analytic epidemiological research to developing effective interventions. Intervention methods are described in detail in Chapter 4.

The epidemiology of each chronic disease, condition, or risk factor is summarized in a figure in each chapter of Parts II through IV, as shown in Figure 2-2. This figure shows the "upstream causes" on the left side, the "high-risk groups" in the middle, and the "downstream consequences" on the right side. The next sections will provide a brief overview of the methods used to measure the

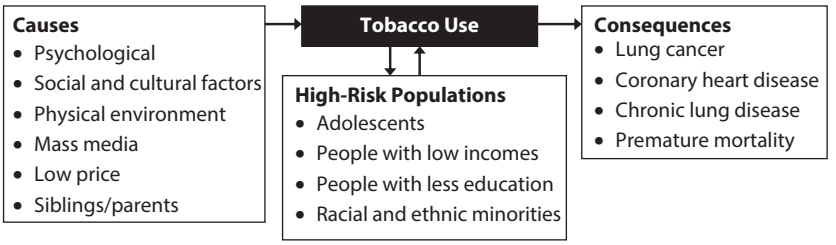


Figure 2-2. Tobacco Use: Causes, Consequences, and High-Risk Groups

burden, describe high-risk groups and trends in the population, and identify modifiable risk factors.

Descriptive Epidemiology

The first step in developing approaches to chronic disease control is to understand the nature and extent of the problem. Echoing the Institute of Medicine's (1988) landmark report on the future of public health, the director of the CDC (Frieden 2015) stressed the importance of epidemiology in carrying out public health assessment—one of the core functions of public health. Frieden and the Institute of Medicine report emphasize the governmental role in public health and affirm that every public health agency should regularly and systematically collect, analyze, interpret, and disseminate information on the health of the community, including statistics on health status, community health needs, and epidemiological and other studies of health problems.

The type of epidemiological studies used to support this assessment function varies, depending on the nature of the public health program. Descriptive epidemiology makes use of available data to examine the distribution of diseases in populations by time, person, or place. This type of information can be used to assess the burden of disease, identify high-risk groups, or monitor trends over time. The public and health policymakers frequently underestimate the health and economic burden from chronic diseases. Research has demonstrated that the public is more concerned about diseases and exposures that are unknown or out of one's individual control. For example, although social determinants of health (tobacco use, poor nutrition, and physical inactivity) account for nearly two thirds of preventable deaths in the United States (Mokdad et al. 2004; Galea et al. 2011), the public often perceives much greater

risk from AIDS, homicide, and environmental pollution (Tinker and Vaughan 2002; so-called involuntary risks).

In the first section of each chapter of Parts II through IV, the authors describe the burden of the chronic disease, condition, or risk factor in the population and how it is distributed by "person, place, and time." For example, in Chapter 16, the following questions are addressed for breast cancer:

1. Significance: What are the "downstream consequences" of this problem? This is addressed through measures of disease occurrence: incidence rate, cumulative incidence, and prevalence. Both incidence rate and cumulative incidence indicate the risk of newly acquiring the disease, whereas prevalence is the number of people who have the disease at a point in time. Mortality rates are often used when incidence rates are unavailable.
2. High-risk populations: What groups have the highest risk? This question is addressed by comparing the risk or incidence of disease among people within the population who have some characteristic (e.g., older age) with those who do not have the characteristic. To do this, we use measures of association—rate ratio—and difference measures.
3. Geographic distribution: How does the burden of disease in one area compare with that in other areas? In calculating statistics that quantify the disease burden in one area versus other areas, consider the potential for distortion attributable to such factors as differing age, sex, and race in your population. To avoid distortion, you may need to divide measures of disease occurrence into categories to make them age-, sex-, or race-specific or standardize (adjust) them to make the populations comparable.
4. Time trends: What are the trends in the disease over time? Monitoring trends over time can provide insight on the etiology of chronic diseases (e.g., increases in lung cancer followed increases in smoking), burden (e.g., increasing rates of obesity), and control (e.g., declines in breast cancer mortality). The methods used to answer these questions are described in more detail in the following sections.

Measuring the Burden of Disease

Several measures are used to quantify the magnitude of disease occurrence, each one valid for a slightly different purpose. The number or actual count of persons affected by a chronic disease, condition, or risk factor is often used as

the most fundamental measure of burden in the population. This measure is useful when one is assessing the need for health care or public health services as a direct measure of the burden on these systems. The actual burden of cancer is described in Chapter 16 (American Cancer Society 2015):

- Cancer is the leading cause of death among persons aged younger than 85 years and the second leading cause of death overall in the United States.
- Cancer accounted for an estimated 1.5 million new cases and 560,000 deaths in the United States in 2012.
- Breast cancer is the most common cancer type among women in the United States and the second leading cause of cancer death.
- The American Cancer Society expects 249,260 new cases of breast cancer and 40,890 breast cancer deaths in the United States in 2016.
- Worldwide, breast cancer is the second most common cancer and the fifth among cancer deaths.
- Breast cancer rarely occurs in men, with only 2,600 new cases and 440 deaths expected in the United States in 2016.

The actual number of people with a chronic disease or the number of deaths is an easy way to communicate burden to the general public and to policymakers. However, the absolute number of people affected is almost entirely dependent on the size of the population under consideration. Therefore, other measures must be used when one is making comparisons across populations and over time.

Calculating Rates in Populations

Rates must be calculated to compare populations of various sizes and characteristics. The principal rates used in chronic disease epidemiology are incidence and prevalence. Measures such as rate ratios and rate differences are then used to compare these different measures of risk across populations.

Incidence (or incidence rate) refers to the number of new cases over a defined period divided by the "person-time experience of the population"—that is, the number of persons multiplied by the period over which they were monitored; this is often called "person-years." However, in practice, the incidence rate is typically used to describe the number of new cases that develops in a year in a specified population. For example, if 500 women develop breast cancer in a

year in a population of 342,000 women, the incidence rate would be 146 cases of breast cancer per 100,000 population of women. For many chronic diseases, such as coronary heart disease and diabetes, mortality rates are calculated because incidence data are unavailable.

Cumulative incidence is defined as the probability or risk of developing a disease over a defined period. It ranges from 0 to 1 and indicates the probability that the disease will develop in a population that is monitored over a set period. For example, according to a 2012 report from the National Cancer Institute (Howlader et al. 2012), a woman's lifetime risk of developing breast cancer is 0.127, or about 1 in 8. Thus, we can estimate that a girl born today has a 12.7% chance of eventually being diagnosed with breast cancer, although such estimates ignore any competing causes of mortality and are based on the somewhat unrealistic assumption that the present incidence rates will persist over time.

Prevalence also is measured as a proportion—that is, existing cases of disease divided by total population—but the occurrence of disease is measured at a point in time rather than over some interval. At the time of a disease survey, prevalence is defined as the number of existing cases divided by the population count. The prevalence of disease is influenced by the incidence (more new cases yield more existent cases) and persistence of the disease (rapid recovery or rapid death reduces the number of affected individuals at any point in time). For breast cancer, for example, the prevalence greatly exceeds the annual incidence because most women diagnosed with breast cancer survive for at least five years.

Because it is influenced by survival and recovery, prevalence is less valuable than incidence for identifying etiologic factors. For assessing public health needs, however, prevalence may be exactly the measure of interest. For example, women diagnosed with breast cancer are at a higher risk of a second breast cancer. Therefore, a prevalence estimate of the number of breast cancer survivors in a given area may be useful in targeting limited public health resources to women in high-risk groups.

Comparing Rates across Populations

Many local, state, and federal health agencies now calculate incidence rates for various conditions. Likewise, mortality rates (which can be viewed as incidence rates of death) are also published, often for such geographic areas as

counties or cities. Public health officials have a natural interest in comparing incidence or mortality rates from other areas with their own to determine which areas have a greater problem.

Once rates have been calculated for various populations or population subgroups, these rates can be compared by using rate ratios or relative risks. For example, data from the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute provide breast cancer incidence rates for men and women. We will use a hypothetical example that approximates actual SEER data, in which the age-adjusted incidence rate for women is 120 per 100,000 person-years and the comparable incidence rate for men is 1.2 cases per 100,000 person-years. Having determined the incidence rate in each of the two groups we wish to compare, our next challenge is to determine how that comparison should be summarized.

The ratio of the incidence rate in one group to that in another is referred to as a rate ratio. Likewise, the ratio of the cumulative incidence or risk in two groups is termed the "risk ratio" or "relative risk." Considering the incidence rate in men as the reference, we can calculate the rate ratio for women compared with men as 120 divided by 1.2 equals 100. Thus, we can say that women have an incidence rate 100 times that of men: breast cancer is predominantly a disease of women, although it does occur infrequently in men.

One advantage of such ratio measures is the ease of intuitive understanding (i.e., disease occurrence is increased about 100-fold among the women). Also, this ratio is independent of the absolute incidence rates in the two groups and, therefore, is directly interpretable: there is a strong association between gender and breast cancer.

Other risk factors besides gender can be discussed in similar terms. For example, research has demonstrated that the risk of breast cancer increases as the age of first birth increases. The five-year risk of breast cancer among women who have their first child before the age of 20 years is 0.7%, compared with a risk of 1.3% for women who have their first child after the age of 30 years (Table 2-1). In this instance, the relative risk for developing breast cancer in the next five years is 1.86 times (1.3/0.7 = 1.86).

An alternative to the ratio measure is the rate difference, calculated by subtracting the rates from one another. For example, the rate difference between women who have their first child after the age of 30 years, compared with women who have their first child before the age of 20 years, is calculated as follows: 1.3% minus 0.7% equals 0.6%, or 0.6 more cases per

Table 2-1. Five-Year Breast Cancer Risk, Rate Ratios, and Rate Differences for Women Based on Age of First Pregnancy

Age at First Birth	Rate, %[a]	Rate Ratio	Rate Difference
<20	0.7	Reference group	Reference group
20–24	0.9	0.9/0.7=1.29	0.9−0.7=0.2%
25–29	1.1	1.1/0.7=1.57	1.1−0.7=0.4%
30+	1.3	1.3/0.7=1.86	1.3−0.7=0.6%
No births	1.1	1.1/0.7=1.57	1.1−0.7=0.4%

Source: NCI (2016b).
Note: Based on women aged 50 years with no family or personal history of breast cancer and age at menarche of 12 to 13 years.
[a]Rate=percentage who will develop breast cancer in the next 5 years.

100 women (or six cases per 1,000 women) over the next five years. This difference indicates how much, in absolute rather than relative terms, the risk differs depending on childbearing histories. The advantage of this measure is that the actual amount by which the disease has increased in one group as opposed to another has public health importance beyond the ratio of the two rates.

A doubling or even tripling of rate of a disease (e.g., the incidence or mortality rate) may not indicate an important public health problem if the baseline rate is extremely low. That is, the increment in disease burden from doubling a very small number would be very small. The difference in incidence rates, however, provides direct information about the public health effects of a particular exposure. A large difference indicates an important problem, regardless of the size of the baseline rate.

The difference between rate ratios and rate differences is shown in Table 2-2. If one compares smokers with nonsmokers, the relative risk is much greater for developing lung cancer than for coronary heart disease. However, the rate difference for heart disease (the rate in smokers minus the rate in nonsmokers) is actually greater than the rate difference for lung cancer because the baseline mortality rate of heart disease is so much greater than that for lung cancer. The public health impact attributable to increasing heart disease risk twofold is similar to the impact of increasing lung cancer risk tenfold. Both ratio and difference measures contribute to our understanding of the effect of an exposure on disease occurrence, and both measures should be examined when the data permit.

Table 2-2. Smoking-Related Rate Ratio and Rate Difference Estimates for Coronary Heart Disease and Lung Cancer among Women

	Mortality Rate			
Disease	Smokers (a)	Nonsmokers (b)	Rate Ratio (a/b)	Rate Difference (a−b)
Lung cancer	131	11	11.9	120
Coronary heart disease	275	153	1.8	122

Source: Adapted from Thun et al. (1997).

Finally, rates can be used to identify subgroups at need for program targeting. For example, identifying factors associated with compliance with cancer screening (Boehm et al. 2013; Shelby et al. 2012) helps program planners and managers to select subpopulations at greater risk for negative outcomes. Also, identifying these associations simultaneously with that of health-related behaviors (Liang et al. 2006; Coughlin et al. 2007; Ward et al. 2014) and other screening practices provides managers with practical information on whom and how to offer joint interventions.

Controlling for Differences in Age When Comparing Populations

Perhaps the most important issue to consider when one is comparing rates across populations is the potential that the two populations differ in average age. The risk of most chronic diseases increases dramatically with increasing age. For example, a retirement community may have a higher rate of breast cancer because they have a greater proportion of older women. Thus, this higher rate might not be attributable to some risk factors (e.g., genetics, childbearing practices). How then could we compare the rates in one area with those in another area, accounting for the differing ages, to determine whether some other factors are influencing the rate of breast cancer?

Two approaches are possible. First, age-specific rates can be calculated. That is, incidence rates may be given only for people in a specific age range. If the incidence rate of breast cancer for women aged 50 through 59 years is much greater in one area than in another, the difference could not have been caused by differences in age distributions between the two groups. A drawback with such calculations is that the problem of precision may be worsened: fewer

cases are diagnosed (therefore, the numerator for the incidence rate is smaller) in a narrow age range than in the entire population. Such age-specific calculations typically require cases diagnosed over longer periods or from larger populations.

A second way to ease comparison of incidence rates is to adjust the rate to a standard population. This, in effect, combines many age-specific rates into a single age-adjusted rate. For example, the breast cancer incidence rate given in the previous example (110 cases per 100,000 women) is age-adjusted to the 1970 U.S. standard population. This means that statistical adjustments were made to the initial calculations to provide the rate that would be expected in the area's population if it had the same age distribution as did the U.S. population in 1970. Rates adjusted to the same standard population can be compared directly.

Small Area Analyses

Once differences in the ages of populations have been taken into account (either by comparing age-specific rates or through age adjustment), the problem of precision must be addressed by considering the statistical precision of the incidence or mortality rate. When rates are calculated for small areas such as a county, the numerator of the rate will usually be small. When a numerator is smaller than 20, rates are "unstable" and may vary by chance alone. If only a few cases are detected a year early or a year late, the incidence rate calculated for a particular year may appear much higher or much lower than it is on average over a longer period.

For example, if the incidence rate for breast cancer is 110 new cases per 100,000 women, a county with 25,000 women would be expected to have about 27 new cases each year (27/25,000 = 110 cases per 100,000). If only four cases were diagnosed too late to be counted for a particular year, the rate would decrease to 23 cases that year (23/25,000 = 92 cases/100,000), yielding an incidence rate ratio of 92 divided by 110 or 0.84. If only six cases from the next year were diagnosed a bit early, the rate would appear to increase dramatically to 33 cases per year (33/25,000 = 132 cases/100,000 women), for an incidence rate ratio of 1.20 and an apparent 20% increase in breast cancer incidence.

Thus, rates calculated from cases diagnosed in a single year from an area with a small population are subject to a large amount of variation from year to year and are said to be imprecise. For this reason, one should calculate rates

from cases diagnosed over several years to increase the number of cases in the numerator of the incidence rate, thereby increasing the precision of the calculated rate.

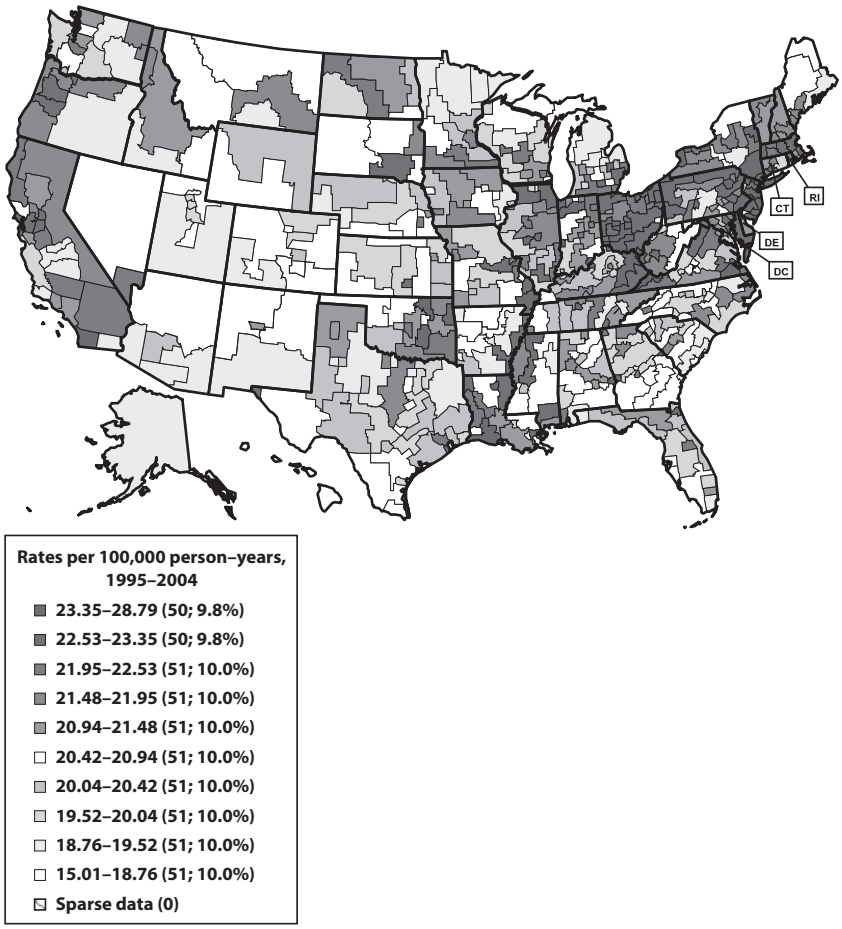
The use of these methods can be seen in the maps of breast cancer mortality for white women and lung cancer mortality for white men for the United States (produced by the authors; NCI 2016a). These maps (Figures 2-3 and 2-4) compare the breast cancer mortality rates for white women and lung cancer mortality for white men for 1995 to 2004. First, notice that rates are presented by grouping counties into "state economic areas" and that many years of data were combined to produce more reliable rates. In addition, these rates are age-adjusted (to the 1970 U.S. population) to account for differences in the ages of the populations, not only among counties but also over time.

These maps provide insight into the epidemiology of breast and lung cancer. First, the mortality rates for breast cancer are higher in the northeastern United States. Research has shown that differences in cancer incidence and survival explain some, but not all, of these geographic differences (Kohler et al. 2015). Considerable interest remains in the possibility that environmental exposures contribute to these differences (Reynolds et al. 2005). In contrast, the reasons for the higher rates of lung cancer mortality among white men are almost entirely due to higher rates of smoking among men living in the southeastern United States.

Analytic Epidemiology

The second major function for epidemiology in chronic disease prevention and control is to understand the "determinants" (i.e., causes) of chronic diseases—with a focus on those determinants that can be modified. In the second section of each of the chapters in Parts II through IV, authors describe the modifiable factors for each of the major chronic diseases, conditions, and, when possible, chronic disease risk factors.

One of the most important and challenging roles for epidemiology is to differentiate between factors that are simply *associated* with chronic diseases and those factors that actually *cause* those chronic diseases. Criteria have been established to help epidemiologists determine which associations are truly causal and which are not, such as the Bradford Hill criteria (Hill 1965). Experimental evidence provides the strongest evidence for a causal association, but it is most useful in examining the effects of drugs and clinical interventions,

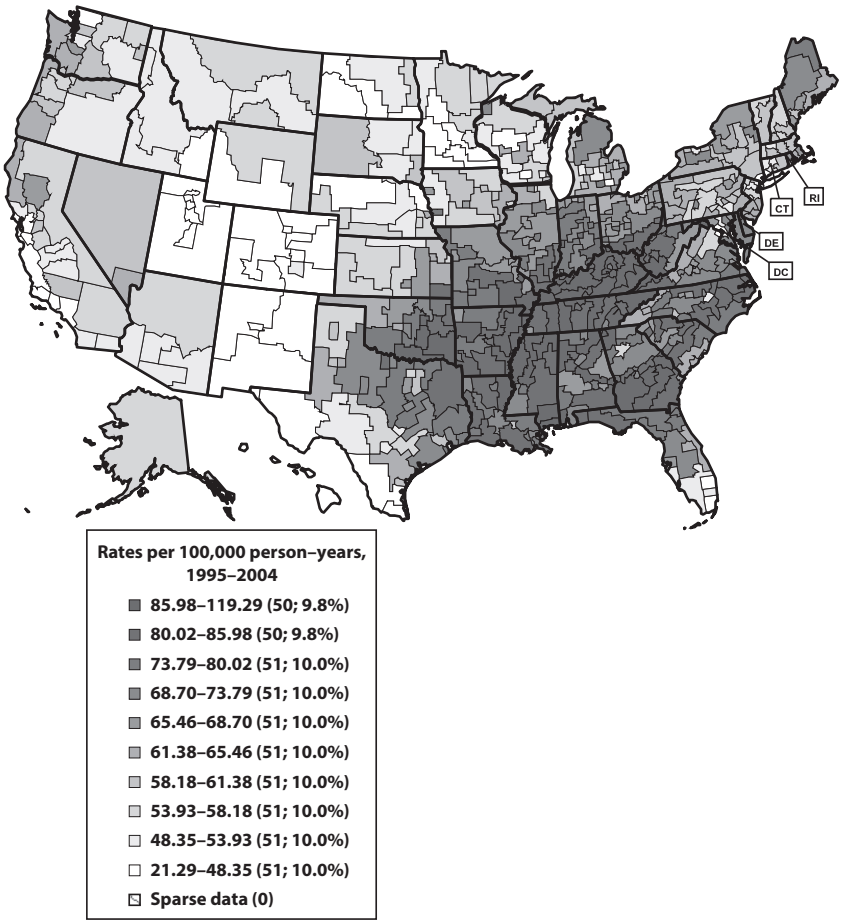


Source: Maps were calculated on the Cancer Mortality Maps website (NCI 2016a).
Note: Rates are age-adjusted to the 1970 U.S. population.

Figure 2-3. Breast Cancer Mortality Rates for White Women by State Economic Area, United States, 1995–2004

rather than the complex associations between environmental, social, or lifestyle factors and chronic diseases. Ultimately, subjective judgments must be made.

Epidemiology can not only be used to identify modifiable risk factors for chronic diseases, but it can also be used to estimate the morbidity and mortality that might be prevented by interventions. Although interventions such as education and screening programs often seem valuable, many have little



Source: Maps were calculated on the Cancer Mortality Maps website (NCI 2016a).
Note: Rates are age-adjusted to the 1970 U.S. population.

Figure 2-4. Lung Cancer Mortality Rates for White Men by State Economic Areas, United States, 1995–2004

effect on the number of people with the disease. Therefore, we need to evaluate the effects of intervention programs and use those results to design new and better programs for controlling chronic diseases. Even when there are strong theoretical reasons to expect benefit, a determination is needed that the intervention reduces mortality or improves quality of life when applied in a real setting. Every intervention, even those that reduce the burden of disease,

has a "downside," both through undesired consequences and by consuming resources. Calculating the proportion of a disease that is attributable to a particular risk factor may help to quantify the effect of its reduction or elimination (i.e., the population-attributable risk).

Study Designs in Chronic Disease Epidemiology

Experimental Studies

Randomized controlled trials (RCTs) are considered the most scientifically rigorous type of epidemiological study. In an RCT, participants are randomly assigned to either receive or not receive a preventive or therapeutic procedure, such as a clinical smoking cessation intervention or a new drug. The disease course or mortality patterns are then observed over time to assess the effectiveness of the preventive or therapeutic procedure.

The advantage of using a clinical trial, when possible, was best demonstrated by an RCT of the health effects of estrogen replacement therapy among women, called the Women's Health Initiative. This study was halted in 2002, when, contrary to its original hypothesis, the estrogen replacement therapy actually increased risk of breast cancer and heart disease (Writing Group for the Women's Health Initiative Investigators 2002). These results were definitive and conflicted with results that had been obtained from observational research decades before.

In practice, however, RCTs are either impossible or impractical when one is studying many interventions, especially those based outside clinical settings and in the community focusing on policy change. For example, it may be impractical to design a study to examine the effects of a community-wide educational intervention encouraging women to get a mammogram. In these situations, interventions can be tested in "quasi-experimental studies," in which a program (e.g., education program, screening program, new treatment regimen) is systematically offered to a population and the effect on health is measured. One key objective of such a study is to draw specific conclusions about the intervention. If the health of the population receiving the intervention improves, the investigators must demonstrate that a similar population that did not receive the intervention did not improve, at least to the same extent. This implies that to clearly interpret the results, such studies require comparison groups, which, unfortunately, are often lacking.

Comparison groups vary in their appropriateness for disease intervention studies. Least convincing are comparisons with national data or populations in other studies. Previous data on the same population are most appropriate if the disease can be shown to have been stable for a long period and if there are no other reasons for a change in disease incidence. The best control groups are those that can be shown to be similar to the intervention population and from which disease information is collected concurrently.

Despite certain limitations, quasi-experimental studies have shown that low-cost interventions can significantly improve the population's health. In one such study, low-cost community-based information campaigns were conducted in four low-income rural counties to encourage women to get routine breast and cervical cancer screening examinations. The rates of screening from baseline to follow-up were significantly higher among women living in these counties, compared with the trends among women living in four similar rural counties in Wisconsin (Eaker et al. 2001; Jaros et al. 2001). The strengths and limitations of the different types of experimental studies are summarized in Table 2-3.

Observational Studies

Randomized controlled trials or quasi-experimental interventions are either impossible or impractical when one is studying many of the causes of chronic diseases in the population. In some cases, it would be unethical or impossible to randomly assign the exposure (e.g., asbestos exposure, cigarette smoking, hypertension). In other cases, it is impractical to design a study to examine the effects of a randomly assigned exposure (e.g., effects of exercise during adolescence on the risk of postmenopausal breast cancer). In these instances, investigators must use observational methods. In observational studies, in contrast to RCTs (which are considered experimental studies), the risk factor or disease process is allowed to take its course without intervention from the researcher.

The observational study that is closest to an experiment is the prospective cohort design. In this approach, exposed and unexposed participants are identified and then observed over time for the development of disease. Unlike a true experiment, however, the exposures are observed rather than randomly assigned. Exposures are implicitly "assigned" on the basis of physician recommendations (e.g., x-ray use), genetic heritage (e.g., family history of breast cancer), or individual behavior (e.g., dietary fat intake). Having obtained measures of disease occurrence (incidence rates or cumulative incidence) for both exposed and unexposed groups, the influence of

Table 2-3. Summary of Strengths and Limitations of Various Study Designs

Study Type	Strengths	Limitations
Experimental studies		
Randomized clinical trial	Controls for bias by random assignment	High cost Not practical for many exposures (e.g., lifestyle, environmental, social/economic) Not practical for long latency periods
Randomized community trial	Can examine population-wide exposures Multicomponent interventions may be more effective	Very high cost Often involves small number of study groups
Quasi-experimental studies	Can be used to study real-world program and policy interventions Can use multiple comparison groups, repeated baseline measures to strengthen design	Potential for bias in comparison groups Lack of control of confounding factors
Observational studies		
Prospective cohort	Opportunity to measure risk factors before disease occurs Can study multiple disease outcomes Can yield incidence rates as well as relative risk estimates	Often expensive Requires large number of participants Requires long follow-up period
Case–control	Useful for rare diseases Relatively inexpensive Relatively quick results	Possible bias in measuring risk factors after disease has occurred Possible bias in selecting control group Identified cases may not represent all cases

exposure can be quantified as either the ratio of exposed to unexposed (the relative risk) or the difference between exposed and unexposed (the rate difference), as discussed earlier.

The primary advantage of a true prospective study over other observational designs is the opportunity to actively and intensively measure the exposures of interest before the period of disease induction. For example, instead of having participants recall dietary fat intake over many years, intake could be calculated more accurately if periodic direct measurements were made. For rare diseases and those with a prolonged period between exposure and the

manifestation of adverse effects (typical of chronic diseases), however, true prospective studies must be extremely large and prolonged and, therefore, are very expensive.

One way to overcome these problems is to study exposures that occurred in the past and to monitor disease either up to the present or into the future. The ability to conduct these historical (or retrospective) cohort studies depends on the availability of exposure records that can be linked to disease outcomes. For example, in a study of the effect of chest x-ray and the subsequent rate of breast cancer occurrence, Andrieu et al. (2006) were able to identify a roster of exposed and unexposed participants from an earlier study of women with a gene that increases the risk of breast cancer (*BRCA1* and *BRCA2*). This study demonstrated that women who had a chest x-ray were 1.5 times more likely to develop breast cancer compared with women who had not had a previous chest x-ray.

A case–control study is another type of observational study that can be used to investigate the causes of chronic diseases. Case–control studies study participants on the basis of their health outcomes (e.g., disease); this design intentionally oversamples persons who have developed the disease of interest. In case–control studies, researchers identify a sample of people with the disease, or case patients (often, all available case patients), and a sample of people without the disease, or control participants, from the same population that yielded the case patients.

The historical exposures that may have influenced disease risk are ascertained for all participants, and the frequency of exposure among case patients is compared with that among control participants. Such studies cannot yield incidence rates or cumulative incidence because the population at risk is not comprehensively defined; the control participants are typically a sample of that source population, but the sampling fraction is unknown. An estimate of the ratio of incidence rates or risks can be obtained, however, by calculating the ratio of the odds of exposure among cases (the number exposed divided by the number not exposed) to the odds of exposure among controls, referred to as the odds ratio. The odds ratio is intended to approximate the relative risk. No estimate of the risk difference can be obtained from a case–control study.

The obvious advantages of case–control studies are the relative speed with which they can be conducted, because the latent period is all in the past, and the small number of participants who have to be enrolled. In contrast to a cohort study, in which many participants who never develop the disease must

be monitored, a case–control study includes only a small but adequate fraction of the nondiseased population. The principal weakness is the study's vulnerability to some forms of bias that arise from the fact that the disease has usually occurred before risk factor information was ascertained. The strengths and limitations of prospective and case–control studies are summarized in Table 2-3.

To evaluate the potential causes of breast cancer mortality, the National Cancer Institute has supported an ongoing case–control study of breast cancer among residents of Wisconsin, Massachusetts, and New Hampshire since 1988 (Sprague et al. 2015). These studies compare the exposures (e.g., physical activity patterns during adolescence, childbearing history) of women diagnosed with breast cancer (i.e., cases) with the exposure histories of women of the same age, who have not had breast cancer (i.e., controls). These studies have led to the identification of potentially modifiable risk factors for breast cancer, including a long menstrual history (early menarche and late menopause), having no children or not having breastfed, previous chest irradiation, postmenopausal hormone therapy use, obesity (particularly postmenopausal), and moderate to heavy alcohol intake.

How Do We Measure Associations between Exposures and Outcomes?

The "relative risk" is the primary measure used to determine whether there is an association between an exposure (e.g., cigarette smoking) and a health outcome (e.g., lung cancer). As described previously, the ratio of the risk in two groups is termed the "relative risk." Associations can be assessed anywhere along the continuum shown in Figure 2-1. For example, the risk of lung cancer mortality among smokers is 131 deaths per 100,000 persons who smoke, compared with only 11 deaths per 100,000 persons who do not smoke—yielding a relative risk of 11.9.

Similarly, studies have shown that the risk of becoming a smoker is higher among children who view smoking in the movies versus children who do not view smoking in the movies (Charlesworth and Glantz 2005). For example, if 34% of children who view smoking become smokers, compared with only 20% of children who do not view smoking in the movies, the relative risk of smoking would be 34% divided by 20% equals 1.7. If this association is causal, one could say that children who view smoking in the movies are 1.7 times more

likely to become smokers, compared with children who do not view smoking in the movies.

How Do We Evaluate Whether the Study Results Are Valid?

Despite the increasing sophistication of research studies, uncertainty continues to exist in our understanding of the causes and consequences of chronic diseases. As most chronic disease epidemiology studies use observational methods, it is possible that errors in measurement or in selection of study participants lead to spurious results. For example, there may be uncertainty in epidemiological studies about a measure of association, such as relative risk. If we obtain an odds ratio or relative risk of 2.0, are we fully confident that becoming exposed will truly double the risk of disease or that removing exposure will cut the risk in half?

Considerable expertise is needed to address uncertainty in research and to determine the quality of a research study. Riegelman uses an organizing framework to evaluate whether results from a research study are valid (Riegelman 2013). Two of the most important factors to consider include "confounding" and "bias."

Confounding

One reason for uncertainty is the possibility for confounding, in which the influence of an exposure of interest is mixed with the effect of another. This arises when a risk factor for the disease of interest is also associated with the exposure of interest. For example, people who drink alcohol are also more likely to drink coffee. When a study is conducted to examine alcohol use and breast cancer, the possible confounding role of coffee and caffeine must be taken into consideration. The estimated effect of alcohol on breast cancer will be confounded by caffeine intake if (1) caffeine and alcohol use are correlated and (2) caffeine use independently influences the risk of breast cancer.

In an experiment or RCT, these potential confounders can be balanced among the study groups through the design of the study and the random manner in which the exposure is assigned. Conversely, in an observational study, potential confounders must be measured and adjusted statistically. The strategy involves creating groups that are similar with respect to the potential confounder and examining the impact of the exposure of interest within each of those groups. For example, we could measure caffeine consumption and create

strata of nonconsumers, low caffeine consumers, and high caffeine consumers and assess the role of alcohol use on breast cancer within each of those strata. As long as the potential confounder can be measured, the adjustments will be effective; however, some potential confounders, such as psychological stress, health consciousness, or dietary intake, may be difficult to measure, or we simply may be unaware of the risk factors that should be considered for adjustment.

Selection and Information Bias

A different source error comes from bias regarding how the participants enter the study or how information is collected from study participants. A faulty sampling mechanism, caused by such problems as nonresponse or refusal to participate, could produce a sample that has a higher or lower disease risk. Note that the only kind of sampling distortion that matters is the selection that influences disease risk. Similarly, in a case–control study, the selected case patients should reflect the exposure distribution of all case patients of interest, and the selected control participants should reflect exposure in the overall population that produced the cases. The potential for a poorly constituted control group is a major threat to the validity of case–control studies. For example, when we choose control participants from a hospital, health problems that led to their hospitalization may be associated with the exposure of interest. Similarly, when we choose control participants by telephone screening, omitting households without telephones could introduce a bias.

Another category of bias that can occur in epidemiological studies is the result of errors in classification of exposure or disease; this is referred to as information bias. Although efforts should be made to minimize such bias, errors in classification of exposure are unavoidable. Past exposures such as dietary intake, alcohol use, or physical activity are impossible to measure perfectly, even if we know what aspect of such exposures was the most relevant to disease causation. In many instances, the errors in exposure classification can be assumed to be similar for those who do and do not develop disease. This situation, referred to as nondifferential misclassification of exposure, results in a predictable bias in which the measure of association (such as odds ratio or relative risk) will be biased toward the null value of no association (1.0 for ratio measures, 0 for difference measures). This means that virtually all reported associations between exposure and disease will be diluted to some degree or missed entirely.

When the patterns of misclassification are different for the study groups, this is referred to as differential misclassification. This can occur when exposure is classified differently for diseased and nondiseased participants or when disease is classified differently for exposed and unexposed participants. Now the distortion in the measure of association can be in either direction (exaggerated or understated), depending on the precise pattern of error. A particular worry in case–control studies is the possibility of recall bias, which is a particular type of differential misclassification. Recall bias exists when the recall of exposure information is different for case patients than for control participants, presumably because the illness experience of the case patients has in some way altered their memory or reporting of past events. Intuitively, one might expect case patients to overreport exposures that did not occur in an effort to explain their illness. Also, studies suggest that case patients may report accurately (presumably because their memory search is more thorough), whereas control participants tend to underreport past exposures. In either situation, the reported exposures of case patients are artificially greater than the reported exposures of control participants, and the relative risk is falsely elevated.

Information bias was thought to be responsible for the apparent association between breast cancer and abortion. Case–control studies demonstrated that women with breast cancer were more likely to report having had an abortion in the past, compared with similar women who did not have breast cancer. In contrast, prospective studies comparing the future risk of breast cancer among women who had an abortion with women who had not had an abortion did not find an association. Researchers suspect that information bias in the case–control studies led to a spurious association, as women may underreport having had an abortion. This underreporting may be more common among controls, as women diagnosed with breast cancer may have provided a more accurate history (Rookus and van Leeuwen 1996). A review of the literature a decade later confirmed that researchers have not found a cause-and-effect relationship between abortion and breast cancer (Kitchen et al. 2005).

How Do We Assess Whether Associations between Potential Etiologic Factors and Disease Are Causal?

Any intervention program or public health action is based on the presumption that the associations found in epidemiological studies are causal rather than

arising through bias or for some other spurious reason. Unfortunately, in most instances in observational epidemiology, such as the research showing an association between viewing smoking in the movies and becoming a smoker, there is no opportunity to absolutely prove that an association is causal. Nonetheless, some principles are helpful when one must make this judgment.

The Bradford Hill criteria (Hill 1965) are often cited as a checklist for causality in epidemiological studies. These criteria have value but only as general guidelines. Most of Hill's nine criteria relate to particular cases of refuting biases or drawing on nonepidemiological evidence.

1. Strength of association: Stronger associations are less likely to be the result of some subtle confounding or bias, presuming that major distorting influences would be more readily recognized than small ones.
2. Consistency of association: The association is observed across diverse populations and circumstances, making a particular bias unlikely to explain a series of such observations.
3. Specificity of association: The exposure causes one rather than many diseases, and the disease is associated with one rather than many exposures, suggesting that the association is not the result of bias. This is the weakest of the criteria for chronic diseases and might as well be eliminated because we now know that many, perhaps most, exposures that influence one health outcome affect others (e.g., tobacco, radiation, diet) and that virtually all diseases have multiple causes.
4. Temporality: The exposure must precede the disease. This is the only absolute criterion for causality. Prospective study design helps establish the case for temporality.
5. Biological gradient: A dose–response curve, in which the risk of disease increases with increasing exposure, indicates that an association probably is not the result of a confounder or other bias. This criterion is generally valid, but the absence of a perfect dose–response pattern does not negate the possibility of a causal explanation because true thresholds or ceilings of effects may exist. Conversely, the presence of a dose–response gradient may be the result of a strong confounder that closely tracks the exposure.
6. Plausibility: Evidence from other disciplines suggests that the agent is biologically capable of influencing the disease. This is useful supportive evidence when it is available, but the lack of advancements in the other

biological sciences should not be used to negate an epidemiological observation.

7. Coherence: The evidence should not be contradictory to the known biology and natural history of the disease (similar to the plausibility criterion).
8. Experimental evidence: When attainable, experimental evidence for causality—obtained by removing or randomly assigning exposure—is very strong because both known and unknown confounders are controlled when exposure is randomly allocated.
9. Analogy: When other similar agents have been established as causes of disease, then the credibility of theories regarding a new disease operating in a similar manner is enhanced. This is the epidemiological counterpart to plausibility; however, the supportive evidence comes from other areas of epidemiology rather than from other disciplines.

In practice, the establishment of evidence for causality is largely through the elimination of noncausal explanations for an observed association. Consider, for example, the evidence that alcohol use may increase the risk of breast cancer. A series of further studies might confirm that this relationship is valid and not a result of confounding or other biases such as detection bias (in which disease is more thoroughly diagnosed among alcohol users) or nonresponse bias. By whittling away alternative explanations, the hypothesis that asserts that alcohol use causes breast cancer becomes increasingly credible. It is the job of critics to propose and test noncausal explanations, so that when the association has withstood a series of such challenges, the case for causality is strengthened.

The danger of formalizing the process of declaring causality on the basis of a checklist or any other mechanistic process is that it can only lead to endless debates about the degree of certainty and can impede needed public health actions. Those who argue that causality must be established with absolute certainty before interventions can begin fail to appreciate that their two alternatives—action and inaction—each have risks and benefits. Decisions must therefore be based on evidence that exposure causes diseases and must take into account the costs of intervention, the potential for the intervention to produce adverse side effects, and the potential costs of failing to act.

For example, the tobacco companies have argued, until recently, that the association between smoking and disease is uncertain. In a technical sense, we will always have some degree of uncertainty, especially with no definitive

data from large numbers of participants randomly assigned to be smokers and nonsmokers. We have no doubt, however, that the evidence indicates a need to intervene because smoking has no clear health benefits and scientists have exhausted all reasonable noncausal explanations for the strong associations observed between smoking and a number of diseases. Nonetheless, to the extent that tobacco companies continue to argue that the evidence for causality is not definitive, they create enough controversy to distract some policymakers from supporting needed interventions. Establishing causality is an important goal for epidemiological research, but absolute proof is not needed to justify action.

How Much Morbidity and Mortality Might Be Prevented by Interventions?

In each chapter in Parts II through IV, authors report the relative risk (discussed previously) and the "population-attributable risk." The population-attributable risk is particularly useful in evaluating the potential benefits of intervention. When presented with an array of potential causal factors for disease, we need to evaluate how much might be gained by reducing or eliminating each of the hazards. Relative risk estimates indicate how strongly exposure and disease are associated, but this measure does not indicate directly the benefits that could be gained through modifying the exposure.

Attributable Risk

The attributable risk is a measure of how much of the disease burden could be eliminated if the exposure were eliminated. The attributable risk represents the proportion of disease among exposed people that actually results from the exposure. This issue might arise in a court case in which an exposed individual claims that the agent to which he or she was exposed caused the disease. Note that we are presuming that the associations reflect causality for the purposes of estimation. The attributable risk among exposed individuals is calculated as follows:

[relative risk–1]/relative risk

Thus, a relative risk of 2.0 (risk is doubled by exposure) yields an attributable proportion among exposed people of 0.5. This suggests a 50% chance that the disease resulted from the exposure in this study population.

Population-Attributable Risk

Of still greater potential value is the incorporation of information on how common the exposure is. Although some exposures exert a powerful influence on individuals (i.e., a large relative risk), they are so rare that their public health impact is minimal. Conversely, some exposures have a modest impact but are so widespread that their elimination could have great benefit. To answer the question, "What proportion of disease in the total population is a result of the exposure?" the population-attributable risk or etiologic fraction is used. The population-attributable risk can be calculated in two ways. First, if the rate in the exposed and unexposed population is known, then the population-attributable risk is

$$[\text{rate (total population)} - \text{rate (unexposed)}]/[\text{rate (total population)}]$$

The population-attributable risk can also be calculated with information only about the relative risk (usually obtained from research studies) and exposure in the population (often obtained from surveys). It is calculated as follows:

$$[P_e\ (\text{relative risk} - 1)]/[1 + P_e\ (\text{relative risk} - 1)]$$

where P_e represents the proportion of the population that is exposed.

The population-attributable risk is used most commonly to estimate the proportion of disease caused by a certain risk factor, such as the proportion of lung cancer caused by smoking. Assuming that the relative risk of lung cancer attributable to cigarette smoking is 15 and that 30% of the population are smokers, the population-attributable risk is 0.81, or 81%. This would suggest that 81% of the lung cancer burden in the population is caused by cigarette smoking and could be eliminated if the exposure were eliminated.

Population-attributable risk could also be used to estimate the proportion of a risk factor attributable to various "upstream" determinants. As described previously, research has shown that children exposed to smoking in the movies are at increased risk of becoming regular smokers (Charlesworth and Glantz 2005). If one assumes that this is a causal relationship, about 35% of smoking among adolescents could be attributed to exposure to smoking in the movies. Intuitively, the 35% figure sounds very high, which would cast doubt on the assumption of a casual relationship. It is also important to note that, because of the effects of interactions between various risk factors, population-attributable risk estimates for a given disease can sometimes add up to more than 100%.

The relationship between risk, relative risk, and population-attributable risk is illustrated in Figure 2-5 as a hypothetical example for 100 children:

- Risk (e.g., 34% of children exposed to smoking in the movies become smokers, compared with only 20% of children not exposed to smoking in the movies).
- Relative risk (e.g., children exposed to smoking in the movies are 1.7 times more likely to become smokers).
- Population-attributable risk (e.g., 35% of smoking among adolescents is because of smoking in the movies).

Although population-attributable risk estimates provide a useful estimate of the public health burden, they may be unrealistic as absolute goals because only rarely can a risk factor be completely eliminated.

Interactions among exposures, also known as effect modification, in the causation of disease are of particular importance in fully understanding etiology. Effect modification occurs when the effect of one exposure on disease risk is modified by the presence of another exposure. In the purest form, which is rarely observed, each of two exposures alone may have no effect on disease, but when the two are combined, a synergism occurs, causing an increase in disease. Conversely, two exposures that each can influence disease risk independently may be antagonistic, so that in combination they have a smaller effect on disease risk.

In epidemiological studies, interaction is measured as a combined effect of exposures that is larger than would be expected by simply adding the effects

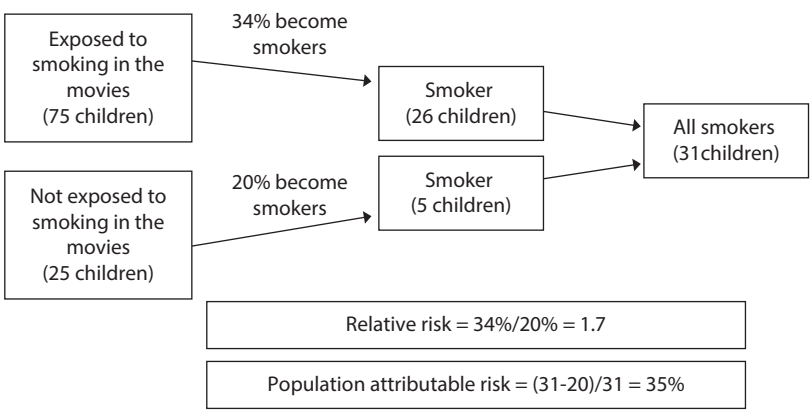


Figure 2-5. Risks of Becoming a Smoker among 100 Children Exposed or Not Exposed to Smoking in the Movies

of the two separate exposures. Interaction is most easily detected by comparing the disease risk in groups with all combinations of exposure to the group exposed to neither agent. For example, cigarette smoking and asbestos exposure have been demonstrated to multiply the risk of lung cancer (Table 2-4). The risk of lung cancer among nonsmokers who are exposed to asbestos (relative risk = 5.2) is approximately half that of smokers who are not exposed to asbestos (relative risk = 10.8). However, a multiplicative effect is observed among smokers who are exposed to asbestos (relative risk = 53.2).

This measure of interaction has direct implications for developing intervention and prevention programs. If two factors interact, then the benefit of removing a given exposure will be greater if the other exposure is also present. For example, eliminating smoking has even more benefit for a group of workers exposed to asbestos than for individuals who are not exposed to asbestos.

Obviously, many factors enter into decisions about interventions, including certainty of causality, amenability to intervention, and social and political issues. However, in the traditional role of epidemiology as the basic science of public health, quantitative considerations of preventable disease can help us make a rational choice. How can we predict what benefits one or more of these interventions might yield in the community? This is where estimates of population-attributable risk may be particularly useful. If one considers the earlier example of smoking and lung cancer, it is apparent that lung cancer incidence could be reduced by more than 80% if cigarette smoking were eliminated.

Additional Uses of Epidemiology in Chronic Disease Prevention and Control

Although descriptive epidemiology is necessary, it is not sufficient to prevent and control chronic diseases in populations. Developing evidence-based

Table 2-4. Example of Relative Risk Estimates for Lung Cancer Associated with Smoking and Asbestos Exposure

	Relative Risk Estimate	
Smoking Category	**No Asbestos Exposure**	**Asbestos Exposure**
Nonsmoker	1	5
Smoker	10	50

Source: Data from IARC (2004).

interventions requires going beyond epidemiology to use other disciplines, such as behavioral and social sciences, environmental health sciences, health management, and policy analysis. This section briefly describes a set of questions and issues to be considered when one is interpreting data in chronic disease epidemiology. It also describes several analytic tools and processes of interest to practitioners (e.g., meta-analysis, expert panels).

Why Should Action Not Be Taken on the Basis of a Single Epidemiological Study?

Scientists and practitioners committed to improving the public's health have a natural tendency to scrutinize the epidemiological literature for new findings that would serve as the basis for prevention or intervention programs. In fact, application to public health practice is a principal motive for conducting such research. Adding to this inclination to intervene may be claims from investigators regarding the critical importance of their findings, media interpretation of the findings as the basis for immediate action, and even community support for responding to the striking new research findings with new or modified programs or elimination of existing programs. John Oliver, a television satirist who blends investigative journalism with slapstick humor, scrutinized the misuse and misinterpretation of scientific studies by the media on his program *Last Week Tonight.* The program illustrates how the misuse of science can lead to erroneous and often bizarre conclusions such as a news anchor's assertion that "women are more open to romance when they are full instead of hungry" (Last Week Tonight 2016). The statement is presented as fact, and may indeed be true; however, it originated from a pilot study with a sample size of 20 that was published in a niche journal (Ely et al. 2015). Oliver points out these necessary disclaimers, which the news anchor failed to do. Although it is presented in a humorous fashion, Oliver's satire highlights an important concept: a single study is never sufficient for making such assertions. This is particularly true for epidemiological research, in which an appreciation of epidemiological methods applied to chronic disease prevention and control leads to the inevitable conclusion that multiple studies are required to ascertain the truth behind the statistical noise. Well-designed and carefully conducted research adds evidence to assist in setting policy, but the stakes are so high in economic terms and public credibility that cautious interpretation of research findings that have not been replicated is required.

We have already discussed the validity of epidemiological research and criteria for causality, but a more direct consideration of why a single study is insufficient for action should be helpful. Breast cancer had rarely been considered a disease that might be affected by environmental pollutants, in contrast, for example, to lung or bladder cancer. However, a paper published by Wolff et al. (1993) changed that perception. This study, which was well designed, carefully conducted, and certainly worthy of publication, reported that pesticide residues from dichlorodiphenyltrichloroethane (DDT) in the blood were positively associated with the risk of breast cancer, with relative risks on the order of 3.0. Media interest was high, and the notion that a common, life-threatening disease could be related in part to environmental agents was both terrifying and promising of the potential for intervention. Although these exposures had been accrued over a lifetime, if found to be related to breast cancer, interventions to reduce body burden of DDT would be worthy of consideration. For other reasons, largely based on adverse effects on wildlife, DDT was banned more than 40 years ago.

How could a study be methodologically sound, published in a reputable journal, generate substantial media and public interest, and yet not be worthy of any action on the part of public health practitioners? First, despite the talents of the investigators and quality of the resulting study, its findings may simply be wrong or misleading. That is, for reasons discussed earlier, even within the population studied, there may well be no causal association between DDT residues and breast cancer. One critical concern is whether the measured levels of DDT residues were affected by early stages of breast cancer rather than the reverse. Also, potential confounding by lactation was examined but produced rather anomalous results, and the number of cases available for analysis was small.

Second, even if valid for the population under study, the generalizability of the findings has to be examined. Can results from women in the New York area be applied to other populations with different exposure histories, ethnic backgrounds, and risk factor profiles? Although we look for universal explanations for disease and occasionally find strong causal factors that account for most of the disease in most populations (e.g., smoking and lung cancer), these successes are rare. More often, there is a multitude of interacting factors that must be considered for extrapolating findings from one population to apply to another.

Finally, even if findings are valid and generalizable, thus making a contribution to public health decision, what action should be taken for an agent that was banned long ago? At the margins, in evaluating costs and benefits, this

research would (if valid) add to the evidence of health harm from exposure, but in this case, there is not a clear decision to be made. If we had two comparably effective pesticides, one of which was associated with health harm and one that was not, such evidence might tip the balance. Public health decisions must integrate the full array of considerations regarding risks and benefits of different courses of action, as discussed in the following section.

How Do We Assess a Series of Epidemiological Studies and Integrate the Evidence to Make Decisions?

Several important methods and tools are available to epidemiologists and practitioners to assist in determining when public health action is warranted. It is often necessary to use these because exposure–disease relationships in chronic disease epidemiology typically show relatively weak associations, in which the relative risk estimate is not too different from the null value of 1.0. Most accepted risk factors for breast cancer are associated with relative risks of less than 5.0, and often less than 3.0, considered by some to be weak (Wynder 1987; Boffetta 2010). The closer a relative risk estimate comes to unity (i.e., 1.0), the more likely that it can be explained by methodological limitations such as confounding, misclassification, and other sources of bias. Yet, as noted earlier in the description of population-attributable risk, even when a risk factor is weak, if highly prevalent in the population, the public health impact can be large. Therefore, we have great interest in determining when even relatively weak associations provide the basis for public health intervention, which can only come from a series of well-designed studies.

Although it is tempting to intervene rather than conduct yet more studies in the face of a serious health problem, in the long run, evaluation of all major interventions is essential. To conduct such an evaluation, investigators must study a comparison group and rule out other factors as the cause of any observed change. The diagnosis of breast cancer in a prominent local citizen, for example, could be the main cause of increased breast cancer screening, rather than a community education program. In programs that combine several interventions, investigators may be able to determine which intervention(s) actually produced the health benefits and whether the results of an evaluation are generalizable to a different population. The debate continues, for example, as to whether the results of studies of reduced serum cholesterol in men can be applied to women. Thus, determining whether a proposed intervention will actually bring about more good than harm can be difficult.

It is important to note that even ostensibly useful interventions may not have positive effects and that almost all programs may have unintended negative effects. An education program to increase breast cancer screening, for example, could have no effect on women who need screening, yet it could raise anxiety among younger women at low risk who do not need screening. The usefulness of a particular screening test is based on several characteristics, including its accuracy, reproducibility, sensitivity, and specificity.

This section provides an overview of three related methods that have proven useful in assimilating large bodies of evidence in chronic disease epidemiology. In turn, the summarized evidence can be useful in shaping public health interventions and policies. Because this consideration is necessarily brief, readers are referred to other sources for more detail.

Systematic Reviews and Meta-analyses

The volume of information published about chronic diseases and their risk factors in journals every day is far beyond that which any person can remain current on. To address this problem, researchers conduct "systematic reviews" to consolidate all the information from studies addressing a single clinical or public health question. Systematic reviews use explicit and comprehensive (systematic) methods to identify, select, and critically assess all relevant research on the issue under consideration. To avoid bias, all Cochrane reviews start as a published protocol stating in advance how the review will be carried out (searching for data, appraising and combining study data). Over the past two decades, systematic reviews have been increasingly using "meta-analysis" to provide a more quantitative approach for integrating the findings of individual studies (Petitti 2000). Petitti describes four steps in undertaking a meta-analysis.

1. Identify relevant studies: Relevant studies must first be identified for inclusion in the meta-analysis. These can be identified through computerized sources such as MEDLINE, review articles, other journal articles, doctoral dissertations, and personal communications with other researchers.
2. Inclusion and exclusion criteria: Explicit criteria distinguish a meta-analysis from a qualitative literature review. Criteria for inclusion should specify the study designs to be included; the years of publication or of data collection; the languages in which the articles are written (e.g., English only or English plus other specified languages); the minimum sample size and the extent of follow-up; the treatments and/or exposures; the

manner in which the exposures, treatment, and outcomes were measured; and the completeness of information. Study quality should also be considered. As a minimum, studies whose quality falls below some specified rating criteria should be excluded. Rating scales may be developed to assess the quality of the included studies, although the basis for rating can be controversial and it may be preferable to consider the actual study attributes rather than a summary quality score.

3. Data abstraction: In this step, important features of each study are abstracted such as design, number of participants, and key findings. The abstraction summary should produce findings that are reliable, valid, and free of bias. Blinding of abstractors and reabstraction of a sample of studies by multiple abstractors may be beneficial.
4. Statistical analysis and exploration of heterogeneity: The data are combined to produce a summary estimate of the measure of association along with confidence intervals. Data are also examined to determine if the effect across studies was homogeneous and, if not, the reasons for heterogeneity.

Meta-analysis is most useful for combining the results of multiple, small RCTs whose results are generally consistent yet imprecision is a problem in each individual trial. The method is less useful in situations in which intervention trials have found truly heterogeneous results through different methods or because the relationship of interest varies across populations. Partly for that reason, one must be careful not to be overwhelmed by the impressively large numbers that can be accrued in a meta-analysis. Although the improved precision is a strength, one must also consider the validity of combining results across studies and whether the estimated size of the effect is large enough to warrant action.

Risk Assessment

Quantitative risk assessment is a widely used term for a systematic approach to characterizing the risks posed to individuals and populations by environmental pollutants and other potentially adverse exposures. Risk assessment has been described as a "bridge" between science and policy-making. In the United States, its use is either explicitly or implicitly required by a number of federal statutes, and its application worldwide is increasing. There has been considerable debate over the U.S. risk-assessment policies, and the most widely

recognized difficulties in risk assessment are because of extrapolation-related uncertainties (i.e., extrapolating low-dose effects from higher exposure levels). Risk assessment has become an established process through which expert scientific input is provided to agencies that regulate environmental or occupational exposures.

Four key steps in risk assessment are hazard identification, risk characterization, exposure assessment, and risk estimation (USEPA 2005). An important aspect of risk assessment is that it frequently results in classification schemes that take into account uncertainties about exposure–disease relationships. For example, the U.S. Environmental Protection Agency has developed a five-tier scheme for classifying potential and proven cancer-causing agents, which includes the following: (1) group A, carcinogenic to humans; (2) group B, probably carcinogenic to humans; (3) group C, possibly carcinogenic to humans; (4) group D, not classifiable as to human carcinogenicity; and (5) group E, evidence of noncarcinogenicity for humans.

Sources of Evidence

Most government agencies, in both executive and legislative branches, and voluntary health organizations, such as the American Cancer Society, use expert panels when examining epidemiological studies and their relevance to health policies and interventions (Brownson 2006). Ideally, the goal of expert panels is to provide peer review by scientific experts of the quality of the science and scientific interpretations that underlie public health recommendations, regulations, and policy decisions. If conducted well, peer review can provide an important set of checks and balances for the regulatory process. Optimally, the expert review process has the following common properties: experts are sought in epidemiology and related disciplines (e.g., clinical medicine, biomedical sciences, biostatistics, economics, ethics); panels typically consist of 8 to 15 members and meet in person to review scientific data (written guidance is provided to panel members); panel members should not have financial or professional conflicts of interest; and draft findings from expert panels are frequently released for public review and comment before final recommendations.

The Community Guide is a source of information about the effectiveness of chronic disease prevention and control interventions, based on the findings of the Community Preventive Services Task Force and the related

systematic reviews. The Community Guide is a credible resource with many uses because it is based on a scientific systematic review process and answers questions critical to almost everyone interested in community health and well-being such as

- What interventions have and have not worked?
- In which populations and settings has the intervention worked or not worked?
- What might the intervention cost?
- What should I expect for my investment?
- Does the intervention lead to any other benefits or harms?
- What interventions need more research before we know if they work or not?

The goal of the Community Guide is to promote the use of interventions that have been shown to work, discourage the use of interventions that have been shown not to work, and promote research on interventions for which there is not enough evidence to say whether or not they work (Community Preventive Services Task Force, 2016).

Conclusions

In this chapter, we have discussed the issues that help to determine whether associations are causal, the role of intervention research in disease control, and the uses of epidemiological evidence to make public health decisions. We have also outlined several types of epidemiological study designs and the biases that can complicate interpretation of results. We have discussed issues that help to determine whether associations are causal, as well as the role of intervention research in disease control. Such information serves as a foundation that will help readers address the more complex issues of chronic disease etiology and control. Some of the key epidemiological concepts that the reader will encounter in other chapters are summarized in Table 2-5.

The final section in each chapter of Parts II through IV describes the evidence that exists for effective programs and policies. Epidemiological, behavioral, social science, and other research methods have been used to identify effective intervention strategies that can be implemented at the individual and community level (Zaza et al. 2005). These methods are described in more detail in the following two chapters.

Table 2-5. Key Concepts in Chronic Disease Epidemiology

Term	Definition
Incidence rate=A/B	A=Number of new events in a specified time period B=Number of persons exposed to risk during this period
Relative risk=C/D	C=Risk of disease or death in the exposed population D=Risk of disease or death in the unexposed population
Population-attributable risk[a]=E/F	E=Rate of a disease in a population that is associated with (attributed to) a certain risk factor F=Total rate of a disease in the population

[a]This concept implies causality and should be used with caution.

Suggested Reading

Brownson, RC, Petitti DB, eds. *Applied Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 2006.

Cochran Collaboration and Library. Available at: http://www.cochrane.org.

Friss RH, Sellers TA. *Epidemiology for Public Health Practice*. Frederick, MD: Aspen; 1996.

Gordis L. *Epidemiology.* 2nd ed. Philadelphia, PA: W. B. Saunders; 2000.

Riegelman RK. *Studying a Study and Testing a Test: How to Read the Medical Evidence.* 6th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2013.

Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. Gaithersburg, MD: Aspen; 2000.

References

American Cancer Society. *Breast Cancer Facts & Figures 2015–2016*. Atlanta, GA: American Cancer Society Inc; 2015.

Andrieu N, Easton DF, Chang-Claude J, et al. Effect of chest X-rays on the risk of breast cancer among BRCA1/2 mutation carriers in the international BRCA1/2 carrier cohort study. *J Clin Oncol*. 2006;24(21):3361–3366.

Boehm JE, Rohan EA, Preissle J, DeGroff A, Glover-Kudon R. Recruiting patients into the CDC's colorectal cancer screening demonstration program: strategies and challenges across 5 sites. *Cancer.* 2013;119(suppl 15):2914–2925.

Boffetta P. Causation in the presence of weak associations. *Crit Rev Food Sci Nutr.* 2010;50(suppl 1):13–16.

Brownson RC. Epidemiology and health policy. In: Brownson RC, Petitti DB, eds. *Applied Epidemiology: Theory to Practice*. 2nd ed. New York, NY: Oxford University Press; 2006:361–363.

Brownson RC, Petitti DB, eds. *Applied Epidemiology*. 2nd ed. New York, NY: Oxford University Press; 2006.

Charlesworth A, Glantz SA. Smoking in the movies increases adolescent smoking: a review. *Pediatrics*. 2005;116(6):1516–1528.

Centers for Disease Control and Prevention (CDC). *Principles of Epidemiology*. 3rd ed. 2016. Available at: http://www.cdc.gov/ophss/csels/dsepd/ss1978/lesson1/section1.html. Accessed May 20, 2016.

Community Preventive Services Task Force. *The Guide to Community Preventive Services: What Works to Promote Health?* 2016. Available at: http://www.thecommunityguide.org. Accessed August 4, 2016.

Coughlin SS, Berkowitz Z, Hawkins NA, Tangka F. Breast and colorectal cancer screening and sources of cancer information among older women in the United States: results from the 2003 Health Information National Trends Survey [erratum in: *Prev Chronic Dis*. 2007;4(4):A114]. *Prev Chronic Dis*. 2007;4(3):A57.

Eaker ED, Jaros L, Vierkant RA, Lantz P, Remington PL. Women's Health Alliance Intervention Study: increasing community breast and cervical cancer screening. *J Public Health Manag Pract*. 2001;7(5):20–30.

Ely AV, Childress AR, Jagannathan K, Lowe MR. The way to her heart? Response to romantic cues is dependent on hunger state and dieting history: an fMRI pilot study. *Appetite*. 2015;95:126–131.

Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) study. *Am J Prev Med*. 1998;14(4):245–258.

Foege WH, Hogan RC, Newton LH. Surveillance projects for selected diseases. *Int J Epidemiol*. 1976;5(1):29–37.

Frieden TR. Shattuck lecture: the future of public health. *N Engl J Med*. 2015;373(18):1748–1754.

Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. *J Natl Cancer Inst Monogr*. 2006;(36):37–47.

Galea S, Tracy M, Hoggatt KJ, Dimaggio C, Karpati A. Estimated deaths attributable to social factors in the United States. *Am J Public Health*. 2011;101(8):1456–1465.

Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58(5):295–300.

Howlader N, Noone AM, Krapcho M, et al., eds. *SEER Cancer Statistics Review, 1975–2009 (Vintage 2009 Populations)*. Bethesda, MD: National Cancer Institute; 2012.

Institute of Medicine. *The Future of Public Health*. Washington, DC: National Academy Press; 1988.

International Agency for Research on Cancer (IARC). Tobacco smoke and involuntary smoking. In: *Monographs on the Evaluation of Carcinogenic Risks to Humans*. 2004;83:913–972. Available at: https://monographs.iarc.fr/ENG/Monographs/vol83/mono83-6B-4.pdf. Accessed August 25, 2016.

Jaros L, Eaker ED, Remington PL. Women's Health Alliance Intervention Study: description of a breast and cervical cancer screening program. *J Public Health Manag Pract*. 2001;7(5):31–35.

Kitchen AJ, Trivedi P, Ng D, Mokbel K. Is there a link between breast cancer and abortion: a review of the literature. *Int J Fertil Womens Med*. 2005;50(6):267–271.

Kohler BA, Sherman RL, Howlader N, et al. Annual report to the nation on the status of cancer, 1975–2011, featuring incidence of breast cancer subtypes by race/ethnicity, poverty, and state. *J Natl Cancer Inst*. 2015;107(6):djv048.

Koplan JP, Thacker SB. Fifty years of epidemiology at the Centers for Disease Control and Prevention: significant and consequential. *Am J Epidemiol*. 2001;154(11):982–984.

Last Week Tonight. John Oliver discusses how and why media outlets so often report untrue or incomplete information as science. May 8, 2016. Available at: https://www.youtube.com/watch?v=0Rnq1NpHdmw. Accessed August 4, 2016.

Liang SY, Phillips KA, Nagamine M, Ladabaum U, Haas JS. Rates and predictors of colorectal cancer screening. *Prev Chronic Dis*. 2006;3(4):A117.

McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd ed. New York, NY: Oxford University Press; 2006.

Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004;291(10):1238–1245.

National Cancer Institute (NCI). Cancer mortality maps. 2016a. Available at: http://ratecalc.cancer.gov/ratecalc/index.jsp. Accessed August 29, 2016.

National Cancer Institute (NCI). Breast Cancer Risk Assessment Tool. 2016b. Available at: http://www.cancer.gov/bcrisktool/Default.aspx. Accessed August 4, 2016.

Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. 2nd ed. New York, NY: Oxford University Press; 2000.

Reynolds P, Hurley SE, Gunier RB, Yerabati S, Quach T, Hertz A. Residential proximity to agricultural pesticide use and incidence of breast cancer in California, 1988–1997. *Environ Health Perspect*. 2005;113(8):993–1000.

Riegelman RK. *Studying a Study and Testing a Test: How to Read the Medical Evidence*. 6th ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2013.

Rookus MA, van Leeuwen FE. Induced abortion and risk for breast cancer: reporting (recall) bias in a Dutch case–control study. *J Natl Cancer Inst*. 1996;88(23):1759–1764.

Shelby RA, Scipio CD, Somers TJ, Soo MS, Weinfurt KP, Keefe FJ. Prospective study of factors predicting adherence to surveillance mammography in women treated for breast cancer. *J Clin Oncol*. 2012;30(8):813–819.

Sprague BL, Gangnon RE, Hampton JM, et al. Variation in breast cancer–risk factor associations by method of detection: results from a series of case–control studies. *Am J Epidemiol*. 2015;181(12):956–969.

Thun MJ, Myers DG, Day-Lally C, et al. Age and the exposure–response relationships between cigarette smoking and premature death in Cancer Prevention Study II. In: *Monograph 8: Changes in Cigarette-Related Disease Risks and Their Implications for Prevention and Control*. National Cancer Institute. 1997:383–413. Available at: http://cancercontrol.cancer.gov/brp/tcrb/monographs/8/m8_5.pdf. Accessed August 4, 2016.

Tinker T, Vaughan E. Risk communication. In: Nelson DE, Brownson RC, Remington PL, Parvanta C, eds. *Communicating Public Health Information Effectively: A Guide for Practitioners*. Washington, DC: American Public Health Association; 2002:185–203.

US Environmental Protection Agency (USEPA). Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum. PA/630/P-03/001F. Washington, DC: US Environmental Protection Agency; 2005.

Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. *Prev Chronic Dis*. 2014;11:E62.

Wolff M, Toniolo P, Lee E, Rivera M, Bubinn N. Blood levels of organochlorine residues and risk of breast cancer. *J Natl Cancer Inst*. 1993;85(8):648–652.

Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. Principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321–333.

Wynder EL. Workshop on guidelines to the epidemiology of weak associations. *Prev Med*. 1987;16(2):139–141.

Zaza S, Briss PA, Harris KW, eds. *The Guide to Community Preventive Services: What Works to Promote Health?* New York, NY: Oxford University Press; 2005:431–448.