CHAPTER 2

Measures of Disease Frequency

▶ Introduction

Epidemiologists study the distribution and determinants of **disease** frequency in human populations to control health problems.[1(p1),2(p95)] Thus, the objectives of epidemiology are to determine the extent of disease in a population, identify patterns and trends in disease occurrence, identify the causes of disease, and evaluate the effectiveness of prevention and treatment activities. Measuring how often a disease arises in a population is usually the first step in achieving these goals.

This chapter describes how epidemiologists quantify the occurrence of disease in a population. Readers will learn that this quantification process involves developing a definition of a particular disease, counting the number of people who are affected by the disease, determining the size of the population from which the diseased cases arose, and accounting for the passage of time.

▶ Definition of a Population

Because epidemiology is concerned with the occurrence of disease in groups of people rather than in individuals, populations are at the heart of epidemiologists' measurements. A **population** can be defined as a group of people with a common characteristic, such as place of residence, religion, gender, age, use of hospital services, or life event (such as giving birth).

Location of residence, such as a country, state, city, or neighborhood, is one of the most common ways to define a population. For example, the people who reside in the Brooklyn borough of New York City, the city of Boston, the state of Oregon, and the country of Sweden are members of distinct populations defined by geopolitical entities ranging in size from a neighborhood to an entire country. Residence near natural geographic features, such as rivers, mountains, lakes, or islands, can also be used to define a population. For example, people who live along the 2,350-mile length of the Mississippi River, around Mount St. Helens in Washington State, and on Nantucket Island off the coast of Massachusetts are members of populations defined by geographic formations.

Because epidemiology focuses on disease occurrence, populations are commonly defined in relation to a medical facility, such as a medical professional's office, clinic, or hospital. The service population of a medical facility (also called **catchment population**) consists of the people who use the facility's services. This population is often difficult to define because an individual's decision to use a facility may depend on how far it is from home, the person's particular medical condition, his or her type of medical insurance, and so forth.

Consider a situation in which a county has only one general hospital that provides the complete range of medical services, including preventive care, birthing services, and diagnostic and therapeutic services for acute and chronic conditions. The catchment population for this general hospital is likely to consist of all people who live in the county where the hospital is located (see **FIGURE 2-1**).

Now, suppose that this hospital enhances its cardiology department, adding many well-trained clinicians and the latest diagnostic equipment. As the cardiology department's reputation for excellent care grows, patients travel from greater distances to receive care. As a result, the catchment population for the cardiology department expands to the surrounding counties, whereas the catchment population for the other hospital services, particularly those dealing with acute conditions requiring prompt treatment, remains the single county where the hospital is situated (see Figure 2-1).

Socioeconomic status is still another determinant of hospital catchment populations. Consider a city in which there are two hospitals—one public and one private—located within a few miles of each other. The private hospital generally treats patients with medical insurance, and the public hospital mainly treats patients without insurance. Even though each catchment population resides roughly within the same geographic
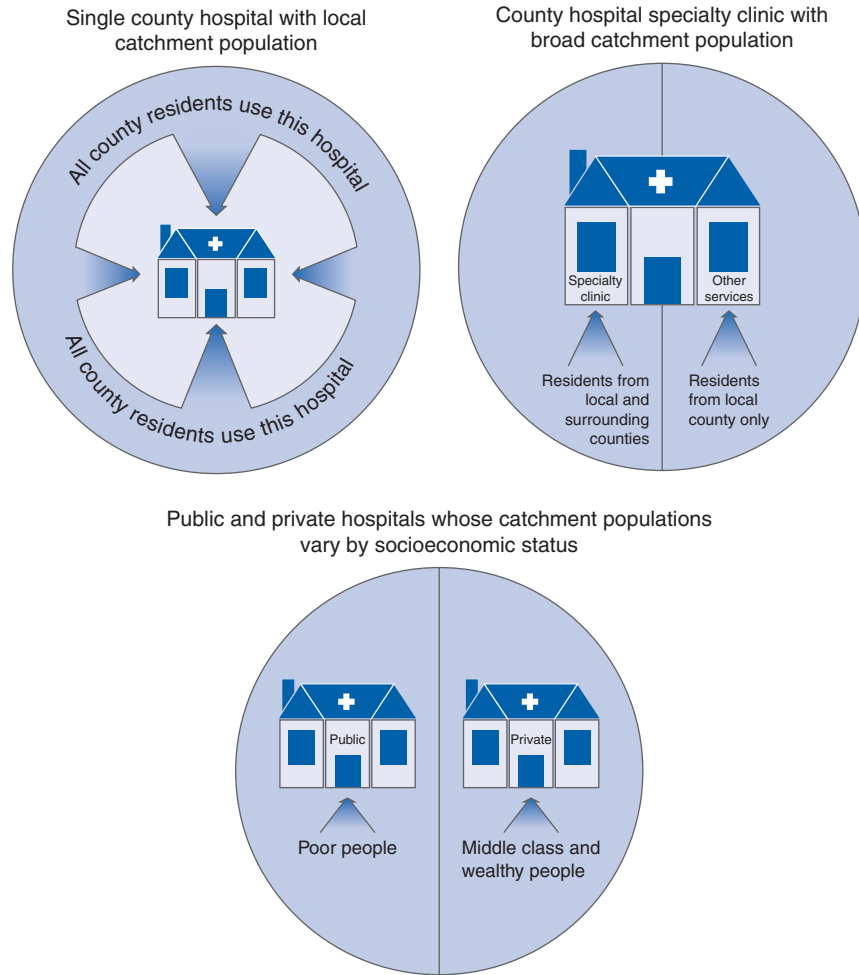
Single county hospital with local
catchment population

County hospital specialty clinic with
broad catchment population

Public and private hospitals whose catchment populations
vary by socioeconomic status

FIGURE 2-1 Types of hospital catchment areas.

area, the two service groups are distinct in terms of income and probably many other factors (see Figure 2-1).

Still another way that a population can be defined is by the occurrence of a life event, such as undergoing a medical procedure, giving birth to a child, entering or graduating from school, or serving in the military. For example, students who graduated from college in 2017 are members of the population known as the "Class of '17," and the men and women who served in the U.S. military during the War in Iraq are members of the population known as Iraq War veterans. Populations are often defined by other characteristics such as age, gender, religion, or type of job.

A unifying framework for thinking about a population is whether its membership is permanent or transient (see **TABLE 2-1**). A population whose membership is permanent is called a **fixed population**. Its membership is always defined by a life event. For example, the people who

TABLE 2-1 Types of Populations		
Type of population	**Key element**	**Example**
Fixed	Membership is based on an event and is permanent.	Japanese atomic bomb survivors
Dynamic or open	Membership is based on a condition and is transitory.	Residents of a city, hospital patients

were in Hiroshima, Japan, when the atomic bomb exploded at the end of World War II are members of a fixed population. This population will never gain any new members because only people who were at this historical event can be members.

The opposite of a fixed population is a **dynamic population**, also called an open population. Its membership is defined by a changeable state or condition and therefore is transient. A person is a member of a dynamic population only as long as he or she has the defining state or condition. For example, the population of the city of Boston is dynamic because people are members only while they reside within the city limits. Turnover is always occurring because people enter the city by moving in or by birth, and people leave the city by moving away or by death. The term *steady state* describes a situation in which the number of people entering the population is equal to the number leaving. Dynamic populations include groups defined by geographic and hospital catchment areas, religious groups, and occupations.

Regardless of the way in which it is defined, a population can be divided into subgroups on the basis of any characteristic. For example, men who undergo coronary bypass surgery are a gender subgroup of a fixed population defined by a life event, and all children up to 6 years of age who live along the Mississippi River are an age subgroup of a dynamic population defined by a geographic formation.

▶ Definitions of Health and Disease

In 1948, the World Health Organization defined **health** as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity."[3] More recently, a national plan for improving the health of the American people stated that health is a key aspect of the quality of life and that it "reflects a personal sense of physical and mental health and the ability to react to factors in the physical and social environments." The plan also states that the "health-related quality of life is more subjective than life expectancy and therefore can be more difficult to measure."[4]

Because measurement is a cornerstone of epidemiology and "health" and "a sense of well-being" are nonspecific and difficult to quantify, epidemiologists have almost entirely focused their activities on the "absence of health," such as specific diseases, injuries, disabilities, and death. Consequently, epidemiologists must first compose a definition of the "absence of health" or "disease" before they can begin measuring its frequency.

The definition of a disease is usually based on a combination of physical and pathological examinations, diagnostic test results, and signs and symptoms. Which and how many criteria are used to define a "case" (a person who meets the disease definition) has important implications for accurately determining who has the disease.

Consider the various criteria that can be used to define a heart attack case. One could use the symptoms of chest pain; the results of diagnostic tests, such as electrocardiograms; or blood enzyme tests for cardiac damage. What are the implications of using only chest pain to define heart attack cases? Using only this nonspecific symptom will capture most but not all people who have heart attacks because it will miss people who have "silent" heart attacks, which occur without chest pain. In addition, it will erroneously include many people who have other conditions that produce chest pain, such as indigestion.

A definition that includes more specific criteria, such as the results of diagnostic tests, will be more accurate. For example, if positive blood enzyme tests are included, silent heart attacks are likely to be picked up and the other conditions that cause chest pain omitted. In practice, epidemiologists use all available information from physical and pathological examinations and laboratory and other tests to define a case of a disease as accurately as possible.

▶ Changes in Disease Definitions

Even when clear-cut criteria are used, disease definitions often change over time as more is learned about a disease and its various manifestations. For example, the official definition of HIV/AIDS (human immunodeficiency virus infection/acquired immune deficiency syndrome) has been changed several times as researchers have gained knowledge about the disease's cause and natural course (see **TABLE 2-2**).

The story began in the summer of 1981 when the Centers for Disease Control and Prevention (CDC) reported that Kaposi's sarcoma and *Pneumocystis carinii* pneumonia had been observed among previously healthy gay men.[5] Previously, Kaposi's sarcoma, a malignant neoplasm of the blood vessels, had been seen primarily among elderly males, and opportunistic infections, such as *Pneumocystis carinii* pneumonia, had occurred almost exclusively in people with compromised immune systems.

In 1982, the CDC composed the first definition of AIDS for national reporting as "a disease, at least moderately predictive of a defect in cell-mediated immunity, occurring in a person with no known

TABLE 2-2 Changes in the Definition of HIV/AIDS over Time

Year	State of knowledge and practices	Criteria for case definition
1982	Knowledge very limited	Only a few conditions, including Kaposi's sarcoma, *Pneumocystis carinii* pneumonia, and other severe opportunistic infections
1985	HIV virus discovered as cause of AIDS; antibody test developed	23 clinical conditions with laboratory evidence of infection
1993	Discovered importance of CD4 T lymphocytes in monitoring immunodeficiency and disease progression	26 clinical conditions and asymptomatic HIV infected cases with low CD4 T lymphocyte counts
2008	HIV testing widely available and diagnostic testing improved	Single case definition for HIV infection that includes AIDS and requires laboratory evidence of infection; 26 clinical conditions retained
2014	Further improvements in diagnostic testing, including recognition of early HIV infection	Single case definition ranging from early HIV infection to AIDS diagnosis; 27 AIDS-defining opportunistic illnesses included

Data from Centers for Disease Control and Prevention. Update on acquired immune deficiency syndrome (AIDS)—United States. *MMWR*. 1982;31:507-514; Centers for Disease Control and Prevention. Revision of the case definition of acquired immune deficiency syndrome for national reporting—United States. *MMWR*. 1985;34:373-375; Centers for Disease Control and Prevention. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR*. 1992;44:1-19; Centers for Disease Control and Prevention. Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged < 18 months and for HIV infection and AIDS among children aged 18 months to < 13 years—United States, 2008. *MMWR*. 2008;57:1-12; and Centers for Disease Control and Prevention. Revised surveillance case definition for HIV infection–United States, 2014. *MMWR*. 2014; 63:1-10.

cause for diminished resistance to that disease … including Kaposi's sarcoma, *Pneumocystis carinii* pneumonia, and serious other opportunistic infections."[6] Even then, the CDC acknowledged that the case definition was imperfect because it could possibly miss the full range of AIDS manifestations and falsely include individuals who were not truly immunodeficient. However, because the cause of AIDS was unknown at that time, the definition at least served as a consistent tool for monitoring the epidemic.

The CDC made the first major change in the official AIDS case definition in 1985 after HIV was determined to be the cause of AIDS and a highly accurate laboratory test was developed to detect the antibody to the HIV virus.[7] Because epidemiologists now knew what to look for and how to find its trail, it was possible to expand the AIDS case

definition from a few severe clinical conditions to a wide range of conditions accompanied by laboratory evidence of HIV infection. Conversely, it also became possible to exclude patients who had AIDS-like symptoms but had negative HIV antibody test results.

At the end of 1992, the CDC changed the AIDS case definition again after natural history studies revealed the importance of lymphocytes in the HIV infection.[8] The new definition was expanded to include individuals who had no disease symptoms but had low levels of CD4 T lymphocytes, a type of white blood cell that is responsible for fighting off infections. (These cells are the primary target of HIV and an excellent marker of disease progression.) In 2008, the CDC further refined the disease definition by combining HIV infection and AIDS into a single case definition and requiring laboratory confirmation of infection.[9] In 2014, further refinements were made primarily to recognize early HIV infection.[10]

In summary, as epidemiologists discovered the cause of AIDS, learned more about its manifestations, and developed a test to detect HIV, the case definition was expanded from a few severe diseases to many varied conditions and the results of sophisticated laboratory tests (see Table 2-2).

What effect did these changes in case definition have on the number of reported HIV/AIDS cases? In general, any expansion in the case definition will increase the number of reportable cases, and any contraction will decrease that number. With HIV/AIDS, the definition changes increased the number of people who were officially counted as cases in national statistics from the 1980s until 1993.[8] However, incidence subsequently declined despite the 1993 expanded case definition because of HIV testing and prevention programs, which reduced the rate of transmission, and treatment programs that included highly effective antiretroviral therapy.[11]

▸ Measuring Disease Occurrence

Epidemiologists must always consider three factors when they measure how commonly a disease occurs in a group of people: (1) the number of people who are affected by the disease, (2) the size of the population from which the cases of disease arise, and (3) the length of time that the population is followed. Failure to consider all three components will give a false impression about the effect of the disease on a population.

Consider the following hypothetical data on the frequency of breast cancer in two counties. In County A, with a population of 50,000, a total of 100 new cases of breast cancer occurred over a 1-year period. In County B, with a population of 5,000, 75 new cases occurred over a 3-year period. Which county has a higher frequency of new breast cancer cases? If one considers only the number of new cases, it appears that County A has a higher frequency (100 versus 75). However, simply comparing the number of cases in each county does not provide a full picture because the cases occurred over different lengths of time (1 versus 3 years) and among populations of different sizes (50,000 versus 5,000) (see **TABLE 2-3**).

TABLE 2-3   Hypothetical Data on the Frequency of Breast Cancer in Two Counties

Type of data	County A	County B
Number of cases	100	75
Population size	50,000	5,000
Follow-up period	1 year	3 years
Comparable disease frequency	200/100,000/year	500/100,000/year

To make a meaningful comparison between the two counties, it is necessary to convert the data into the same population size and time period. Let us estimate the frequency of breast cancer in the two counties over a 1-year period and as if each population consisted of 100,000 people. The frequency of breast cancer in County A is 100 cases/50,000 population/1 year; if the county's population were 100,000, the frequency would double to become 200 cases/100,000 population/1 year.

Two steps are needed to make a similar conversion for County B: (1) divide the numerator by 3 to convert the frequency of breast cancer from a 3- to a 1-year period: 25 cases/5,000 population/1 year; (2) multiply both the numerator and denominator by 20 to estimate the frequency for a population of 100,000. Thus, the frequency of new cases in County B is 500 cases/100,000 population/1 year.

Now, it is clear that the "rate" at which new cases are occurring is much higher in County B than County A (500 cases/100,000 population/1 year versus 200 cases/100,000 population/1 year, respectively) and that examining only the number of new cases gives a false impression. Note that the decision to convert the frequencies to a population size of 100,000 and a 1-year time period, although commonly done, is arbitrary. Other population sizes (such as 1,000 or 10,000) and time periods (such as 1 month or 5 years) could be used. The guiding principle is that the same population size and time period should be used for the compared groups.

▶ Types of Calculations: Ratios, Proportions, and Rates

Three types of calculations are used to describe and compare measures of disease occurrence: ratios, proportions, and rates (see **TABLE 2-4**). A **ratio** is simply one number divided by another. The entities represented by the two numbers are not required to be related to one another. In other words, the individuals in the numerator can be different from those in

TABLE 2-4 Types of Calculations	
Type of calculation	**Characteristics**
Ratio	Division of two unrelated numbers
Proportion	Division of two related numbers; numerator is a subset of denominator
Rate	Division of two numbers; time is always in denominator

the denominator. For example, the "gender ratio" is a ratio of two unrelated numbers: the number of males divided by the number of females, usually expressed as the number of males per 100 females. According to 2016 U.S. Census estimates, the gender ratio among U.S. residents in Florida was 95.5 males per 100 females.[12]

A **proportion** is also one number divided by another, but the entities represented by these numbers are related to one another. In fact, the numerator of a proportion is always a subset of the denominator. Proportions, also known as fractions, are often expressed as percentages and range from 0 to 1 or 0% to 100%. For example, the proportion of U.S. residents who are Black is the number of Black residents divided by the total number of U.S. residents of all races. According to 2016 U.S. Census estimates, the proportion of Black U.S. residents was 0.141, or 14.1%.[12]

A **rate** is also one number divided by another, but time is an integral part of the denominator. We are familiar with rates in our daily travels because a rate is a measure of how fast we travel. For example, in many areas of the United States, the maximum speed or rate at which cars are permitted to travel is 55 miles per hour. This rate can also be written as 55 miles/1 hour. The measure of time in the denominator is what makes this number a rate. The measures of disease occurrence calculated previously for Counties A and B are also rates (200 cases/100,000 population/1 year and 500 cases/100,000 population/1 year, respectively). Unfortunately, the term *rate* is often incorrectly used to describe ratios and proportions.[13]

▶ Measures of Disease Frequency

The two basic measures of disease frequency in epidemiology are incidence and prevalence. **Incidence** measures the occurrence of new disease, and **prevalence** measures the existence of current disease. Each measure describes an important part of the natural course of a disease. Incidence deals with the transition from health to disease, and prevalence focuses on the period of time that a person lives with a disease.

Incidence

Incidence is defined as the occurrence of new cases of disease that develop in a candidate population over a specified time period. There are three key ideas in this definition.

First, incidence measures *new disease events*. For diseases that can occur more than once, it usually measures the first occurrence of the disease. Second, new cases of disease are measured in a **candidate population**, which is a population of people who are "at risk" of getting the disease. Someone is at risk because he or she has the appropriate body organ, is not immune, and so forth. For example, a woman who still has an intact uterus (i.e., she has not undergone a hysterectomy) is a candidate for getting uterine cancer, and a child who has not been fully immunized against the measles virus is a candidate for contracting a measles infection. Although it is possible to define and measure the incidence of disease in a population not at risk (e.g., the incidence of uterine cancer in women who have undergone hysterectomies is, by definition, zero), it is not a particularly interesting pursuit.

Third, incidence takes into account the *specific amount of time* that the members of the population are followed until they develop the disease. Because incidence measures a person's transition from a healthy to a diseased state, time must pass for this change to occur and be observed. As described in the following sections, there are two types of incidence measures: cumulative incidence and incidence rate. Although closely related, each measure has different strengths and weaknesses and is used in different settings.

Cumulative Incidence

Cumulative incidence is defined as the proportion of a candidate population that becomes diseased over a specified period of time. Mathematically, it is expressed as follows:

$$\frac{\text{Number of new cases of disease}}{\text{Number in candidate population}} \text{ Over a specified time period}$$

Note that the numerator (new cases of disease) is a subset of the denominator (candidate population) and therefore the possible value of cumulative incidence ranges from 0 to 1, or if expressed as a percentage, from 0% to 100%. Time is not an integral part of this proportion but rather is expressed by the words that accompany the numbers of the cumulative incidence measure. Thus, cumulative incidence is dimensionless (see **TABLE 2-5**).

Cumulative incidence can be thought of as the average risk of getting a disease over a certain period of time. (A risk is the probability of getting a disease.) A commonly cited measure of cumulative incidence is the "lifetime risk of breast cancer" among women. Currently estimated at "one in eight" among U.S. women, it means that about 12% of women will develop breast cancer sometime during the course of their lives.[14]

TABLE 2-5 Distinguishing Characteristics of Incidence and Prevalence

Measure	Type of number	Units	Range	Numerator	Denominator	Major uses
Cumulative incidence	Proportion	None	0 to 1	New cases	Population at risk	Research on causes, prevention, and treatment of disease
Incidence rate	True rate	1/time, or t^{-1}	0 to infinity	New cases	Person-time at risk	Research on causes, prevention, and treatment of disease
Prevalence	Proportion	None	0 to 1	Existing cases	Total population	Resource planning

Cumulative incidence is influenced by the length of time to which it applies. Generally, the cumulative incidence over a long period of time (such as a lifetime) will be higher than that over a few years.

Cumulative incidence is mainly used in fixed populations when there are no or small losses to follow-up. Consider, for example, the estimated 255,000 residents of Hiroshima, Japan, who were present in the city when the atomic bomb was dropped on August 6, 1945 (see **FIGURE 2-2**). During the blast and in the weeks and months immediately thereafter, an estimated 65,000 people died from physical trauma, burns, and acute radiation sickness, resulting in a 25% cumulative incidence of mortality over a 4-month period.[15] Officials estimate that another 21% of the population died during the year following that initial 4-month period. During the subsequent years and decades, the cumulative incidence of death was much lower, and different causes of death, such as cancer and circulatory disease, predominated.[16] The cumulative incidence of death was 58% over the 53-year period from 1950 through 2003 among Hiroshima and Nagasaki survivors who were enrolled in the Life Span Cohort Study of the effects of high-dose radiation on the incidence of disease and death.[16]

Note that the cumulative incidence of death during the first 16 months after the bomb was dropped (25% + 21% = 46%) was close to that during the 53-year follow-up period of the Life Span Cohort Study (58%). In addition, the size of the candidate population dwindles as members die. For example, all 255,000 residents were at risk of dying when the bomb was dropped in August 1945, whereas approximately 190,000 survivors of the initial period were at risk of dying during the subsequent years (see Figure 2-2).
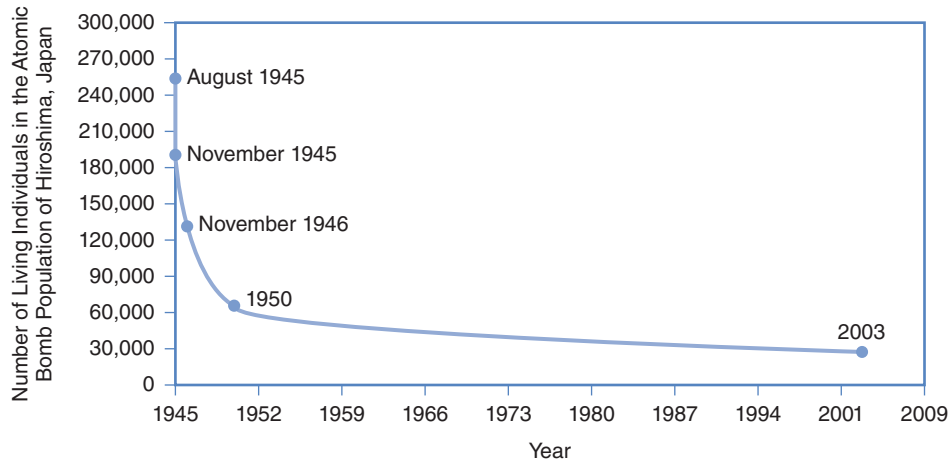
FIGURE 2-2 Dwindling size of the Hiroshima atomic bomb population over time: 1945–2003.

Data from Liebow, AA. *Encounter with Disaster: A Medical Diary of Hiroshima*. New York, NY: WW Norton and Co; 1970:175-176; Committee for the Compilation of Materials on Damage Caused by the Atomic Bombs in Hiroshima and Nagasaki. *Hiroshima and Nagasaki: The Physical, Medical, and Social Effects of the Atomic Bombings*. New York, NY: Basic Books; 1981:113; and Ozasa K, Shimizu Y, Suyama A, Kasagi F, Soda M, Grant EJ, et al. Studies of the mortality of atomic bomb survivors, Report 14, 1950-2003: an overview of cancer and noncancer diseases. *Radiat Res.* 2012;177:229-243.

The following hypothetical example further highlights the importance of selecting the appropriate time period for measuring cumulative incidence measure. Suppose that a pharmaceutical company has developed a new drug for the treatment of acute migraine headaches. Before the U.S. Food and Drug Administration (FDA) will approve its use, the company is required to conduct a small study (N = 20 patients) to determine whether the new drug is more effective than the most popular drug currently approved for migraine treatment. The FDA requests that the study answer two key questions: First, what proportion of patients experience symptom relief within 10 hours of taking the new drug as compared to the current drug? Second, does the new drug relieve symptoms faster than the current drug? As the study results indicate in **FIGURE 2-3**, the proportion of patients who experience symptom relief within 10 hours is identical for the two drugs. Stated in epidemiologic terms, the 10-hour cumulative incidence of relief is 60% for both drugs. However, all patients who took the new drug had symptom relief within 3 hours, whereas those who took the current drug had symptom relief several hours later. Note that the 10-hour measure of cumulative incidence misses this important finding and makes the drugs appear equally effective. Only when you examine the timing of relief can you see this important difference. Furthermore, if a 5-hour cumulative incidence had been selected as the measure of symptom relief instead of the 10-hour measure, this notable difference would have been captured.

A final critical assumption that underlies the cumulative incidence measure is that everyone in the candidate population has been followed for the specified time period. Thus, in the examples above, the cumulative incidence measures assume that everyone in the migraine study population was followed for 10 hours and that everyone in the Hiroshima Life Span
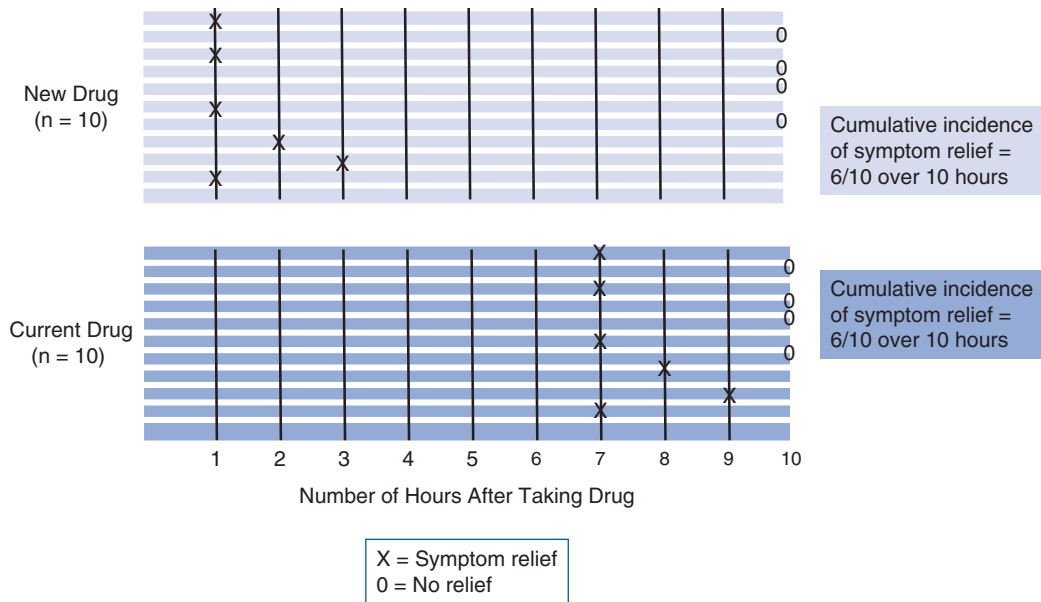
New Drug
(n = 10)

Cumulative incidence
of symptom relief =
6/10 over 10 hours

Current Drug
(n = 10)

Cumulative incidence
of symptom relief =
6/10 over 10 hours

1   2   3   4   5   6   7   8   9   10

Number of Hours After Taking Drug

X = Symptom relief
0 = No relief

FIGURE 2-3 Importance of selecting appropriate time period for cumulative incidence.

Cohort Study was followed for 53 years. Clearly, the former is much easier
to accomplish than the latter. In fact, complete follow-up is difficult to attain
when there is a long follow-up period, particularly in a dynamic population
in which members are continually entering and exiting. Thus, cumulative
incidence is usually reserved for fixed populations, particularly when the
follow-up period is short and there are no or few losses to follow-up.

Incidence Rate

Incidence rate is defined as the occurrence of new cases of disease that
arise during **person-time** of observation. Mathematically, the incidence
rate is expressed as follows:

$$\frac{\text{Number of new cases of disease}}{\text{Person-time of observation in candidate population}}$$

Note that the numerator for incidence rate is identical to that of
cumulative incidence. The difference between the two measures lies in
the denominator. The incidence rate's denominator integrates time (t)
and therefore is a true rate.[13] Thus, its dimension is 1/t, or t^{-1}, and its
possible values range from zero to infinity (Table 2-5). An incidence rate
of infinity is possible if all members of a population die instantaneously.

The concept of person-time can be difficult to understand.
Person-time is accrued only among candidates for the disease. Thus, a
person contributes time to the denominator of an incidence rate only up
until he or she is diagnosed with the disease of interest. However, unlike
cumulative incidence, the incidence rate is not based upon the assumption

that everyone in the candidate population has been followed for a specified time period. Person-time is accrued only while the candidate is being followed. Accrual of person-time stops when the person dies or is lost to follow-up (such as when a person moves to an unknown community). The incidence rate can be calculated for either a fixed or dynamic population. However, because it directly takes into account population changes (such as migration, birth, and death), it is especially useful as a measure of the transition between health and disease in dynamic populations.

Consider the following population of a hypothetical town (**FIGURE 2-4**). This dynamic population of five individuals is followed for the 10-year period from 2006 to 2016. Person A moved into town in 2007, was diagnosed with disease in 2011, and therefore accrued 4 years of person-time. Person B was a resident of the town at the start of the observation period in 2006, died in 2012, and therefore accrued 6 person-years. Person C moved to town in 2008 and remained healthy until he moved away in 2011. The investigator could not determine where he moved and therefore could not learn whether he became diseased later. Person C was considered lost to follow-up as of the time he moved and therefore accrued 3 person-years. Person D was a resident of the town at the start of the observation period, remained healthy for the entire period, and therefore accrued 10 years of person-time. Person E was born to Person A in 2009, remained healthy for the entire observation period, and therefore accrued 7 person-years.

The incidence rate in this hypothetical population is 1/30 person-years. Only one person became diseased, and 30 person-years of observation were accrued by the population. The denominator of the incidence rate is the sum of person-time accrued by each member of the population at risk.

Now, consider an actual population of 59,000 U.S. Black women who were studied to quantify the effect of racism and segregation on
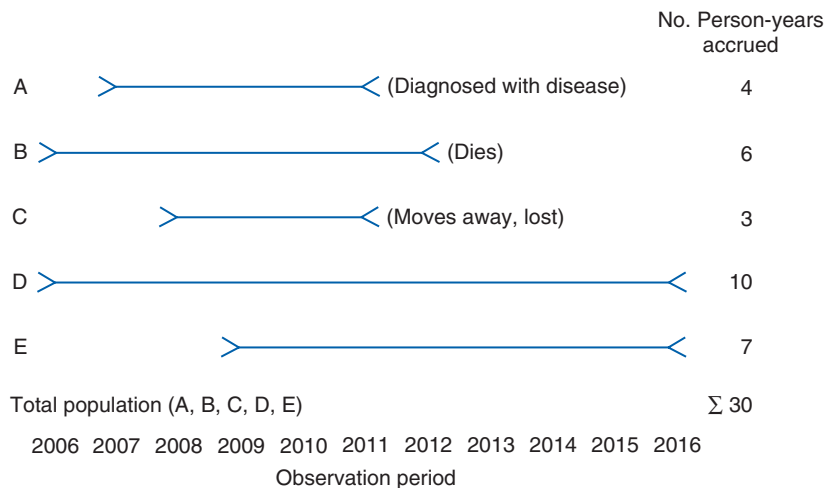

FIGURE 2-4 Measurement of person-time in a hypothetical population.

the incidence of obesity.[17] The women were enrolled in 1995 and fol-
lowed through 2009 to determine their weight over time. Each woman's
follow-up time began with the date of return of her questionnaire in 1997
when questions of racism were first asked and continued until one of the
following events occurred: the occurrence of obesity (defined as a body
mass index ≥ 30 kg/m^2), loss to follow-up, death, or the end of follow-up
for the study, whichever came first. Investigators identified 1,105 inci-
dent cases of obesity among 33,235 person-years of follow-up (33.2/1,000
person-years) among women who experienced the lowest levels of daily
racism in contrast to 1,043 incident cases of obesity among 23,403
person-years of follow-up (44.6/1,000 person-years) among women who
experienced the highest levels of daily racism. Although it is the most
accurate to determine person-time on an individual basis, as was done in
this study, this method is very time consuming and labor intensive. Thus,
researchers sometimes use shortcuts, such as multiplying the number of
people under observation by an estimate of the average follow-up time.

In the hypothetical example portrayed in Figure 2-4, five individuals
who were followed for varying amounts of time accrued 30 person-years
of follow-up. However, depending on the size of the population and
length of follow-up, the 30 person-years could have been accrued in many
ways, such as 5 people each contributing 6 years, 3 people each contribu-
ting 10 years, and so forth. Regardless of how the person-time is accrued
(e.g., from 5 or 50 people), the person-time units are assumed to be equiv-
alent. This assumption is usually reasonable, except in extreme situations
in which a small number of people are followed for a long period of time.

The particular time unit used to measure person-time can vary, but
decisions are guided by how long it takes for the disease to develop. For
example, person-years are commonly used for diseases that take many
years to develop (such as cancer), and person-months or person-days are
used for diseases that develop rapidly (such as infection).

The number of person-time units in the denominator is arbitrary. For
example, the same incidence rate can be expressed in terms of 1 person-year,
10 person-years, or 100 person-years. Epidemiologists generally use 100,000
person-years for rare diseases and those that take a long time to develop.

Relationship Between Cumulative Incidence and Incidence Rate

It is possible to obtain cumulative incidence from an incidence rate. The
simplest situation to demonstrate this relationship is in a fixed popula-
tion with a constant incidence rate and small cumulative incidence (less
than 10%). Here, the mathematical relationship is as follows:

$$CI = IR_i \times t_i$$

where CI is cumulative incidence, IR_i is incidence rate, and t_i is the
specified period of time.

When the incidence rate is not constant, it is necessary to take into account the different rates that prevail during each time period:

$$CI = \sum (IR_i \times t_i)$$

For example, the **mortality rate** (a type of incidence rate) among Hiroshima residents was much higher shortly after the atomic bomb explosion than during subsequent years (see Figure 2-2).

Incidence Summary

In summary, two measures of disease frequency—cumulative incidence and incidence rate—focus on measuring the transition from health to disease. These measures have complementary strengths and weaknesses. The cumulative incidence is easy to calculate and understand, although it is less accurate when its assumptions are not met. Although the incidence rate has greater accuracy, its person-time denominator is more difficult to calculate and understand. Finally, the incidence rate is more useful for dynamic populations, and cumulative incidence is usually reserved for fixed populations.

Prevalence

Whereas incidence measures the frequency with which new disease develops, prevalence measures the *frequency of existing disease*. It is simply defined as the proportion of the total population that is diseased. There are two types of prevalence measures—point prevalence and period prevalence—that relate prevalence to different amounts of time (see **FIGURE 2-5**). **Point prevalence** refers to the proportion of the population that is diseased at a single point in time and can be thought of as a single snapshot of the population. The point can be either a particular calendar date such as July 1, 2017, or a point in someone's life, such as college graduation. **Period prevalence** refers to the proportion of the population that is diseased during a specified duration of time, such as
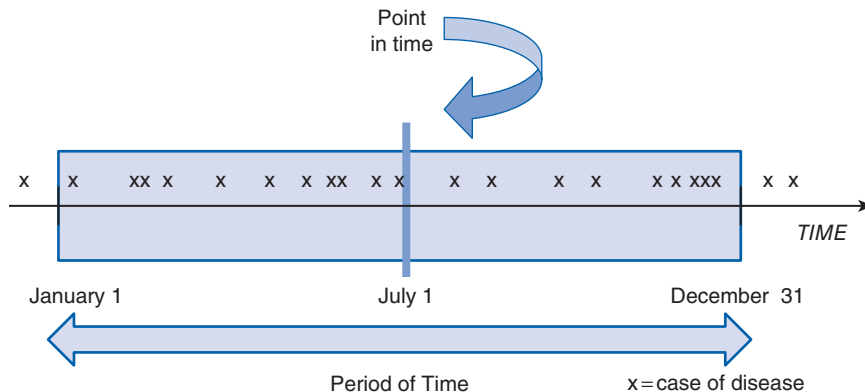

FIGURE 2-5 Time frame for point and period prevalence.

during the year 2017. The period prevalence includes the number of cases that were present at any time over the course of the year.

Mathematically, point prevalence is expressed as follows:

$$\frac{\text{Number of existing cases of disease}}{\text{Number in total population}} \text{ At a point in time}$$

Period prevalence can be expressed as follows:

$$\frac{\text{Number of existing cases of disease}}{\text{Number in total population}} \text{ During a period of time}$$

Let's use these formulas to calculate the point and period prevalence of pneumonia in a nursing home population. The point and period of interest are July 1, 2017, and January 1 through December 31, 2017, respectively. On July 1, 2017, there were 5 cases of pneumonia among the 500 nursing home residents. Thus, the point prevalence of pneumonia was 5/500, or 1%, on that date. During the period January 1 through December 31, 2017, there were 45 cases of pneumonia among the 500 nursing home residents; therefore, the period prevalence was 45/500, or 9%, during the year. Note that, in this example, the size of the nursing home population remained stable over the year, but if it had gained or lost members, the average size of the nursing home population during 2017 would have been the appropriate denominator for the period prevalence measure.

Note that the numerator (existing cases) is a subset of the denominator (total population). Unlike the numerator for the two incidence measures, the prevalence numerator includes all currently living cases regardless of when they first developed. The denominator includes everyone in the population—sick, healthy, at risk, and not at risk. Because prevalence is a proportion, it is dimensionless, and its possible values range from 0 to 1, or 0% to 100% (see Table 2-5).

Relationship Between Prevalence and Incidence

Prevalence depends on the rate at which new cases of disease develop (the incidence rate) as well as the duration or length of time that individuals have the disease. The duration of a disease starts at the time of diagnosis and ends when the person either is cured or dies. Mathematically, the relationship between prevalence and incidence is as follows:

$$\frac{P}{(1-P)} = IR \times D$$

where P is prevalence (the proportion of the total population with the disease), (1 − P) is the proportion of the total population without the disease, IR is incidence rate, and D is the average duration (or length of time) that an individual has the disease.
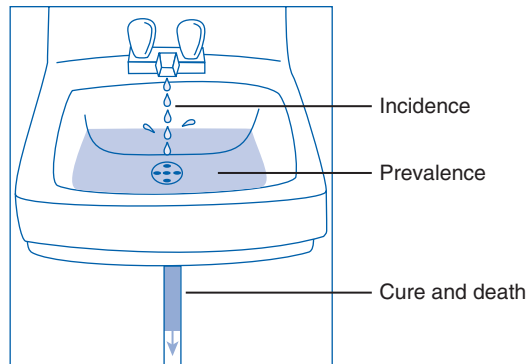
FIGURE 2-6 Relationship among incidence, prevalence, mortality, and cure.

This equation assumes that the population is in steady state (i.e., inflow equals outflow) and that the incidence rate and duration do not change over time. If the frequency of disease is rare (i.e., less than 10%), the equation simplifies to

$$P = IR \times D$$

To better understand this relationship, think of the variables that influence the level of water in a sink (see **FIGURE 2-6**). The water level is influenced by both the inflow from the faucet and the outflow down the drain. The water level will be high if the inflow is large, if the outflow is low, or if both occur. The water level will be low if inflow is low, outflow is high, or both. Now, consider the water level in the sink as prevalence, the incoming water as incidence, and the outflowing water as diseased people who either are cured or die. The number of cases of people currently living with a disease (prevalence) will be influenced by the rate at which new cases develop as well as by the rate at which they are eliminated through cure or death.

Uses of Incidence and Prevalence

Epidemiologists and other public health professionals use each measure of disease frequency for specific purposes (see Table 2-5). Incidence is most useful for evaluating the effectiveness of programs that try to prevent disease from occurring in the first place. In addition, researchers who study the causes of disease prefer to study new cases (incidence) over existing ones (prevalence) because they are usually interested in exposures that lead to developing the disease. Prevalence obscures causal relationships because it combines incidence and survival. In addition, many researchers prefer to use incidence because the timing of exposures in relation to disease occurrence can be determined more accurately.

On the other hand, prevalence is useful for estimating the needs of medical facilities and allocating resources for treating people who already have a disease. In addition, researchers who study diseases such as birth defects (wherein it is difficult to gather information on defects present in miscarried

and aborted fetuses) and chronic conditions such as arthritis (whose beginnings are difficult to pinpoint) have no choice but to use prevalence. Unfortunately, results of such studies are difficult to interpret because it is unclear how much the association is influenced by using a group of survivors.

▶ Commonly Used Measures of Disease Frequency in Public Health

There are many measures of disease frequency that are commonly used in the public health disciplines. Some are incidence measures, some are prevalence measures, some are ratios. Descriptions and examples of the major measures follow. Note that the word *rate* is often used incorrectly to describe a proportion or ratio.

Crude mortality (or death) rate: Total number of deaths from all causes per 100,000 population per year. The term **crude** means that the rate is based on raw data. In 2015 the crude mortality rate in the United States was 844.0/100,000 population/year.[18]

Cause-specific mortality rate: Number of deaths from a specific cause per 100,000 population per year. In 2015, the cause-specific mortality rate from heart disease in the United States was 197.2/100,000/year.[18]

Age-specific mortality rate: Total number of deaths from all causes among individuals in a specific age category per 100,000 population per year in the age category. In 2015, the age-specific death rate was 589.6/100,000/year among U.S. children under the age of 1 year.[18]

Years of potential life lost: The number of years that an individual was expected to live beyond his or her death. In 2015, a total of 957 years were lost from heart disease, 1,283 years were lost from cancer, and 1,172 were lost from unintentional injuries before age 75 per 100,000 population younger than 75 years of age in the United States.[18] The number of years of potential life lost reflects both the number of individuals who died of a particular cause and the age at which the death occurred. For example, a cause of death that is more common among children and young adults (such as unintentional injuries) will result in more years of life lost per individual than a cause of death that is common among the elderly (such as heart disease).

Livebirth rate: Total number of livebirths per 1,000 population per year. A livebirth is a pregnancy that results in a child who, after separation, breathes or shows any other evidence of life. Sometimes, the denominator includes only women of childbearing age. In 2015, the crude livebirth rate among women who were residents of the United States was 12.4/1,000/year.[18]

Infant mortality rate: Number of deaths of infants less than 1 year of age per 1,000 livebirths per year. This statistic is often divided into neonatal deaths (those occurring during the first 27 days following birth) and postneonatal deaths (those occurring from 28 days through 12 months). In 2014, the infant mortality rate in the United States was 5.8/1,000 livebirths/year, the neonatal mortality rate was 3.9/1,000 livebirths/year, and the postneonatal death rate was 1.9/1,000 livebirths/year.[17]

Birth defect rate *(also called congenital anomaly or malformation rate):* Number of children born with defects, usually per 10,000 births. The numerator and denominator often include both livebirths and stillbirths. In 2016–2017, the prevalence of brain malformations, including microcephaly, was 5% among women with evidence of recent possible Zika virus infection.[19]

Morbidity rate*:* Number of existing or new cases of a particular disease or condition per 100 population. The time period that is covered and the population size in the denominator vary. *Morbidity* is a general word that can apply to a disease, condition, or event. For example, from 2011 to 2014, the prevalence of physician-diagnosed diabetes among U.S. adults aged 65 years and over was 20.6%.[18]

Attack rate*:* Number of new cases of disease that develop (usually during a defined and short time period) per the number in a healthy population at risk at the start of the period. This cumulative incidence measure is usually reserved for infectious disease outbreaks. For example, the 24-hour attack rate for food poisoning was 50% among people who ate chicken salad at the banquet.

Case fatality rate*:* Number of deaths per number of cases of disease. Note that this measure is a type of cumulative incidence and therefore it is necessary to specify the length of time to which it applies. For example, in 2014 in the Democratic Republic of Congo, the 5-month case fatality rate among individuals with Ebola virus disease was 74.2%.[20]

Survival rate*:* Number of living cases per number of cases of disease. This rate is the complement of the case fatality rate and is also a cumulative incidence measure. Five-year relative survival rates for cancer compare people with a particular cancer to similar people in the general population. For example, from 2007 to 2013, 5-year relative survival rates for prostate cancer were 100% among men diagnosed while the tumor was still confined to the prostate or had spread only to the regional lymph nodes and 29.8% among men whose tumor had metastasized to distant sites.[14]

Summary

A population is defined as a group of people with a common characteristic, such as place of residence, age, or the occurrence of an event. There are two main types of populations, fixed and dynamic (or open). The membership of a fixed population is defined by a life event and is permanent, whereas the membership of a dynamic population is defined by a changeable characteristic and is transient.

Three factors should always be considered when measuring how commonly a disease occurs in a population: (1) the number of affected individuals or cases, (2) the size of the population from which the cases arise, and (3) the amount of time that this population is followed. Before epidemiologists can count the number of affected cases, they must compose a disease definition that is usually based on physical and pathological examinations, diagnostic tests, and signs and symptoms. Disease definitions often

change over time as more is learned about a disease and its manifestations. For example, the official case definition of HIV/AIDS expanded when its cause was discovered and improvements in detection were made.

Incidence and prevalence are the two basic measures of disease frequency. Incidence measures the occurrence of new disease and therefore captures the transition from health to disease. Cumulative incidence and incidence rate are the two main types of incidence measures. Cumulative incidence is defined as the proportion of a candidate population that becomes diseased over a specified time period. It is a dimensionless proportion that measures the average risk of contracting a disease over a certain time period. Incidence rate is the occurrence of new cases of disease that arise during person-time of observation; therefore, it is a true rate. It is important to remember that person-time accumulates only among candidates for disease. Cumulative incidence and incidence rate are related mathematically. Both measures are most useful for evaluating the effectiveness of disease-prevention activities and for etiologic studies of disease.

Prevalence measures existing disease and therefore focuses on the period when a person is ill. Prevalence measures the proportion of the total population that is diseased at a point in time or during a period of time. Its numerator consists of the number of existing cases, and its denominator includes the total population, including sick, healthy, at-risk, and immune individuals. Point prevalence refers to a single point in time and is like a snapshot. Period prevalence refers to a specific duration of time that may be derived from a series of snapshots. Prevalence is typically used for estimating the needs of medical facilities and allocating resources for treating diseased individuals. The incidence rate and prevalence are mathematically related.

Many measures of disease frequency are commonly used in public health, including the crude, cause-specific, and age-specific mortality rates; morbidity rate; livebirth rate; infant mortality rate; attack rate; case fatality rate; and survival rate. Note that the term *rate* is often incorrectly used to refer to proportions and ratios.

References

1. MacMahon B, Trichopoulos D. *Epidemiology: Principles and Methods.* 2nd ed. Boston, MA: Little, Brown and Co; 1996.

2. Porta M. *A Dictionary of Epidemiology.* 6th ed. New York, NY: Oxford University Press; 2014.

3. *Constitution of the World Health Organization.* who.int. http://www.who.int/governance/eb/who_constitution_en.pdf. Accessed June 8, 2018.

4. U.S. Department of Health and Human Services. *Healthy People 2010: Understanding and Improving Health.* 2nd ed. Washington, DC: U.S. Government Printing Office; 2000.

5. Centers for Disease Control and Prevention. Kaposi's sarcoma and *Pneumocystis* pneumonia among homosexual men—New York City and California. *MMWR.* 1981;30:305-308.

6. Centers for Disease Control and Prevention. Update on acquired immune deficiency syndrome (AIDS)—United States. *MMWR.* 1982;31:507-514.

7. Centers for Disease Control and Prevention. Revision of the case definition of acquired immune deficiency syndrome for national reporting—United States. *MMWR.* 1985;34:373-375.

8. Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Report.* 1997;9(2): 1-43. https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-1997-vol-9-2.pdf. Accessed June 8, 2018.

9. Centers for Disease Control and Prevention. Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged < 18 months and for HIV infection and AIDS among

children aged 18 months to < 13 years—United States, 2008. *MMWR*. 2008;57:1-12.

10. Centers for Disease Control and Prevention. Revised surveillance case definition for HIV infection—United States, 2014. *MMWR*. 2014;63:1-10.

11. New HIV Infections in the United States. December 2012. cdc.gov. https://www.cdc.gov/nchhstp/newsroom/docs/2012/hiv-infections-2007-2010.pdf. Accessed September 12, 2016.

12. Annual Estimates of the Resident Population by Sex, Race Alone or in Combination, and Hispanic Origin for the United States, States and Counties: April 1, 2010 to July 1, 2016. census.gov. https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk. Accessed December 2017.

13. Elandt-Johnson RC. Definition of rates: some remarks on their use and misuse. *Am J Epidemiol*. 1975;102:267-271.

14. Howlander N, Noone AM, Krapcho M, et al. (eds.). *SEER Cancer Statistics Review, 1975-2014*. cancer.gov. https://seer.cancer.gov/csr/1975_2014/. Published April 2017. Accessed December 2017.

15. Liebow AA. *Encounter with Disaster: A Medical Diary of Hiroshima*. New York, NY: W. W. Norton & Co; 1970.

16. Ozasa K, Shimizu Y, Suyama A, et al. Studies of the mortality of atomic bomb survivors, report 14, 1950-2003: an overview of cancer and noncancer diseases. *Radiat Res*. 2012;229-243.

17. Cozier YC, Yu J, Coogan PF, Bethea TN, Rosenberg L, Palmer JR. Racism, segregation, and risk of obesity in the Black Women's Health Study. *Am J Epidemiol*. 2014;179:875-883.

18. Centers for Disease Control and Prevention. *Health, United States, 2016, With Chartbook on Long-term Trends in Health*. https://www.cdc.gov/nchs/data/hus/hus16.pdf. Accessed June 8, 2018.

19. Shapiro-Mendoza CK, Rice ME, Galang RR, et al. Pregnancy outcomes after maternal Zika virus infection during pregnancy—U.S. territories, January 1, 2016–April 25, 2017. *MMWR*. 2017;66:615-621.

20. Nanclares NC, Kapetshi J, Lionetto F, de la Rosa O, Tamfun JJ, Alia M, et al. Ebola virus disease, Democratic Republic of the Congo, 2014. *Emerg Infect Dis*. 2016;22:1579-1586.

Chapter Questions

1. What measure of disease frequency is each of the following?
   a. The percentage of freshman girls who become pregnant over the course of their high school years
   b. The percentage of senior boys who are fathers at the time of graduation
   c. The number of live-born babies who die of sudden infant death syndrome during the first year of life per 100,000 baby-years of follow-up
   d. The percentage of infants weighing less than 2500 grams at birth
   e. The lifetime risk of breast cancer

2. Briefly describe the main similarities and differences between each of the following:
   a. Prevalence and incidence
   b. Incidence rate and cumulative incidence
   c. Fixed and dynamic populations

3. What are the lowest and highest possible values of each of the following measures of disease frequency?
   a. Prevalence
   b. Cumulative incidence
   c. Incidence rate

4. Suppose that there were 2900 new cases of breast cancer diagnosed among women in Boston, Massachusetts, and 200 new cases diagnosed among women in Anchorage, Alaska, in 2017. Based on these data, is it accurate to say that the incidence rate of breast cancer is higher in Boston than Anchorage? Why or why not?

5. A study of 100 injection drug users who tested negative for HIV infection at enrollment had their HIV status retested at 3-month intervals over a 2-year follow-up period. All of the injection drug users were followed for the entire 2-year period. None died and none were lost to follow-up. Which of the following frequency measures of HIV infection can be calculated at the end of the study?

a. Prevalence
b. Cumulative incidence
c. Incidence rate
d. All of the above

6. Consider a class with 100 enrolled students. None of the students were ill at the beginning of the school year. On September 30, a total of 5 students reported having gastroenteritis. All 5 continued to be ill on October 1, but all 5 recovered within 3 days. On October 14, another 3 students developed gastroenteritis. All of these students continued to be ill on October 15, but all 3 recovered 5 days later. In this example, assume that a person cannot get gastroenteritis more than once.
   a. Calculate the prevalence of gastroenteritis in the class on October 1.
   b. Calculate the prevalence of gastroenteritis in the class on October 30.
   c. Calculate the cumulative incidence of gastroenteritis in the class during the month of October.

7. The incidence rate of a nonfatal disease is 500/100,000 person-years. People usually have the disease for an average of 3 years, at which time the disease resolves spontaneously. Estimate the prevalence of this disease using this information. Assume that the population is in steady state.

8. A population of 100 healthy men was followed for the development of prostate cancer. After being followed for 5 years, 20 men developed prostate cancer. Another 10 men were followed for 1 year and then were lost. The remaining men who never developed the disease were followed for 10 years. Calculate the number of person-years of observation accrued by this population.

9. Consider the following hypothetical data on the occurrence of hepatitis in two cities:

City	New cases	Observation period	Starting population at risk
City A	25	January–December 2017	25,000
City B	30	January–December 2017	50,000

   a. Calculate the cumulative incidence of hepatitis in each city.
   b. Which city has the higher cumulative incidence?

10. A total of 60 cases of myocardial infarction were reported over a period of 2 years in a city with a population of 100,000 people. Using these data, estimate the incidence rate of myocardial infarction per 100,000 person-years. State any assumptions that are needed.

11. The incidence rate of postpartum depression among 250,000 women who recently experienced a pregnancy was 12 cases per 100,000 woman-years of follow-up. Exactly how many incident cases of postpartum depression developed in this population?

12. State the type of population (fixed or dynamic) that best describes each of the following:
   a. People who live in New York City
   b. Male residents of Paris who had coronary bypass surgery between 2010 and 2017
   c. Children residing in California who were vaccinated against polio in 1955
   d. Women who are practicing physicians in the United States

13. How does each of the following conditions influence the prevalence of a disease in a population? For each scenario, assume that no other changes occur. Your choices are increases prevalence, decreases prevalence, or has no effect on prevalence.

    a.   A cheap new clinical test becomes widely available that allows doctors to diagnose previously latent (i.e., hidden) disease.

    b.   A new treatment is developed that cures people of the disease very soon after they are diagnosed.

    c.   There is migration of a large number of healthy people into the population.

14. Indicate whether the following statements are true or false:

    a.   Only the population at risk contributes to the denominator of the cumulative incidence.

    b.   When calculating the incidence rate of a disease, it is necessary to follow all subjects for the same length of time.

    c.   If the incidence rate of a very serious disease is 75/100,000 person-years and the prevalence of this disease in the population is 25/100,000, then the average duration of this disease must be 3 years.

    d.   All other things being equal, when a new prevention measure for a disease is developed, the prevalence of the disease will decrease over time.

    e.   All other things being equal, when a treatment is developed that prolongs the life of people suffering from a disease, the prevalence of the disease will increase over time.

15. An epidemiological investigation that was started on January 1, 2017, identified a population of 1,000 individuals among whom 4 were found to have the disease under study. During the year of the study, 6 new cases were found. Among the total of 10 cases, there were 6 deaths during the year. For the 10 cases, the diagram indicates the time of case recognition, periods of observation during the study, and vital status at the time of the termination of observation. An arrow at the start of the diagram (subjects 1, 2, 3, 4) indicates that the start of disease occurred before the study began. Assume that the 990 remaining individuals in the study did not become ill or die during the year of observation. From the information and diagram given, calculate the following:

    a.   Prevalence of the disease on January 1, 2017; July 1, 2017; and December 31, 2017

    b.   Cumulative incidence of disease during 2017

    c.   Cumulative incidence of death during 2017