CHAPTER 3

Measures of Morbidity and Mortality Used in Epidemiology

LEARNING OBJECTIVES

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

- define and distinguish among ratios, proportions, and rates
- explain the term population at risk
- identify and calculate commonly used rates for morbidity, mortality, and natality
- state the meanings and applications of incidence rates and prevalence
- discuss limitations of crude rates and alternative measures for crude rates
- apply direct and indirect methods to adjust rates
- explain when either direct or indirect rate adjustment should be used

CHAPTER OUTLINE

I. Introduction
II. Definitions of Count, Ratio, Proportion, and Rate
III. Risk Versus Rate; Cumulative Incidence
IV. Interrelationship Between Prevalence and Incidence
V. Applications of Incidence Data
VI. Crude Rates
VII. Specific Rates and Proportional Mortality Ratio
VIII. Adjusted Rates
IX. Conclusion
X. Study Questions and Exercises

Introduction

The 2009 H1N1 influenza pandemic illustrated how a potentially deadly virus could spread rapidly from the United States to other countries worldwide. At one point in the growing epidemic, public health officials pondered whether the 2009 pandemic was a repeat of the 1918 "killer" flu. When a flu outbreak occurs, what quantitative measures inform public health professionals that an epidemic caused by a killer virus is occurring? How fast is the virus spreading? How many deaths is the potentially virulent and lethal agent causing? In order to answer questions such as these, the work of the epidemiologist involves enumerating cases of diseases and health-related phenomena as well as describing the occurrence and patterns of disease in the population. Epidemiology examines risk factors associated with adverse health outcomes and identifies potential causal associations between exposures and diseases.

This chapter explains disease occurrence measures used commonly in public health practice for quantifying health outcomes. The foundation of studies designed to identify etiology, monitor trends, and evaluate public health interventions rests on the bedrock of our ability to measure the occurrence of morbidity and mortality carefully and accurately. This chapter defines four categories of epidemiologic measures (counts, ratios, proportions, and rates), differentiates between the concepts of risk and rate, discusses relationships among measures, and illustrates their applications.

Definitions of Count, Ratio, Proportion, and Rate

The four types of epidemiologic measures covered in this section are counts, ratios, proportions, and rates. Refer to **Figure 3–1** for an overview of the measures discussed in this chapter. The figure identifies the measures and indicates their hierarchy and interrelationships.

Count

The simplest and most frequently performed quantitative measure in epidemiology is a *count*. As the term implies, a count refers merely to the number of cases of a disease or other health phenomenon being studied. Several examples of counts are the number of:

- cases of influenza reported in Westchester County, New York, during January of a particular year
- traffic fatalities in the borough of Manhattan during a 24-hour period

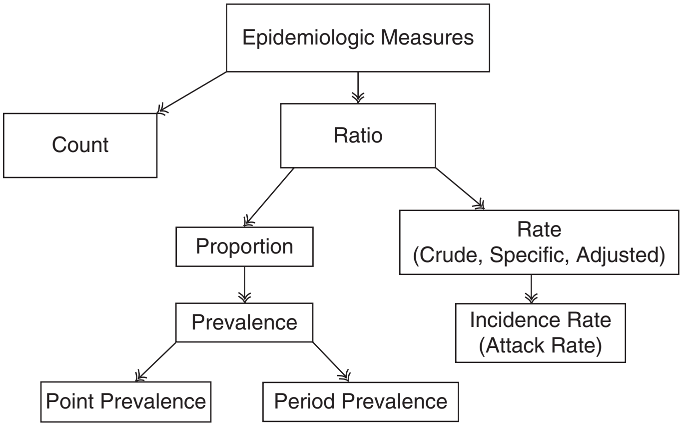


FIGURE 3–1 Overview of epidemiologic measures.

- participants screened positive in a hypertension screening program organized by an industrial plant in northern California
- college dorm residents who had mono
- stomach cancer patients who were foreign born

Ratio

A *ratio* is defined as "[t]he value obtained by dividing one quantity by another. RATE, PROPORTION, and PERCENTAGE are types of ratios."[1] A ratio therefore consists of a numerator and a denominator. The most general form of a ratio does not necessarily have any specified relationship between the numerator and denominator. A ratio may be expressed as follows: ratio = X/Y. An example of a ratio is the sex ratio, which is shown in three variations:

1. **Simple sex ratio**: Of 1,000 motorcycle fatalities, 950 victims are men and 50 are women. The sex ratio for motorcycle fatalities is:

$$\frac{\text{Number of male cases}}{\text{Number of female cases}} = \frac{950}{50} = 19{:}1 \text{ male to female}$$

2. **Demographic sex ratio**: This ratio refers to the number of males per 100 females. In the United States (2010), the sex ratio for the entire population was 96.7, indicating more females than males.

$$\text{Sex ratio} = \frac{\text{Number of males}}{\text{Number of females}} \times 100 = \frac{151{,}781{,}326}{156{,}964{,}212} \times 100 = 96.7$$

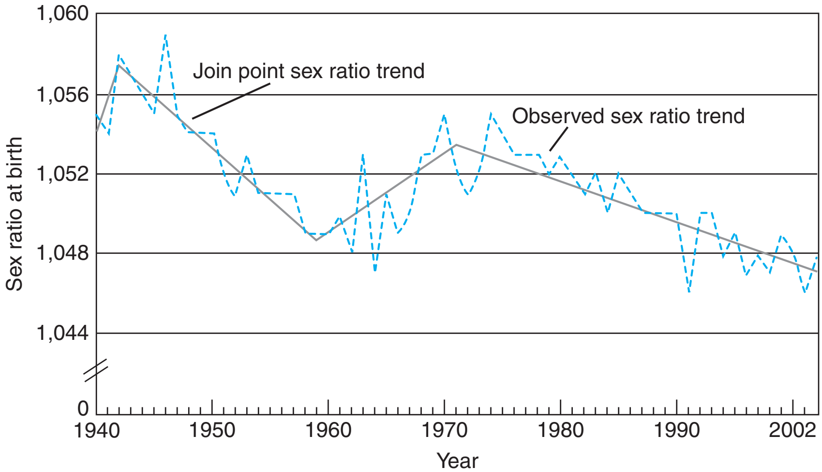
3. **Sex ratio at birth**: the sex ratio at birth is defined as: (the number of male births divided by the number of female births) multiplied by 1,000.

$$\text{Sex ratio at birth} = \frac{\text{Number of male births}}{\text{Number of female births}} \times 1{,}000$$

Figure 3–2 shows that between 1940 and 2002, the sex ratio at birth exceeded 1,000 and made significant transitions in 1942, 1959, and 1971. Nevertheless, the sex ratio trended downward since 1940.

Proportion

A *proportion* is a type of ratio in which the numerator is part of the denominator; proportions may be expressed as percentages. Let us consider how a proportion can be helpful in describing health issues by reexamining a count. For a count to be descriptive of a group, it usually should be seen relative to the size of the group. Suppose there were 10 college dorm residents who had hepatitis. How large a problem did these 10 cases represent? To answer this question, one would need to know whether the dormitory housed 20 students or 500 students. If there were only 20 students, then 50% (or 0.50) were ill. Conversely, if there were 500 students in the dormitory, then only 2% (or 0.02) were ill. Clearly,



NOTES: Sex ratio at birth is the number of male births divided by the number of female births multiplied by 1,000.

FIGURE 3–2 Sex ratio at birth, 1940–2002. *Source:* Reproduced from TJ Mathews, BE Hamilton. Trend Analysis of the Sex Ratio at Birth in the United States. *National Vital Statistics Reports*, Vol. 55, No. 20, p. 1. Hyattsville, MD: National Center for Health Statistics; 2005.

these two scenarios paint a completely different picture of the magnitude of the problem. In this situation, expressing the count as a proportion is indeed helpful.

Table 3–1 illustrates the calculation of the proportion of African-American male deaths among African-American and white boys aged 5–14 years.

In most situations, it will be informative to have some idea about the size of the denominator. Although the construction of a proportion is straightforward, one of the central concerns of epidemiology is to find and enumerate appropriate denominators to describe and compare groups in a meaningful and useful way.

The previous discussion may leave the reader with the impression that counts, in and of themselves, are of little value in epidemiology; this is not true, however. In fact, case reports of patients with particularly unusual presentations or combinations of symptoms often spur epidemiologic investigations. In addition, for some diseases even a single case is sufficient to be of public health importance. For example, if a case of smallpox or Ebola virus were reported, the size of the denominator would be irrelevant. That is, in these instances a single case, regardless of the size of the population at risk, would stimulate an investigation.

Rate

A *rate* also is a type of ratio; however, a rate differs from a proportion because the denominator involves a measure of time. The numerator consists of the frequency of a disease over a specified period of time, and the denominator is a unit size of population (**Exhibit 3–1**). It is critical to remember that to calculate a rate, two periods of time are involved: the beginning of the period and the end of the period.

Medical publications may use the terms ratio, proportion, and rate without strict adherence to the mathematical definitions for these terms. Hence, one must be alert to how a measure is defined and calculated.[2] In the formula shown in Exhibit 3–1, the denominator also is termed the reference population and by

Table 3–1 Calculation of the Proportion of African-American Male Deaths Among African-American and White Boys Aged 5 to 14 Years

A	B	Total (A + B)
Number of deaths among African-American boys	Number of deaths among white boys	Total
1,150	3,810	4,960

Proportion = A/(A + B) × 100 = (1,150/4,960) × 100 = 23.2%

EXHIBIT 3–1

Rate Calculation

Rate: A ratio that consists of a numerator and a denominator and in which time forms part of the denominator.

Epidemiologic rates contain the following elements:

- disease frequency
- unit size of population
- time period during which an event occurs

Example:

$$\text{Crude death rate} = \frac{\text{Number of deaths in a given year}}{\begin{array}{c}\text{Reference population}\\ \text{(during midpoint of the year)}\end{array}} \times 100{,}000$$

(Either rate per 1,000 or 100,000 is used as the multiplier)

Calculation problem (crude death rate in the United States):

Number of deaths in the United States during 2007 = 2,423,712
Population of the United States as of July 1, 2007 = 301,621,157

$$\text{Crude death rate} = \frac{2{,}423{,}712}{301{,}621{,}157} \times 100{,}000 = 803.6 \text{ per } 100{,}000$$

definition is the population from which cases of a disease have been taken. For example, in calculating the annual death rate (crude mortality rate) in the United States, one would count all the deaths that occurred in the country during a certain year and assign this value to the numerator. The value for the denominator would be the size of the population of the country during a particular year. The best estimate of the population would probably be the population around the midpoint of the year, if such information could be obtained. Referring to Exhibit 3–1, one calculates the U.S. crude mortality rate as 803.6 per 100,000 persons for 2007.

Rates improve one's ability to make comparisons, although they also have limitations. Rates of mortality or morbidity for a specific disease (see the section on cause-specific mortality rates later in this chapter) reduce that standard of comparison to a common denominator, the unit size of population. To illustrate, the U.S. crude death rate for diseases of the heart in 2003 was 235.6 per 100,000.

One also might calculate the heart disease death rate for geographic subdivisions of the country (also expressed as frequency per 100,000 individuals). These rates could then be compared with one another and with the rate for the United States to judge whether the rates found in each geographic area are higher or lower. For example, the crude death rates for diseases of the heart in New York and Texas were 288.0 and 188.9 per 100,000, respectively. It would appear that the death rate is higher in New York than in Texas based on the crude death rates. This may be a specious conclusion, however, because there may be important differences in population composition (e.g., age differences between populations) that would affect mortality experience. Later in this chapter, the procedure to adjust for age differences or other factors is discussed.

Rates can be expressed in any form that is convenient (e.g., per 1,000, per 100,000, or per 1,000,000). Many of the rates that are published and routinely used as indicators of public health are expressed in particular conventions. For example, cancer rates are typically expressed per 100,000 population, and infant mortality is expressed per 1,000 live births. One of the determinants of the size of the denominator is whether the numerator is large enough to permit the rate to be expressed as an integer or an integer plus a trailing decimal (e.g., 4 or 4.2). For example, it would be preferable to describe the occurrence of disease as 4 per 100,000 (or 4.2 per 100,000) rather than 0.04 per 1,000 (or 0.042 per 1,000), even though both are perfectly correct. Throughout this chapter, the multiplier for a given morbidity or mortality statistic is provided.

Exhibit 3–2 describes the Iowa Women's Health Study (IWHS). The data collected illustrate the various measures of disease frequency defined in this chapter.

Prevalence

The term *prevalence* refers to the number of existing cases of a disease or health condition in a population at some designated time.[1] As shown in **Figure 3–3**, prevalence is analogous to water that has collected in a pool at the base of a waterfall. Prevalence data provide an indication of the extent of a health problem and thus may have implications for the scope of health services needed in the community. Prevalence can be expressed as a number, a percentage, or number of cases per unit size of population. Consider three examples: The prevalence of diarrhea in a children's camp on July 13 was 15, the prevalence of phenylketonuria-associated mental disabilities in institutions for the developmentally disabled was 15%, and the prevalence of obesity among women aged 55–69 years was 367 per 1,000. These examples illustrate that the designated time can be specified (e.g., one day) or unspecified. When the time period is unspecified,

EXHIBIT 3–2

The Iowa Women's Health Study

The IWHS is a longitudinal study of mortality and cancer occurrence in older women.[3,4] The state of Iowa was chosen as the site of this study because of the availability of cancer incidence and mortality data from the State Health Registry of Iowa. This registry is a participant in the National Cancer Institute's Surveillance, Epidemiology, and End Results Program. The sample was selected from a January 1985 current drivers list obtained from the Iowa Department of Transportation. The list contained the names of 195,294 women aged 55 to 69 and represented approximately 94% of the women in the state of Iowa in this age range.

In December 1985, a 50% random sample of the eligible women was selected, yielding 99,826 women with Iowa mailing addresses. A 16-page health history questionnaire was mailed on January 16, 1986, followed by a reminder postcard one week later and a follow-up letter four weeks later; a total of 41,837 women responded. Information was collected about basic demographics, medical history, reproductive history, personal and family history of cancer, usual dietary intake, smoking and exercise habits, and medication use. A paper tape measure also was provided along with detailed instructions for the subject to record selected body measurements: height, weight, and circumferences of the waist and hips.

The primary focus of the study was to determine whether distribution of body fat centrally (i.e., around the waist) rather than peripherally (i.e., on the hips) is associated with increased risk of cancer. The occurrence of cancer was determined by record linkage with the State Health Registry. A computer program was used to match new cancer cases in the registry with study participants on name, ZIP code, birth date, and Social Security number.

prevalence usually implies a particular point in time. More specifically, these examples refer to *point prevalence*.

A second type of prevalence measure is *period prevalence*, which denotes the total number of cases of a disease that exist during a specified period of time, for instance a week, month, or longer time interval. To determine the period prevalence, one must combine the number of cases at the beginning of the time interval (the point prevalence) with the new cases that occur during the interval. Because the denominator may have changed somewhat (the result of people

FIGURE 3–3 Analogy of prevalence and incidence. The water flowing down the waterfall symbolizes incidence and the water collecting in the pool at the base symbolizes prevalence.

$$\text{Point prevalence} = \frac{\text{Number of persons ill}}{\text{Total number in the group}} \text{ at a time point}$$

Example: In the IWHS, respondents were asked: "Do you smoke cigarettes now?" The total number in the group was 41,837. The total number who responded yes to the smoking question was 6,234. Therefore, the prevalence of current smokers in the IWHS on January 16,1985, was 6,234/41,837. This result could be expressed as a percentage (14.9%) or as a frequency per 1,000 (149.0).

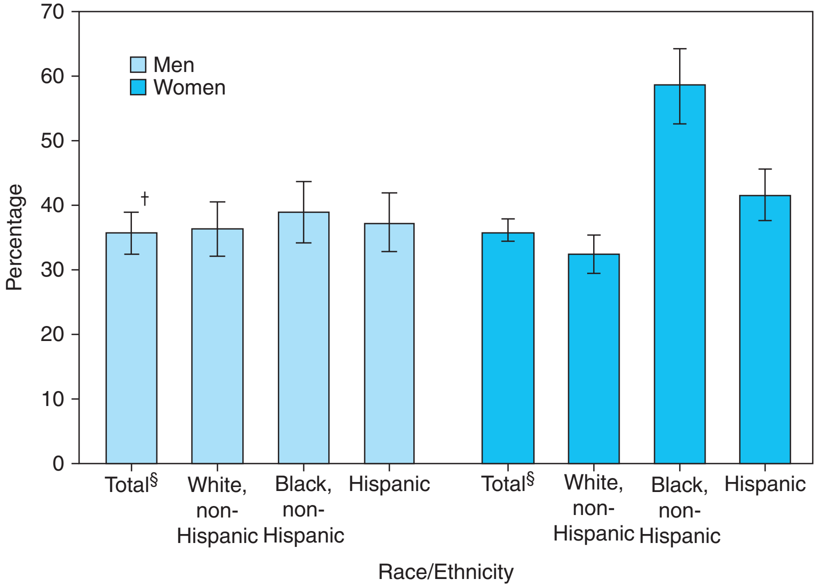
$$\text{Period prevalence} = \frac{\text{Number of persons ill}}{\text{Average population}} \text{ during a time period}$$

Example: In the IWHS, women were asked: "Have you ever been diagnosed by a physician as having any form of cancer, other than skin cancer?" Note that the question did not ask about current disease but rather about the lifetime history. Thus, it refers to period prevalence, the period being the entire life span. To calculate the period prevalence, one needs to know the average population (still 41,837) and the number who responded yes to the question (2,293). Therefore, the period prevalence of cancer in the study population was 2,293/41,837, or 5.5%.

entering or leaving during the period of observation), one typically refers to the average population. Note that for period prevalence, cases are counted even if they die, migrate, or recur as episodes during the period.

A second example is from the National Center for Health Statistics—the National Health and Examination Survey, United States, 2009–2010. **Figure 3–4** shows the prevalence of obesity among adults 20 years of age or older. Obesity was defined as a body mass index (BMI) of 30 or greater. The prevalence (percentage) was 58.5% among non-Hispanic black women.

Technically speaking, both point and period prevalence are proportions. As such, they are dimensionless and should not be described as rates, a mistake that is commonly made. To illustrate the distinction between point and period prevalence, consider as an example the issue of homelessness in the United States. The conditions surrounding homelessness present a serious public health problem, particularly in the control of infectious diseases and the effect on homeless persons' physical and mental health. Consequently, public officials have a legitimate need to estimate the magnitude of the problem, an issue that has produced intense debate. Surveys of currently homeless people pose extremely challenging methodologic difficulties that have led some authorities to believe that point prevalence may lead to serious underreporting. According to Link and colleagues, "The first problem is finding people who are currently homeless. Surveys may miss the so-called hidden homeless, who sleep in box cars, on the roofs of tenements, in campgrounds, or in other places that researchers cannot effectively search. [Even if located] . . . respondents may refuse to be interviewed or deliberately hide the fact that they are homeless."[5(p 1907)] People who experience relatively short or intermittent episodes of literal homelessness are likely



† 95% confidence interval.
§ Includes other races (i.e., Asians and American Indians/Alaska natives) not shown seperately because of small sample sizes, which affect reliability of estimates.

FIGURE 3–4 Prevalence of obesity* among adults aged ≥ 20 years, by race/ethnicity and sex—National Health and Nutrition Examination Survey, United States, 2009–2010. *Source:* Reproduced from Centers for Disease Control and Prevention. QuickStats. *MMWR.* Vol 61, No. 7, p. 130, February 24, 2012.

*Defined as a body mass index (weight [kg]/height [m]^2) ≥ 30.

to be missed in brief surveys. To address these problems, Link et al.[5] conducted a national household telephone survey to provide lifetime and five-year period prevalence estimates. They found that 14% of the sample had ever been homeless, 4.6% in the last five years. Compared with previous estimates based on point prevalence, the investigators concluded that the magnitude of the problem was much greater than previous estimates had indicated.

Prevalence studies are useful in describing the health burden of a population and in allocation of health resources, such as facilities and personnel. The foregoing data on the prevalence of smoking, obesity, and homelessness were illustrations. Also, epidemiologists use prevalence data to estimate the frequency of an exposure in a population. They can survey a sample of respondents in order to determine the types of exposures (e.g., use of drugs, medications, or other types of exposures) they have had; in other cases environmental researchers can make direct measures of toxic contaminants through environmental monitoring.

Typically, prevalence studies are not as helpful as other types of epidemiologic research designs for studies of etiology. Among several reasons, the most important is the possible influence of differential survival. That is, for a case to be included in a prevalence study, he or she would have had to survive the disease long enough to participate. Cases that died before participation would obviously be missed, resulting in a truncated sample of eligible cases. Risk factors for rapidly fatal cases may be quite different from risk factors for less severe manifestations. One situation in which the use of prevalent cases may be justified for studies of disease etiology arises when a condition has an indefinite time of onset, such as occurs with mental disorders.[2]

Incidence Rate

Incidence is defined as "[t]he number of instances of illnesses commencing, or of persons falling ill, during a given period in a specified population. More generally, the number of new health-related events in a defined population within a specified period of time. It may be measured as a frequency count, a rate, or a proportion."[1] Incidence is a measure of the risk of a specified health-related event. (We will explain this concept later in the chapter.) In Figure 3–3 incidence is analogous to water flowing in the waterfall (new cases). An example of incidence measured as a frequency is the number of new cases of HIV infection diagnosed in a population in a given year: A total of 164 HIV diagnoses were reported among American Indians or Alaska natives in the United States during 2009.

The term *incidence rate* describes the rate of development of a disease in a group over a certain time period; this period of time is included in the denominator. An incidence rate (**Exhibit 3–3**) includes three important elements:

1. a numerator: the number of new cases
2. a denominator: the population at risk
3. time: the period during which the cases accrue

Number of New Cases

The incidence rate uses the frequency of new cases in the numerator. This means that individuals who have a history of the disease are not included.

Population at Risk

The denominator for incidence rates is the population at risk. One therefore should exclude individuals who have already developed the disease of interest (e.g., those who have had heart attacks) or are not capable of developing

EXHIBIT 3–3

Incidence Rate

$$\text{Incidence rate} = \frac{\text{Number of new cases over a time period}}{\text{Total population at risk during the same time period}} \times \underset{\text{(e.g., 100,000)}}{\text{Multiplier}}$$

The denominator consists of the population at risk (i.e., those who are at risk for contracting the disease).

Example: Calculate the incidence rate of postmenopausal breast cancer in the IWHS. The population at risk in this example would not include women who were still premenopausal (n = 569), women who had had their breasts surgically removed (n = 1,870), and women with a previous diagnosis of cancer (n = 2,293). Thus, the denominator is 37,105 women. After eight years of follow-up, 1,085 cases were identified through the State Health Registry. The incidence rate is therefore 1,085/37,105 per eight years. To express this rate per 100,000 population: Divide 1,085 by 37,105 (answer: 0.02924). This is the rate over an eight-year period. For the annual rate, divide this number by eight years (answer: 0.003655) and multiply by 100,000.

Answer: 365.5 cases of postmenopausal breast cancer per 100,000 women per year.

the disease. For example, if one wanted to calculate the rate of ovarian cancer in the IWHS, women who had had their ovaries removed (oophorectomized women) should be excluded from the cohort at risk. It is not uncommon, however, to see some incidence rates based on the average population as the denominator rather than the population at risk. This distinction really must be made for those infectious diseases that confer lifetime immunity against recurrence. Regarding chronic diseases to which most people appear to be susceptible, the distinction is less critical. The population at risk may include those exposed to a disease agent or unimmunized or debilitated people, or it may consist of an entire population (e.g., a county, a city, or a nation). The population at risk may represent special risk categories; occupational injury and illness incidence rates are calculated for full-time workers in various occupations, for example, because these are the populations at risk.

Specification of a Time Period

The definition of incidence entails the designation of a time period, such as a week, a month, a year, or a multiyear time period. To determine an incidence rate, one must be able to specify the date of onset for the condition during the time period. Some acute conditions (e.g., a severe stroke or an acute myocardial infarction) may have a readily identifiable time of onset. Other conditions (e.g., cancer) may have an indefinite time of onset, which is defined by the initial definitive diagnosis date for the disease.[6]

Attack Rate

The *attack rate* (AR) is an alternative form of the incidence rate that is used when the nature of the disease or condition is such that a population is observed for a short time period, often as a result of specific exposure.[2] In reporting outbreaks of *salmonella* infection or other foodborne types of gastroenteritis, epidemiologists employ the AR. The formula for the AR is:

$$AR = \text{Ill}/(\text{Ill} + \text{well}) \times 100 \text{ (during a time period)}$$

Calculation example: a total of 87 people at a holiday dinner ate roast turkey.
Among these persons, 63 who consumed roast turkey became ill; the remainder did not become ill.

$$AR \text{ (for the roast turkey)} = 63/(63 + 24) \times 100 = 72.4\%$$

As shown in this formula, the numerator consists of people made ill as a result of exposure to the suspected agent, and the denominator consists of all people, whether well or ill, who were exposed to the agent during a time period. Strictly speaking, the AR is not a true rate because the time dimension is often uncertain or specified arbitrarily.

Although the AR often is used to measure the incidence of disease during acute infectious disease epidemics, it also may be used for the incidence of other conditions where the risk is limited to a short time period or the etiologic factors operate only within certain age groups. An example is hypertrophic pyloric stenosis (a blockage from the stomach to the intestines), which occurs predominantly in the first three months of life and is practically unknown after the age of six months.

Risk Versus Rate; Cumulative Incidence

Epidemiologists have been known to use the terms risk and rate interchangeably. However, if pressed to explain the difference, they would be able (one hopes) to identify several key distinctions. First, risk is a statement of the probability or chance that an individual will develop a disease over a specified period, conditioned on that individual's not dying from any other disease during the period.[7] As such, risk ranges from 0 to 1 and is dimensionless. Statements of risk also require a specific reference period, for example, the five-year risk of developing asthma.

Cumulative Incidence

Cumulative incidence refers to "[t]he number . . . of a group (cohort) of people who experience the onset of a health-related event during a specified time interval."[1] Cumulative incidence is used when all individuals in the population are thought to be at risk of the health-related event being investigated, as in a prospective cohort study in which the population is fixed. The cumulative incidence estimates the risk of a particular health-related outcome in the cohort. If it is possible to follow up every individual in the cohort during a given time period, then the cumulative incidence is the number of events that occur during that time period expressed relative to the denominator. (However, as we will describe later, a problem arises in determining cumulative incidence and incidence when individuals are observed for different periods of time.)

The illustration regarding the incidence of postmenopausal breast cancer in the IWHS is an example of a cumulative incidence. Because the population is fixed, no individuals are allowed to enter the denominator after the start of the observation period, and the numerator can include only individuals who were members of that fixed population. Calculation of cumulative incidence also requires that disease status be determined for everyone in the denominator. That is, once a group of individuals is selected for follow-up for disease occurrence, subsequent information about the occurrence of disease is obtained for everyone selected, which is difficult to achieve even in the best of circumstances. Most of the regions where we live and work contain dynamic populations; people move into and out of the area. Some individuals who were not in the study population at the baseline period may move into the region and become ill. Thus, the numerator has increased but the denominator has not. Conversely, if an individual moves away and then develops the disease, he or she would be counted in the denominator but not in the numerator. One solution to the problem of geographic mobility and loss to follow-up is to use rates as an indicator of risk. A simple perspective is that groups with high rates of disease are at greater risk than are groups with low rates of disease. The issue is a bit more complicated than that perspective (and beyond the scope of this text). The main caveat is that rates can be used to estimate risk only when the period of follow-up is short and the rate of disease over that interval is relatively constant. Thus, to estimate small risks, one simply multiplies the average rate times the duration of follow-up.[8]

Incidence Density

The *incidence density* is "[t]he average person-time incidence rate."[1] This variation in the incidence rate is calculated by using the person-time of observation as the denominator. Person-time ". . . is the sum of the periods of time at risk for each of the subjects. The most widely used measure is person-years."[1] Person-time is used when the amounts of time of observation of each of the subjects in the study varies instead of remaining constant for each subject.

Here is an example of how incidence density becomes useful. A special problem occurs when a population or study group is under observation for different lengths of time. This may occur for a variety of reasons, including attrition or dropout, mortality, or development of the disease under study. An illustration is the calculation of the incidence of postmenopausal breast cancer in the IWHS. Although the study was able to identify all cancers diagnosed within the state, some women may have moved out of state after the initial questionnaire administration. Any cancers diagnosed among these women would be unknown to the investigators. Other women died before the end of the follow-up period. In the previous calculation, we merely counted the number of cases over the 8-year period of follow-up (n = 1,042) and divided by the number of women at risk (n = 37,105). The implicit assumption of this calculation is that each of the 37,105 women was "observed" for the full 8-year period. Clearly, this could not be the case. To allow for varying periods of observation of the subjects, one uses a modification of the formula for the incidence rate in which the denominator becomes person-time of observation. Incidence density is defined in **Exhibit 3–4**.[2] An example of how to calculate person-years, the most common measure of person-time, is shown in **Table 3–2**.

In Table 3–2, person-years were derived simply by summing the product of each category of length of observation and the number of subjects in the category. A more difficult issue is how one actually determines the length of observation for each individual. Visiting again the IWHS example, a computer program

EXHIBIT 3–4

Incidence Density

$$\text{Incidence density} = \frac{\text{Number of new cases during the time period}}{\text{Total person-time of observation}}$$

When the period of observation is measured in years, the formula becomes:

$$\text{Incidence density} = \frac{\text{Number of new cases during the time period}}{\text{Total person-years of observation}}$$

Example: In the IWHS, the 37,105 women at risk for postmenopausal breast cancer contributed 276,453 person-years of follow-up. Because there were 1,085 incident cases, the rate of breast cancer using the incidence density method is 1,085/276,453 = 392.5 per 100,000 per year.

Note that had each woman been followed for the entire eight-year period of follow-up, the total person-years would have been 296,840. Because the actual amount of follow-up was 20,000 person-years less than this, the estimated rate of breast cancer was higher (and more accurate) using the incidence density method.

Table 3–2 Person-Years of Observation for Hypothetical Study Subjects in a 10-Year Heart Disease Research Project

A Number of Subjects	B Length of Observation (Years)	A × B Person-Years
30	10	300
10	9	90
7	8	56
2	7	14
1	1	1
Total 50		461

Number of health events (heart attacks) observed during the 10-year period: 5.

Incidence density = (5/461) × 100 = 1.08 per 100 person-years of observation.

was used to tabulate, for each individual, the amount of time that elapsed from receipt of the mailed questionnaire until the occurrence of one of the following events (listed in order of priority): breast cancer diagnosis, death (if in Iowa), a move out of Iowa (if known through the National Change of Address Service), midpoint of interval between date of last contact and December 30, 1993, or midpoint of interval between date of last contact and date of death (for deaths that occurred out of Iowa, identified through the National Death Index). Women who did not experience any of these events were assumed to be alive in Iowa and contributed follow-up until December 30, 1993. This real-life example illustrates that actual computation of person-years, although conceptually straightforward, can be a fairly complicated procedure.

Interrelationship Between Prevalence and Incidence

> Interrelationship: $P \cong ID$
>
> The prevalence (P) of a disease is proportional to the incidence rate (I) times the duration (D) of a disease.

For conditions of short duration and high incidence, one may infer from this formula that, when the duration of a disease becomes short and the incidence is high, the prevalence becomes similar to incidence. For diseases of short duration, cases recover rapidly or are fatal, eliminating the build-up of prevalent cases. In fact that is the case for infectious diseases of short duration, such as the common cold.

Typically chronic diseases have a low incidence and, by definition, long duration; as the duration of the disease increases, even though incidence is low or stable, the prevalence of the disease increases relative to incidence. An example is HIV/AIDS prevalence as shown in **Figure 3–5**. The line for HIV prevalence is much higher than the line for HIV incidence and shows an increasing trend. The explanation is that the prevalence of HIV is increasing gradually (about 1.1 million cases in 2006); however, the annual incidence of HIV in the United States has remained stable (slightly fewer than about 60,000 cases each year).

Figure 3–6 illustrates a second example of the relationship between incidence and prevalence. Suppose that there is an outbreak of meningococcal disease in a summer school class of 10 students. The frequency of the disease is recorded for 2 weeks. Individual cases plotted by the duration of each case for the period July 1–July 14 are shown in Figure 3–6. For the 10-day period (July 5–July 14), the period prevalence of meningococcal disease was 8/10; the point prevalence of disease on July 5 was 5/10. Because the disease in this example is one that can affect individuals more than once (no lifetime immunity after initial infection),

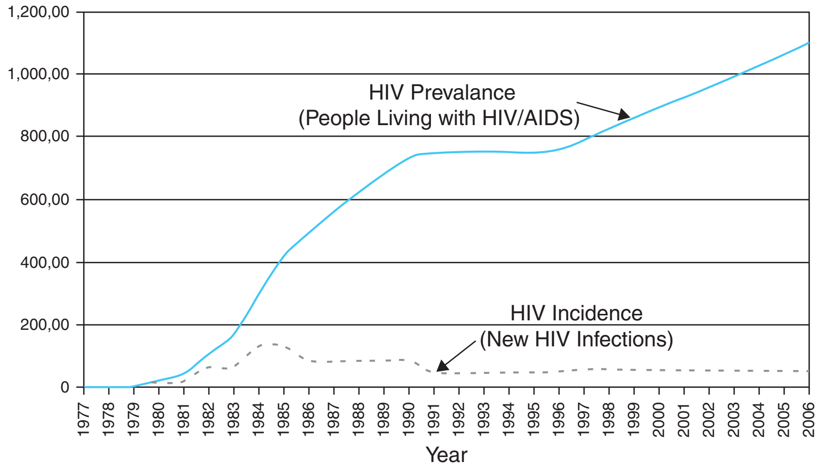


FIGURE 3–5 HIV incidence and prevalence, United States, 1977–2006. *Source:* Adapted and reprinted from HIV and AIDS in the United States. Available at: http://www.cdc.gov/hiv/topics/surveillance/print/united_states.htm. Accessed August 25, 2012.

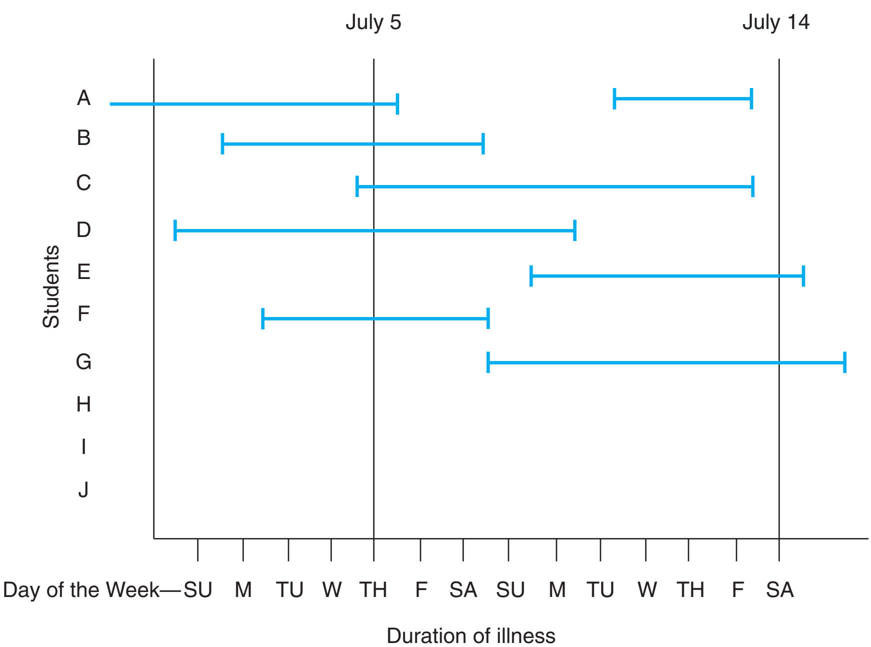


FIGURE 3–6 Outbreak of meningococcal infections in a summer school class of 10 students. *Note:* Students H, I, and J were not ill.

the incidence rate of disease was 3/10. Note that on July 5 cases A, B, C, D, and F were existing cases of disease and were not included in the count for incidence; subsequently, case A was a recurrent case and should be counted once for incidence and twice for period prevalence. The measure of incidence would be more accurate if the cumulative duration of observation (person-days) was used in the denominator. If one was interested only in the first occurrence of meningococcal disease, then students A, B, C, D, and F would not have been included in the estimation of incidence, because they were prevalent cases on July 5. In that situation, the incidence would have been 2/5.

Applications of Incidence Data

It was noted earlier that prevalence data are useful for determining the extent of a disease (particularly chronic diseases) or health problem in the community. Prevalence data are not as helpful as incidence data for studies of etiology because of the possible influence of differential survival. (The prevalent cases may be the survivors who remain after the other cases died; consequently, the prevalent cases may represent an incomplete picture of the outcome variable.) Incidence data (e.g., cumulative incidence rates) help in research on the etiology of disease because they provide estimates of risk of developing the disease. Thus, incidence rates are considered to be fundamental tools in research that pursues the causality of diseases. Note how the incidence rate of postmenopausal breast cancer was calculated in the IWHS. Comparison of incidence rates in population groups that differ in exposures permits one to estimate the effects of exposure to a hypothesized factor of interest. This study design, known as a cohort study, differs from a prevalence study in that it selects participants who have a specific kind of exposure (e.g., exposure to a toxic chemical).

Crude Rates

The basic concept of a rate can be broken down into three general categories: crude rates, specific rates, and adjusted rates. *Crude rates* are summary rates based on the actual number of events in a population over a given time period. An example is the crude death rate, which approximates the proportion of a population that dies during a time period of interest.[1] Refer to the study questions and exercises at the end of this chapter for calculation problems. Some of the more commonly used crude rates are presented in **Exhibit 3–5**. The definitions for *measures of natality* (statistics associated with births) come from Health, United States, 2010.[9]

EXHIBIT 3–5

Examples of Crude Rates: Overview of Measures That Pertain to Birth, Fertility, Infant Mortality, and Related Phenomena

- *Crude birth rate:* used to project population changes; it is affected by the number and age composition of women of childbearing age.
- *Fertility rate:* used for comparisons of fertility among age, racial, and socioeconomic groups.
- *Infant mortality rate:* used for international comparisons; a high rate indicates unmet health needs and poor environmental conditions.
- *Fetal death rate* (and late fetal death rate): used to estimate the risk of death of the fetus associated with the stages of gestation.
- *Fetal death ratio:* provides a measure of fetal wastage (loss) relative to the number of live births.
- *Neonatal mortality rate:* reflects events happening after birth, primarily:
  1. Congenital malformations
  2. Prematurity (birth before gestation week 28)
  3. Low birth weight (birth weight less than 2,500 g)
- *Postneonatal mortality rate:* reflects environmental events, control of infectious diseases, and improvement in nutrition. Since 1950, neonatal mortality in the United States has declined; postneonatal mortality has not declined greatly.
- *Perinatal mortality rate:* reflects events that occur during pregnancy and after birth; it combines mortality during the prenatal and postnatal periods.
- *Maternal mortality rate:* reflects healthcare access and socioeconomic factors; it includes maternal deaths resulting from causes associated with pregnancy and puerperium (during and after childbirth).

Birth Rate

The *crude birth rate* refers to the number of live births during a specified period of time (e.g., one calendar year) per the resident population during the midpoint of the time period (expressed as rate per 1,000). The crude birth rate is a useful measure of population growth and is an index for comparison of developed and developing countries. The crude birth rate is generally higher in less developed areas than in more developed areas of the world. As an illustration of this measure, **Figure 3–7** presents birth rates categorized by age of mother for the United States for the years 1990–2009. The birth rate has trended upward for older women and downward for women in the youngest age group.

$$\text{Crude birth rate} = \frac{\text{Number of live births within a given period}}{\text{Population size at the middle of that period}} \times 1{,}000 \text{ population}$$

Sample calculation: 4,130,665 babies were born in the United States during 2009, when the U.S. population was 307,006,550. The birth rate was 4,130,665/307,006,550 = 13.5 per 1,000.

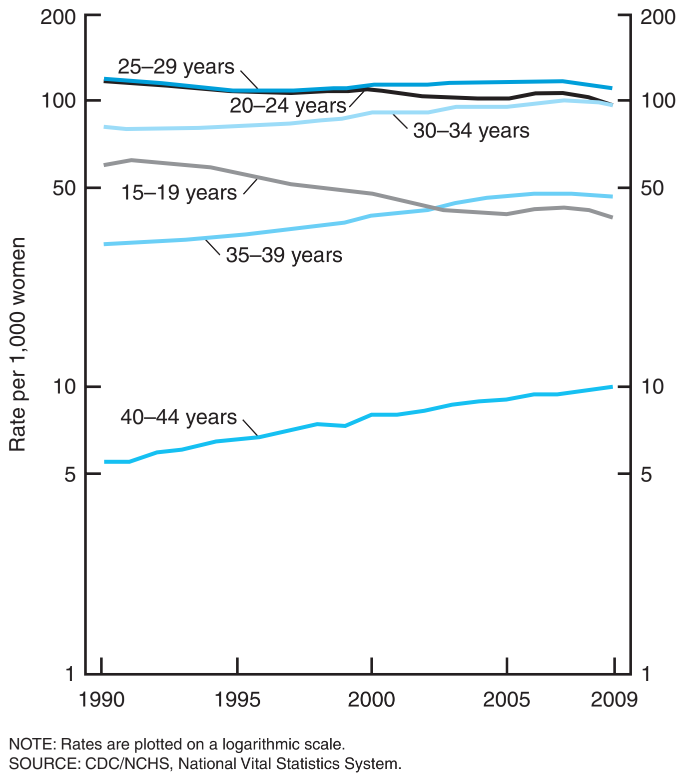


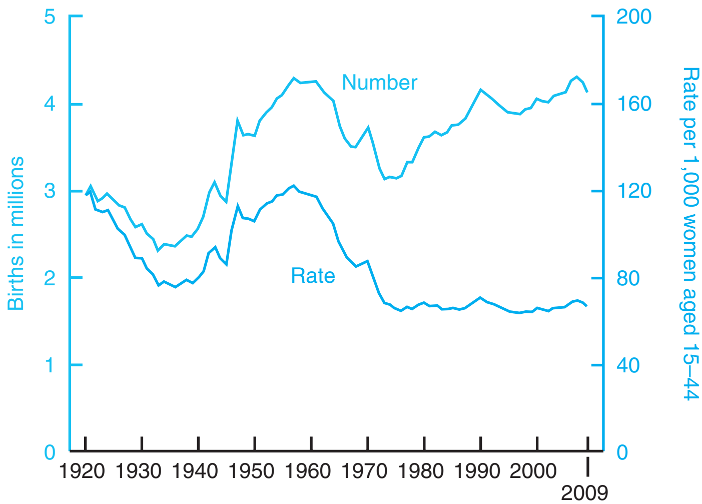
FIGURE 3–7 Birth rates by selected age of mother: United States, 1990–2009. *Source:* Reproduced from JA Martin, BE Hamilton, SJ Ventura, et al. Births: Final data for 2009, *National Vital Statistics Reports*, Vol 60, No 1, p. 6. Hyattsville, MD: National Center for Health Statistics; 2011.

Fertility Rate

Among the several types of fertility rates, one of the most noteworthy is the *general fertility rate*. This rate consists of the number of live births reported in an area during a given time interval (for example, during 1 year) divided by the number of women aged 15–44 years in that area. The population size for the number of women aged 15–44 years is assessed at the midpoint of the year. Sometimes the age range of 15–49 years is used. **Figure 3–8** illustrates fertility rates compared

$$\text{General fertility rate} = \frac{\text{Number of live births within a year}}{\text{Number of women age 15–44 years during the midpoint of the year}} \times \begin{array}{c}\text{1,000 women}\\ \text{aged 15–44}\end{array}$$

Sample calculation: During 2009, there were 61,948,144 women aged 15 to 44 in the United States. There were 4,130,665 live births. The general fertility rate was 4,130,665/61,948,144 = 66.7 per 1,000 women aged 15 to 44.



NOTE: Beginning with 1959, trend lines are based on registered live births; trend lines for 1920–1958 are based on live births adjusted for underregistration.
SOURCE: CDC/NCHS, National Vital Statistics System.

FIGURE 3–8 Live births and rates: United States, 1920–2009.
Source: Reprinted from JA Martin, BE Hamilton, SJ Ventura, et al. Births: Final data for 2009, *National Vital Statistics Reports*, Vol 60, No 1, p. 3. Hyattsville, MD: National Center for Health Statistics; 2011.

with the number of live births for the United States from 1920 to 2009. (The general fertility rate is often referred to more generically as the fertility rate.)

A second type of fertility rate is the *total fertility rate*. This rate is "[t]he average number of children that would be born if all women lived to the end of their childbearing years and bore children according to a given set of age-specific fertility rates."[1] In the United States, the total fertility rate was estimated to be 2.06 in 2012. This rate is close to the replacement fertility rate of 2.1, the rate at which the number of births is equivalent to the number of deaths; consequently, when its fertility rate is about 2.1, the United States does not have a net population gain due to births.

Fetal Mortality

Fetal mortality is an issue of major public health significance, although often overlooked. The term fetal mortality is defined as "spontaneous intrauterine death at any time during pregnancy."[10(p 1)] When such deaths occur during the later stages of pregnancy, they are sometimes referred to as stillbirths. Fetal mortality indices depend on estimation of fetal death after a certain number of weeks of gestation. In the following three definitions, the gestation time is stated or presumed. The *fetal death rate* is defined as the number of fetal deaths after 20 weeks or more gestation divided by the number of live births plus fetal deaths (after 20 weeks or more gestation). It is expressed as rate per 1,000 live births and fetal deaths. The late fetal death rate refers to fetal deaths after 28 weeks or more gestation. Both measures pertain to a calendar year.

The *fetal death ratio* refers to the number of fetal deaths after gestation of 20 weeks or more divided by the number of live births during a year. It is expressed as rate per 1,000 live births.

Fetal death rate (per 1,000 live births plus fetal deaths)

$$= \frac{\text{Number of fetal deaths after 20 weeks or more gestation}}{\begin{array}{c}\text{Number of live births + number of fetal deaths}\\ \text{after 20 weeks or more getation}\end{array}} \times 1{,}000$$

Late fetal death rate (per 1,000 live births plus late fetal deaths)

$$= \frac{\text{Number of fetal deaths after 28 weeks or more gestation}}{\begin{array}{c}\text{Number of live births + number of fetal deaths}\\ \text{after 28 weeks or more getation}\end{array}} \times 1{,}000$$

Fetal death ratio

$$= \frac{\text{Number of fetal deaths after 20 weeks or more gestation}}{\text{Number of live births}} \times \frac{1{,}000}{\text{(during a year)}}$$

Sample calculation: During 1 year there were 134 fetal deaths with 20 weeks or more gestation and 10,000 live births. The fetal death ratio is (134/10,000) = 13.4 per 1,000. Note that the fetal death rate is (134/10,134) = 13.2 per 1,000, which is slightly lower than the fetal death ratio.

See **Figure 3–9** for comparisons of the fetal mortality rate and late fetal mortality rate in the United States between 1990 and 2005. The overall fetal death rate (fetal mortality plus late fetal mortality) declined by 17% between 1990 and 2003. The decline was attributable to decreases in late fetal deaths.[10] The rate stabilized at 6.23 per 1,000 live births in 2003 and was nearly the same (6.22 per 1,000) in 2005. In comparison with other racial/ethnic groups, fetal death rates were highest for non-Hispanic black women, due to the greater risk of preterm delivery.

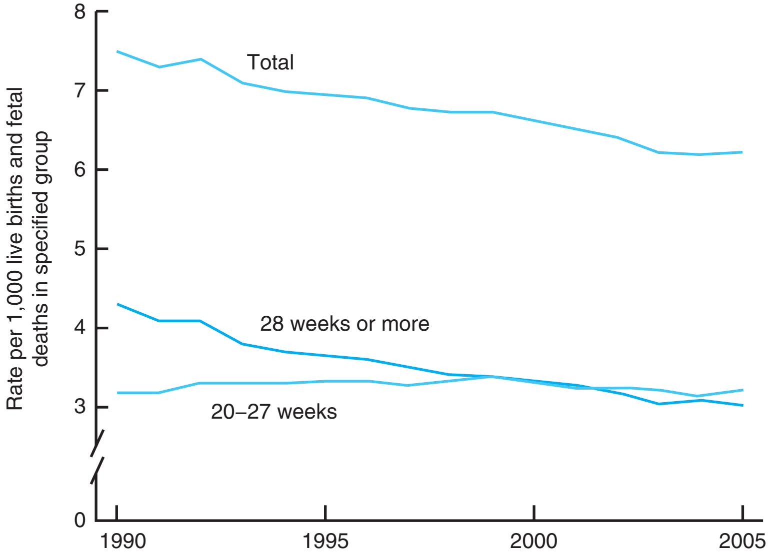


FIGURE 3–9 Fetal mortality rates, by period of gestation: United States, 1990–2005. *Source:* Reproduced from MF MacDorman, S Kirmeyer. The Challenge of Fetal Mortality. *NCHS Data Brief*, No 16, April 2009.

Infant Mortality Rate

The *infant mortality rate* is obtained by dividing the number of infant deaths during a calendar year by the number of live births reported in the same year. The infant mortality rate measures the risk of dying during the first year of life among infants born alive. Note that not all infants who die in a calendar year are born in that year, which represents a source of error. Typically, however, the number of infant deaths from previous years' births is balanced by an equal number of deaths during the following year among the current year's births. The following is the formula for the infant mortality rate:

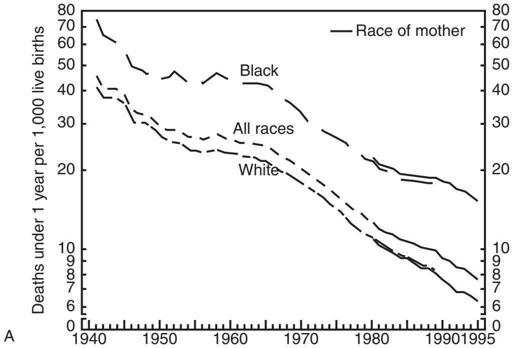
$$\text{Infant mortality} = \frac{\begin{array}{c}\text{Number of infant deaths}\\ \text{among infants aged 0–365 days}\\ \text{during the year}\end{array}}{\begin{array}{c}\text{Number of live births}\\ \text{during the year}\end{array}} \times 1{,}000 \text{ live births}$$

Sample calculation: In the United States during 2007, there were 29,153 deaths among infants under 1 year of age and 4,316,233 live births. The infant mortality rate was (29,153/4,316,233) × 1,000 = 6.75 per 1,000 live births.

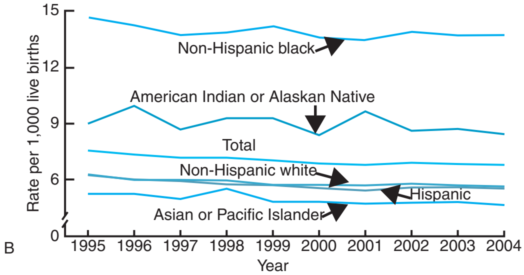
Infant mortality rates are highest among the least developed countries of the world (e.g., Afghanistan with 165 per 1,000 births in 2009) in comparison with some developing countries (e.g., India with 50 per 1,000), less developed countries of Eastern Europe (e.g., Romania with 10 per 1,000), and developed market economies (e.g., Sweden with about 2 per 1,000).[12]

Figure 3–10 shows trends in U.S. infant mortality by race from 1940 to 1995 (part A) and from 1995 to 2004 (part B). Note how total infant mortality rates declined steadily until 2000 and have declined very little since then. Infant mortality rates vary greatly by race/ethnicity in the United States The rate for non-Hispanic blacks is approximately twice the rate for the United States as a whole (part C).

Figure 3–11 presents a comparison of the infant mortality rate of the United States with that reported by other industrialized nations. In 2007, the U.S. infant mortality rate exceeded that of many other nations. Some of the differences observed between the United States and other developed/industrialized nations, may be artifactual (i.e., due to variations in the definition measurement and reporting of infant deaths). It is most likely, however, that the differences are associated with a high rate of preterm births in the United States.[10,11]



NOTES: Infant deaths are classified by race of decedent. For 1940–90, live births are classified by race of child and for 1980–94, by race of mother.



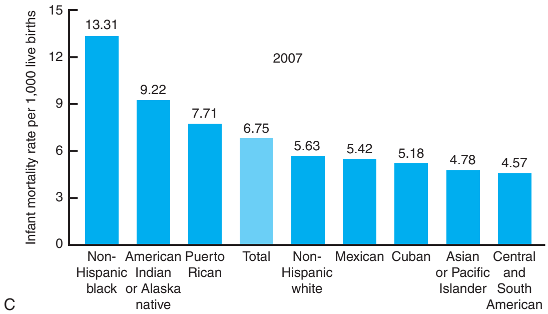


FIGURE 3–10 Infant mortality. Part A: Infant mortality rates by race: United States, 1950–1995. Part B: Infant mortality rates by race and ethnicity, 1995–2004. Part C: Infant mortality rates, by race and Hispanic origin of mother: United States, 2007. *Sources*: Part A: Modified from Anderson RN, Kochanek KD, and Murphy SL. Report of Final Monthly Statistics, 1995, Monthly Vital Statistics Report, Vol 45, No 11, Suppl 2, p. 12. Hyattsville, MD: National Center for Health Statistics; 1997. Part B: From Mathews TJ, MacDorman MF. Infant mortality statistics from the 2004 period linked birth/infant death data set, National Vital Statistics Reports; Vol 55, No 15, p. 1. Hyattsville, MD: National Center for Health Statistics; 2007. Part C: MF MacDorman, TJ Mathews. Understanding Racial and Ethnic Disparities in U.S. Infant Mortality Rates. NCHS Data Brief, No 74, September 2011.

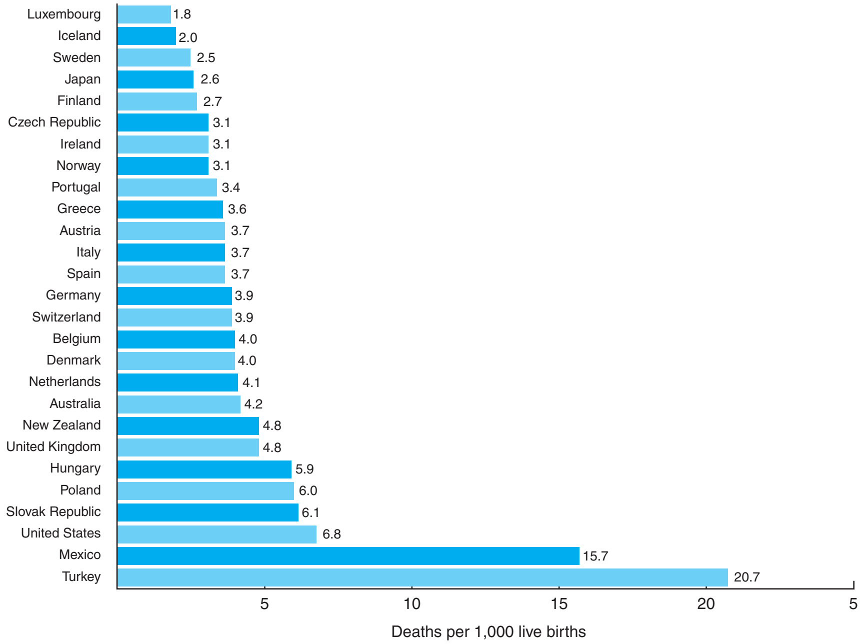


FIGURE 3–11 International infant mortality rates, selected countries,* 2007. *Source:* Reproduced from U. S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau. *Child Health USA 2011*, International Infant Mortality, p. 28.

*2007 data were not available for all Organization for Economic Co-operation and Development (OECD) countries.

Neonatal Mortality Rate

The *neonatal mortality rate* measures risk of dying among newborn infants who are under the age of 28 days (0–27 days) for a given year. The formula is as follows:

$$\text{Neonatal mortality rate} = \frac{\text{Number of infant deaths under 28 days of age}}{\text{Number of live births}} \times 1{,}000 \text{ live births (during a year)}$$

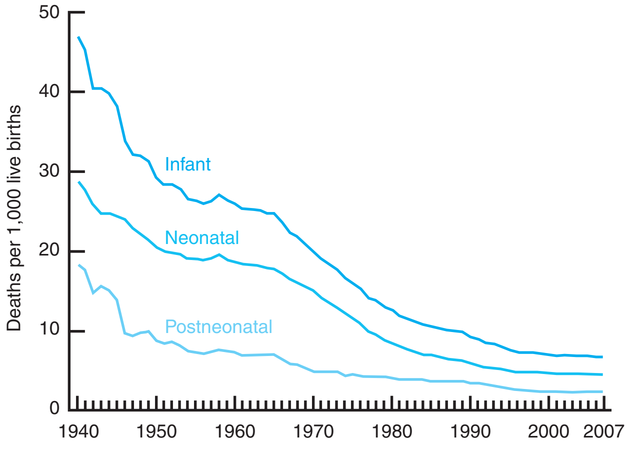
Postneonatal Mortality Rate

A statistic that is related to the neonatal mortality rate is the *postneonatal mortality rate*. The postneonatal mortality rate measures risk of dying among older infants during a given year.

Postneonatal mortality rate

$$= \frac{\text{Number of infant deaths from 28 days to 365 days after birth}}{\text{Number of live births} - \text{neonatal deaths}} \times \begin{matrix}\text{1,000 live} \\ \text{births}\end{matrix}$$

Figure 3–12 illustrates trends in infant mortality rates, neonatal mortality rates, and postneonatal mortality rates in the United States. All three measures showed a declining trend after 1940. However, the infant mortality rate was higher than either the neonatal mortality rate or the postneonatal mortality rate; in addition, both infant and neonatal mortality rates were higher than the postneonatal mortality rate. Between 1997 and 2007 the following mortality trends I occurred: infant mortality decreased by 7%; neonatal mortality decreased by 8%; and postneonatal mortality rate decreased by 5%.[13] Between 2006 and 2007 neonatal mortality did not change significantly. In 2007 the neonatal mortality rate was 4.42. Postneonatal mortality showed a statistically significant increase over 2006 of 3.5% (from 2.24 to 2.34 per 1,000 live births, all races combined).[14]



NOTE: Rates are infant (under 1 year), neonatal (under 28 days), and postneonatal (28 days—11 months) deaths per 1,000 live births in specified group.
SOURCE: CDC/NCHS, National Vital Statistics System, Mortality.

FIGURE 3–12 Infant, neonatal, and postneonatal mortality rates: United States, 1940–2007. *Source:* Reproduced from Xu J, Kochanek KD, Murphy SL, et al. Deaths: Final Data for 2007, *National Vital Statistics Reports*. Vol 58, No 19, p. 13. Hyattsville, MD: National Center for Health Statistics, 2010.

Perinatal Mortality

Two measures of perinatal mortality are the *perinatal mortality rate* and the *perinatal mortality ratio*. The perinatal period used in these measures captures late fetal deaths (stillbirths) plus infant deaths within 7 days of birth. Internationally, perinatal mortality reflects variations in the care of mothers as well as their health and nutritional statuses and also is a quality metric for obstetrics and pediatrics.[15] Approximately 3 million stillbirths and 3 million infant deaths (during the first 7 days of life) occur across the globe each year. According to the World Health Organization, "[t]he perinatal mortality rate is five times higher in developing than in developed regions: 10 deaths per 1,000 births in developed countries; 50 per 1,000 in developing regions and over 60 per 1,000 in least developed countries. It is highest in Africa, with 62 deaths per 1,000 births, and especially in middle and western Africa, which have rates as high as 75 and 76 per 1,000."[15(p 20)] **Figure 3–13** compares rates of perinatal mortality in world regions. The United States had a perinatal mortality rate of 7 per 1,000 in 2000. The formulas for the perinatal mortality rate and perinatal mortality ratio are:

Perinatal mortality rate

$$= \frac{\begin{array}{c}\text{Number of late fetal deaths after}\\ \text{28 weeks or more gestation +}\\ \text{infant deaths within 7 days of birth}\end{array}}{\begin{array}{c}\text{Number of live births +}\\ \text{number of late fetal deaths}\end{array}} \times 1{,}000 \text{ live births and fetal deaths}$$

Perinatal mortality ratio

$$= \frac{\begin{array}{c}\text{Number of late fetal deaths after}\\ \text{28 weeks or more gestation +}\\ \text{infant deaths within 7 days of birth}\end{array}}{\text{Number of live births}} \times 1{,}000 \text{ live births}$$

Maternal Mortality Rate

The *maternal mortality rate* is the number of maternal deaths ascribed to childbirth (i.e., pregnancy and puerperal causes) per 10,000 or 100,000 live births. Factors that affect maternal mortality include maternal age, socioeconomic status, nutritional status, and healthcare access. **Figure 3–14** gives causes of maternal mortality. Direct causes include complications related to the puerperium (period after childbirth), eclampsia (a condition marked by convulsions following delivery), and hemorrhage.

Maternal mortality rate
(per 100,000 live births, including multiple births)

$$= \frac{\text{Number of deaths assigned to causes related to childbirth}}{\text{Number of live births}} \times 100{,}000 \text{ live births (during a year)}$$

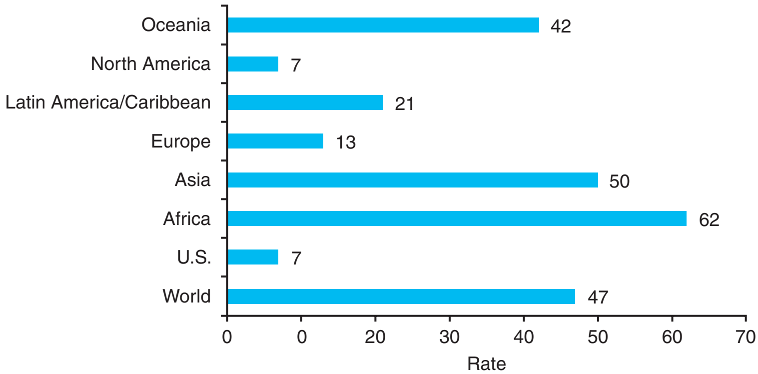


FIGURE 3–13 Global estimates of perinatal mortality rates by geographical (United Nations) region and subregion, 2000. *Source:* Data from World Health Organization, Neonatal and perinatal mortality: country, regional and global estimates. Geneva: Switzerland.

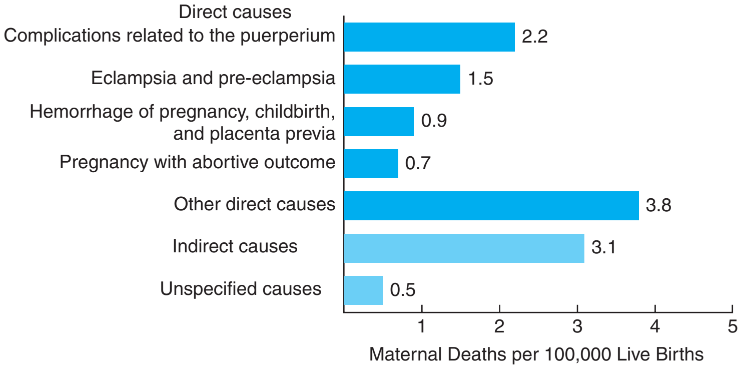


FIGURE 3–14 Leading causes of maternal mortality, 2007. *Source:* Reproduced from U. S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau. *Child Health USA 2011*, International Infant Mortality, p. 29.

Specific Rates and Proportional Mortality Ratio

Specific rates are a type of rate based on a particular subgroup of the population defined, for example, in terms of race, age, or sex, or they may refer to the entire population but be specific for some single cause of death or illness. Although the crude rates described so far are important and useful summary measures of the occurrence of disease, they are not without limitations. A crude rate should be used with caution in making comparative statements about disease frequencies in populations. Observed differences between populations in crude rates of disease may be the result of systematic factors within the populations rather than true variations in rates. Systematic differences in sex or age distributions would affect observed rates. To correct for factors that may influence the make-up of populations and in turn influence crude rates, one may construct specific and adjusted rates. Two other measures that have been defined previously also can be considered specific measures: incidence and prevalence. That is, both are typically specific to a particular end point. Examples of specific rates are cause-specific rates and age-specific rates.

Cause-Specific Rate

A *cause-specific rate* is "[a] rate that specifies events, such as deaths according to their cause."[1] An example of a cause-specific rate is the cause-specific mortality rate. As the name implies, it is the rate associated with a specific cause of death. Sample calculations are shown in **Table 3–3**. The number of deaths among the 25- to 34-year-old age group (population 39,872,598) due to human immunodeficiency virus (HIV) infection was 1,588 during 2003. The cause-specific mortality rate due to HIV was (1,588/39,872,598), or 4.0 per 100,000.

Cause-specific rate

$$= \frac{\text{Mortality (or frequency of a given disease)}}{\text{Population size at midpoint of time period}} \times 100{,}000$$

Age-Specific Rates

An *age-specific rate* is defined as "[a] rate for a specified age group. The numerator and denominator refer to the same age group."[1] To calculate age-specific rates, one subdivides (or stratifies) a population into age groups, such as those

Table 3–3 The 10 Leading Causes of Death, 25–34 Years, All Races, Both Sexes, United States, 2003 (Number in Population Aged 25–34 Years = 39,872,598)

Rank Order	Cause of Death	Number	Proportional Mortality Ratio (%)	Cause-Specific Death Rate per 100,000
1	Accidents (unintentional injuries)	12,541	30.4	31.5
2	Intentional self harm (suicide)	5,065	12.3	12.7
3	Assault (homicide)	4,516	10.9	11.3
4	Malignant neoplasms	3,741	9.1	9.4
5	Diseases of the heart	3,250	7.9	8.2
6	Human immunodeficiency virus (HIV) disease	1,588	3.8	4.0
7	Diabetes mellitus	657	1.6	1.6
8	Cerebrovascular diseases	583	1.4	1.5
9	Congenital malformations, deformations, and chromosomal abnormalities	426	1.0	1.1
10	Influenza and pneumonia	373	0.9	0.9
	All causes	41,300		

Source: Adapted from Heron MP, Smith BL. Deaths: Leading Causes for 2003. *National Vital Statistics Reports*, Vol 55, No 10, p. 18. Hyattsville, MD: National Center for Health Statistics; 2007.

defined by 5- or 10-year intervals. Then, one divides the frequency of a disease in a particular age stratum by the total number of persons within that age stratum to find the age-specific rate. A similar procedure may be employed to calculate sex-specific rates. An example of an age-specific cancer mortality rate is shown in **Exhibit 3–6**. A second example of the calculation of age-specific mortality rates for the U.S. population is shown in **Table 3–4**. (Some age-specific death rates shown in Table 3–4 differ from published rates because of differences in estimation of population size and use of different intervals for age groups.)

In summary, this section has demonstrated how to calculate cause-specific and age-specific rates. It is also possible to define other varieties of specific rates (e.g., sex specific rates). All in all, specific rates are a much better indicator of risk than crude rates, especially for rates specific to defined subsets of the population (e.g., age, race, and sex specific). A disadvantage of specific rates is the difficultly in visualizing the "big picture" in those situations where specific rates for several factors are presented in complex tables. **Table 3–5** shows the age-specific cancer incidence rates by sex, age group, and year of diagnosis. Most people would find it difficult to synthesize the data from a complex table, such as this one, and discern any specific trends. The numbers shown in the table are age-adjusted, a procedure that we will describe later in the chapter.

EXHIBIT 3–6

Age-Specific Rate (R_i)

Age-specific rate: The number of cases per age group of population (during a specified time period).

Example:

$$R_i = \frac{\text{Number of deaths among those aged 5–14 years}}{\text{Number of persons who are aged 5–14 years (during time period)}} \times 100{,}000$$

Sample calculation: In the United States during 2003, there were 1,651 deaths due to malignant neoplasms among the age group 5 to 14 years, and there were 40,968,637 persons in the same age group. The age-specific malignant neoplasm death rate in this age group is (1,651/40,968,637) = 4.0 per 100,000.

Table 3–4 Method of Calculation of Age-Specific Death Rates

Age Group (Years)	Number of Deaths (D_i) in 2003	Number in Population (P_i) as of July 1, 2003*	Age-Specific Rate (R_i) per 100,000
Under 1	28,025	4,003,606	700.0
1–4	4,965	15,765,673	31.5
5–14	6,954	40,968,637	17.0
15–24	33,568	41,206,163	81.5
25–34	41,300	39,872,598	103.6
35–44	89,461	44,370,594	201.6
45–54	176,781	40,804,599	433.2
55–64	262,519	27,899,736	940.9
65–74	413,497	18,337,044	2,255.0
75–84	703,024	12,868,672	5,463.1
85+	687,852	4,713,467	14,593.3
Not stated	342	NA	NA
Totals	2,448,288	290,810,789	841.9**

* Estimated

** The crude mortality rate for the United States

Source: Data from Hoyert DL, Heron MP, Murphy SL, Kung H. Deaths: Final Data for 2003. *National Vital Statistics Reports*, Vol 54, No 13, pp. 23 and 112. Hyattsville, MD: National Center for Health Statistics; 2006.

Proportional Mortality Ratio

The *proportional mortality ratio* (PMR) is the number of deaths within a population due to a specific disease or cause divided by the total number of deaths in the population.

Table 3–5 Trends in Age-Specific Cancer Incidence Rates* (By Sex, Age Group, and Year of Diagnosis, SEER Program, 2000–2009; All Sites Combined, All Races)

	Year of Diagnosis									
Sex/Age	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
All										
Ages < 20	16.0	17.1	17.0	15.9	16.8	17.2	16.2	17.2	17.2	17.3
Ages 20–49	146.2	147.3	145.9	146.0	148.5	149.9	149.3	150.6	151.9	151.2
Ages 50–64	876.9	884.1	881.6	847.1	844.7	835.4	846.1	859.4	845.9	834.7
Ages 65–74	2025.3	2041.9	2014.0	1944.3	1941.2	1903.2	1932.8	1970.1	1919.7	1886.2
Ages 75+	2463.2	2476.9	2457.7	2390.1	2404.0	2381.4	2367.3	2369.3	2318.8	2248.4
Males										
Ages < 20	16.8	18.0	18.5	16.5	17.8	18.1	17.1	18.2	18.5	18.1
Ages 20–49	113.0	113.5	112.1	110.8	112.3	113.1	111.3	112.9	113.4	110.7
Ages 50–64	962.2	971.7	973.8	934.7	933.9	911.2	940.5	957.3	929.5	910.7
Ages 65–74	2644.9	2675.9	2620.4	2514.3	2512.8	2427.1	2485.2	2538.1	2424.9	2356.7
Ages 75+	3349.4	3364.2	3333.3	3228.3	3227.3	3182.1	3162.8	3169.2	3041.9	2923.1
Females										
Ages < 20	15.2	16.0	15.5	15.1	15.7	16.3	15.2	16.2	15.9	16.4
Ages 20–49	179.1	181.0	179.6	181.0	184.8	186.9	187.5	188.6	190.9	192.2
Ages 50–64	798.9	804.0	797.0	766.6	762.9	765.9	759.3	769.2	768.9	764.8
Ages 65–74	1521.5	1522.6	1514.2	1472.3	1465.7	1465.0	1469.4	1492.0	1492.4	1486.6
Ages 75+	1939.3	1947.6	1933.4	1884.1	1904.9	1887.4	1872.2	1868.6	1862.7	1822.1

* Rates are per 100,000 and are age-adjusted to the 2000 US Standard Population (19 age groups—Census P25-1130).

Source: http://seer.cancer.gov/faststats/. Accessed May 22, 2012.

PMR (%)

$$= \frac{\text{Mortality due to a specific cause during a time period}}{\text{Mortality due to all causes during the same time period}} \times 100$$

Sample calculation: In a certain community, there were 66 deaths due to coronary heart disease during a year and 200 deaths due to all causes in that year. The PMR is $(66/200) \times 100 = 33\%$.

Refer to Table 3–3 for a more detailed example of a PMR. In Table 3–3, the PMR (%) is calculated according to the formula given above. For example, the proportional mortality ratio for HIV among the 25- to 34-year-old group was 3.8% (1,588/41,300). This PMR should be used with caution when comparisons are made across populations, especially those that have different rates of total mortality. To illustrate, consider that two countries have identical death rates from cardiovascular disease (perhaps 5 per 100,000 per year) and that each country has exactly 1 million inhabitants. Therefore, one would expect 50 deaths from cardiovascular disease to occur in each country (5 per 100,000 per year × 1,000,000). Suppose further, however, that in country A the total death rate per 100,000 per year is 30 and that it is only 10 in country B. Therefore, the expected total number of deaths would be 300 in country A and only 100 in country B. When these data are used to construct a PMR, one sees that the proportion of deaths from cardiovascular disease is higher in country B (0.50) than in country A (0.17). The PMR is not a measure of the risk of dying of a particular cause. It merely indicates, within a population, the relative importance of a specific cause of death. For a health administrator, such information may be useful to determine priorities and planning. To an epidemiologist, such differences may indicate an area for further study. For example, why does country A have such higher total mortality rates than country B? Is it merely because of differences in age structure? Is the difference a result of access to health care or certain behavioral or lifestyle patterns associated with elevated mortality? The PMR should not be confused with a case fatality rate, which expresses the proportion of fatal cases among all cases of disease during a specific time period.

Table 3–6 presents a summary of unadjusted measures of morbidity and mortality discussed in this chapter. We provide this table to assist you with future review and reference to these measures.

Table 3–6 Summary of Unadjusted Measures of Morbidity and Mortality Discussed in This Chapter

Measure	Numerator (x)	Denominator (y)	Expressed per Number at Risk (10^{n})*
Age-specific rate (e.g., mortality rate)	Number of deaths among a specific age group of the population during a specified time period	Number of persons who comprise that age group during the same time period	100,000
Attack rate	Number ill during a time period	Number ill + well during the same time period	100 (or percent)
Case fatality rate (discussed in Chapter 12)	Number of deaths due to disease "X" during a given time interval	Number of new cases of that disease reported during the same time interval	100 (or percent)
Cause-specific rate, (e.g., cause-specific death or mortality rate)	Number of deaths (mortality) assigned to a specific cause during a year	Estimated population size during the midpoint of that same year	100,000
Crude birth rate	Number of live births within a given period	Population size at the midpoint of that time period	1,000
Crude death rate	Number of deaths during a given year	Reference population during the midpoint of the year	100,000
Fetal death rate	Number of fetal deaths after 20 weeks or more gestation during a year	Number of live births + number of fetal deaths after 20 weeks or more gestation during the same year	1,000
Fetal death ratio	Number of fetal deaths after 20 weeks or more gestation during a year	Number of live births during the same year	1,000
General fertility rate	Number of live births within a year	Number of women aged 15–44 years during the midpoint of the year	1,000
Incidence density	Number of new cases during a time period	Total person-years of observation during the same time period	Varies (e.g., 100,000)
Incidence rate	Number of new cases during a time period	Total population at risk during the same time period	Varies (e.g., 100,000)
Infant mortality rate	Number of infant deaths under 1 year of age (0–365 days) during a year	Number of live births reported during that same year	1,000
Late fetal death rate	Number of fetal deaths after 28 weeks or more gestation during a year	Number of live births + number of fetal deaths after 28 weeks or more gestation during the same year	1,000
Maternal mortality rate	Number of deaths assigned to causes related to childbirth during a year	Number of live births during the same year	100,000
Neonatal mortality rate	Number of infant deaths under 28 days of age during a year	Number of live births during the same year	1,000
Perinatal mortality rate	Number of late fetal deaths after 28 weeks or more gestation + infant deaths within 7 days of birth during a year	Number of live births + number of late fetal deaths during the same year	1,000
Perinatal mortality ratio	Number of late fetal deaths after 28 weeks or more gestation + infant deaths within 7 days of birth during a year	Number of live births during the same year	1,000
Postneonatal mortality rate	Number of infant deaths from 28 days to 365 days after birth during a year	Number of live births – neonatal deaths during the same year	1,000
Proportional mortality ratio	Mortality due to a specific cause during a year	Mortality due to all causes during that same year	100 or 1,000

**Note:* The foregoing measures are expressed in the general form $\frac{x}{y} \times 10^{n}$.

Adjusted Rates

Adjusted rates are summary measures of the rate of morbidity or mortality in a population in which statistical procedures have been applied to remove the effect of differences in composition of the various populations. A common factor for rate adjustment is age, which is probably the most important variable in risk of morbidity and mortality, although rates can be adjusted for other variables. Crude rates mask differences between populations that differ in age and thus are not satisfactory for comparing health outcomes in such populations.[16] Members of older populations have a much greater risk of mortality than those in younger populations. Consequently, when a population is older the crude mortality rate will be higher than when the population is younger. Refer to **Table 3–7**.

The crude death rates (all ages) in Group A and Group B are 50 per 1,000 and 40 per 1,000, respectively; these rates suggest that Group A has a higher mortality rate than Group B. Next, we will examine the age-adjusted death rates for the same populations: These rates are 42 per 1,000 and 52 per 1,000. (For the time being, ignore the procedures for age adjustment; these will be described later.) Group A has a lower age-adjusted mortality rate than Group B because the population of Group A is older.

Figure 3–15 presents trends in U.S. crude and age-adjusted mortality rates between 1960 and 2007, a time interval during which the population has aged. Both rates have trended downward, with the age-adjusted rate declining much more steeply.

Now let's examine methods for adjusting rates: Two methods for the adjustment of rates are the direct method and the indirect method. An easy way to remember how they differ is that direct and indirect refer to the source of the rates. The direct method may be used if age-specific death rates in a population to be standardized are known and a suitable standard population is available. The direct method is presented in **Table 3–8**. Note that each age-specific rate found in Table 3–4 is multiplied by the number of persons in the age group in the standard population. Before the year 2000, the U.S. population in 1940 was used as the standard; now the standard shown in Table 3–8 is the estimated number in the standard population in the year 2000. (See **Exhibit 3–7** for information on the development of the year 2000 standard for age adjustment.) As indicated in the fourth column of Table 3–8, the result is the expected number of deaths in each age group, which is then summed across all age groups to determine the total number of expected deaths. The age-adjusted rate is the total expected number of deaths divided by the total estimated 2000 population times 100,000: $[(2{,}286{,}926.31/274{,}633{,}642) \times 100{,}000] = 832.7$ per 100,000.

Table 3–7 Group Comparison of Crude and Age-Adjusted Death Rates

	Group A				Group B					
Age	Deaths (1)	Population (2)	Rate[1] (3)	Weighted Rate[2] (4)	Deaths (5)	Population (6)	Rate[1] (7)	Weighted Rate[2] (8)	Standard Population (9)	Standard Weight[3] (10)
All ages	500	10,000	50	. . .	400	10,000	40	. . .	10,000	1.0
0–24 years	20	1,000	20	6	180	6,000	30	9	3,000	0.3
25–64 years	120	3,000	40	12	150	3,000	50	15	3,000	0.3
65 years and over	360	6,000	60	24	70	1,000	70	28	4,000	0.4
Age-adjusted death rate[1]	. . .	. . .	. . .	42	. . .	. . .	. . .	52	. . .	. . .

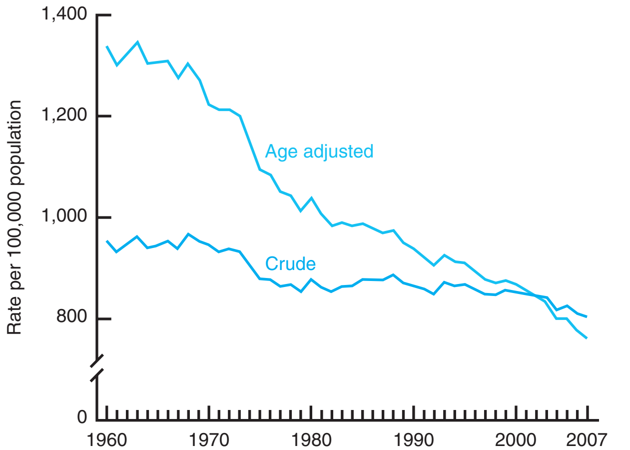
. . . Category not applicable.

[1] Rate per 1,000 population.

[2] The weighted rate is calculated by multiplying the age-specific rate by the standard weight.

[3] The standard weight for each age group is calculated by dividing the standard population at each age by the total standard population.

Source: From Anderson RN, Rosenberg HM. Age Standardization of Death Rates: Implementation of the Year 2000 Standard, *National Vital Statistics Reports*, Vol 47, No 3, p. 2. National Center for Health Statistics; 1998.



NOTE: Crude death rates are on an annual basis per 100,000 population; age-adjusted rates are per 100,000 U.S. standard population.
SOURCE: CDC/NCHS, National Vital Statistics System, Mortality.

FIGURE 3–15 Crude and age-adjusted death rates: United States, 1960–2007. *Source:* Reproduced from Xu J, Kochanek KD, Murphy SL, et al. Deaths: Final Data for 2007, *National Vital Statistics Reports*. Vol 58, No 19, p. 4. Hyattsville, MD: National Center for Health Statistics, 2010.

EXHIBIT 3–7

The National Center for Health Statistics Adopts a New Standard Population for Age Standardization of Death Rates

The crude death rate is a widely used measure of mortality. However, crude rates are influenced by the age composition of the population. As such, comparisons of crude death rates over time or between groups may be misleading if the populations being compared differ in age composition. This is relevant, for example, in trend comparisons of U.S. mortality, given the aging of the U.S. population. . . . The crude death rate for the United States rose from 852.2 per 100,000 population to 880.0 during 1979 to 1995. This increase in the crude death rate was due to the increasing proportion of the U.S. population in older age groups that have higher death rates. Age standardization, often called age adjustment, is one of the key tools used to control for the changing age distribution of the population, and thereby to make meaningful death comparisons of vital

continues

EXHIBIT 3–7 ***continued***

rates over time and between groups. In contrast to the rising crude death rate, the age-adjusted death rate for the United States dropped from 577.0 per 100,000 U.S. standard population to 503.9 during 1979 to 1995. This age-adjusted comparison is free from the confounding effect of changing age distribution and therefore better reflects the trend in U.S. mortality. To use age adjustment requires a standard population, which is a set of arbitrary population weights.

The new standard is based on the year 2000 population and beginning with data year 1999 will replace the existing standard based on the 1940 population. . . . Currently, at least three different standards are used among Department of Health and Human Services agencies. Implementation of the year 2000 standard will reduce confusion among data users and the burden on state and local agencies. Use of the year 2000 standard also will result in age-adjusted death rates that are substantially larger than those based on the 1940 standard. Further, the new standard will affect trends in age-adjusted rates for certain causes of death and will narrow race differentials in age-adjusted death rates. ■

Source: Adapted from Anderson RN, Rosenberg HM. Age Standardization of Death Rates: Implementation of the Year 2000 Standard, *National Vital Statistics Reports*, Vol 47, No 3, p. 1. National Center for Health Statistics; 1998.

To summarize, direct adjustment requires the application of the observed rates of disease in a population to some standard population to derive an expected number (rate) of mortality. The same procedure would be followed for other populations that one might wish to compare. By standardizing the observed rates of disease in the populations being compared to the same reference population, one is thereby assured that any observed differences that remain are not simply a reflection of differences in population structure with respect to factors such as age, race, and sex.

A method of direct adjustment that achieves the same results as those reported in Table 3–8 uses year 2000 standard weights (refer to **Table 3–9**). From the previous discussion, you may have inferred the following relationship:

$$\mathrm{R}_i = \frac{D_i}{P_i}$$

where R_i = the age-specific death rate for the *i*-th interval (row) in Table 3–4 and:

D_i = number of deaths in age interval i

P_i = number of persons in age interval i at midyear

Table 3–8 Direct Method for Adjustment of Death Rates

Age Group (Years)	2003 Age-Specific Death Rate per 100,000†	Number in Standard Population, 2000*	Expected Number of Deaths
Under 1	700.0	3,794,901	26,564.08
1–4	31.5	15,191,619	4,784.22
5–14	17.0	39,976,619	6,785.62
15–24	81.5	38,076,743	31,018.66
25–34	103.6	37,233,437	38,566.36
35–44	201.6	44,659,185	90,042.86
45–54	433.2	37,030,152	160,428.66
55–64	940.9	23,961,506	225,462.73
65–74	2,255.0	18,135,514	408,952.54
75–84	5,463.1	12,314,793	672,765.23
85+	14,593.3	4,259,173	621,555.36
Totals		274,633,642	2,286,926.31

† Age-specific death rates are from Table 3–4.

* Estimated

Age-adjusted rate per 100,000 = 832.7

(total expected number of deaths/estimated 2000 population) × 100,000

(2,286,926.31/274,633,642) × 100,000

Source: Data from Hoyert DL, Heron MP, Murphy SL, Kung H. Deaths: Final Data for 2003. *National Vital Statistics Reports*, Vol 54, No 13, p. 114. Hyattsville, MD: National Center for Health Statistics; 2006.

Table 3–9 Weighted Method for Direct Rate Adjustment

Age Group (Years)	Number in Standard Population, 2000* (P_{si})	Standard Weight (W_{si}) for 2000	Age-Specific Death Rate (R_i), 2003†	$W_{si} \cdot R_i$
Under 1	3,794,901	0.013818	700.0	9.6726
1–4	15,191,619	0.055316	31.5	1.7420
5–14	39,976,619	0.145563	17.0	2.4708
15–24	38,076,743	0.138646	81.5	11.2946
25–34	37,233,437	0.135575	103.6	14.0428
35–44	44,659,185	0.162614	201.6	32.7865
45–54	37,030,152	0.134835	433.2	58.4155
55–64	23,961,506	0.087249	940.9	82.0958
65–74	18,135,514	0.066035	2,255.0	148.9084
75–84	12,314,793	0.044841	5,463.1	244.9682
85+	4,259,173	0.015509	14,593.3	266.3216
Totals	$\sum_i P_{si} = 274{,}633{,}642$	1.0	N/A	$\sum_i W_{si} \cdot R_i = 832.7$

* Estimated

† From Table 3–4.

Source: Data from Hoyert DL, Heron MP, Murphy SL, Kung H. Deaths: Final Data for 2003. *National Vital Statistics Reports*, Vol 54, No 13, p. 114. Hyattsville, MD: National Center for Health Statistics; 2006.

We may assign standard weights (W_{si}) to each interval according to the following formula:

$$W_{si} = \frac{P_{si}}{\sum_i P_{si}}$$

where W_{si} is the standard weight associated with the i-th interval of the year 2000 standard U.S. population and:

P_{si} = the population in the i-th age interval in the standard population

$\sum_i P_{si}$ = total number in the standard population

Example (see Table 3–9): For age interval 1–4 years,

$$W = \frac{15{,}191{,}619}{274{,}633{,}642} = 0.055316$$

Then the age-adjusted death rate (AADR) is:

$$AADR = \sum_i W_{si} \bullet \frac{D_i}{P_i} = \sum_i W_{si} \bullet \mathrm{R}_i = 832.7$$

The formula for AADR indicates that the year 2000 standard weights for each age group are multiplied by the age-specific death rates in that same row. These products are then summed to obtain the AADR (see Table 3–9). Note that the results for this method of standardization are the same as those reported in Table 3–8.

A second method of age adjustment is the indirect method, which may be used if age-specific death rates of the population for standardization are unknown or unstable (e.g., because the rates to be standardized are based on a small population). The stratum-specific rates of a larger population, such as that of the United States, are applied to the number of persons within each stratum of the population of interest to obtain the expected numbers of deaths. Thus, the indirect method of standardization does not require knowledge of the actual age-specific incidence or mortality rates among each age group for the population to be standardized. By applying the rates of disease from a standard population (in this example, the 2003 population) to the observed structure of the population of interest, one is left with an expected number of cases (or deaths) in the study population if the rates of disease were the same as in the standard population. One way to evaluate the result is to construct a standardized morbidity ratio or a *standardized mortality ratio* (SMR).

$$\text{SMR} = \frac{\text{Observed deaths}}{\text{Expected deaths}} \times 100$$

Sample calculation: The number of observed deaths due to heart disease is 600 in a certain county during year 2014. The expected number of deaths is 1,000. The SMR = (600/1,000) × 100 = 60% (0.6).

If the observed and expected numbers are the same, the SMR would be 100% (1.0), indicating that the observed morbidity or mortality in the study population is not unusual. An SMR of 200% (2.0) is interpreted to mean that the death (or disease) rate in the study population is two times greater than expected.

A second example of the indirect method of adjustment is shown in **Table 3–10**. Note that the standard age-specific death rates for the year 2003 (which we will designate as the year for obtaining the standard population) from Table 3–4 were multiplied by the number in each age group of the population of interest to obtain the expected number of deaths. To calculate the SMR, the observed number of deaths was divided by the expected number. The crude mortality rate is 502/230,109 = 218.2 per 100,000. The SMR is (502/987.9) × 100 = 50.8%. From the SMR, one may conclude that the observed mortality in this population falls below expectations, because the SMR is less than 1.0 or 100%.

Note that construction of an SMR is not the only way to interpret the net effect of the indirect adjustment procedure. An alternative is to compute a mortality rate per 100,000 by using the expected number of deaths as the numerator,

Table 3–10 Illustration of Indirect Age Adjustment: Mortality Rate Calculation for a Fictitious Population of 230,109 Persons

Age (Years)	Number in Population of Interest	Death Rates (per 100,000) in Standard Population*	Expected Number of Deaths in Population of Interest
15–24	7,989	81.5	6.5
25–34	37,030	103.6	38.4
35–44	60,838	201.6	122.6
45–54	68,687	433.2	297.6
55–64	55,565	940.9	522.8
Totals	230,109		987.9

Total expected number of deaths = 987.9

Observed number of deaths in this population = 502

* Standard death rates are from Table 3–4.

rather than the observed number of deaths in the study population. If we wanted to focus on an outcome other than mortality, we could use the expected number of morbid events as the numerator. In either case, the calculation would be based on the expected numbers derived from the standard population. Referring to the example in Table 3–10, the total population size was 230,109 and the total expected number of deaths was 987.9. The adjusted death rate would be $987.9/230{,}109 \times 100{,}000 = 429.3$ per 100,000 per year. In comparison, the unadjusted death rate was 502/230,109 or 218.2 per 100,000 per year.

It is important to be aware that the numeric magnitude of an SMR in this situation is a reflection of the standard population. That is, if one were to use the age distribution of the 1970 U.S. population instead of the 2003 U.S. population for age adjustment, the adjusted rates that one would find would be quite different. Accordingly, SMRs for different populations typically cannot be compared with one another unless the same standard population has been applied to them. In addition, SMRs sometimes can be misleading: As a summary index, the overall SMR can be equal to 1.0 across different populations being compared, yet there might still be important differences in mortality in various subgroups. Finally, the longer a population is followed, the less information the SMR provides. Because it is expected that everyone in the population will die eventually, the SMR will tend to be equal to 1.0 over time.

Conclusion

This chapter defined several measures of disease frequency that are commonly employed in epidemiology. Counts or frequency data refer to the number of cases of a disease or other health phenomenon being studied. A ratio consists of a numerator and a denominator that express one number relative to another (e.g., the sex ratio). Prevalence is a measure of the existing number of cases of disease in a population at a point in time or over a specified period of time. A rate is defined as a proportion in which the numerator consists of the frequency of a disease during a period of time and the denominator is a unit size of population. Rates improve one's ability to make comparisons of health indices across contrasting populations. Examples of rates include the crude mortality rate, incidence rates, and infant mortality rates. Other examples of rates discussed were the birth rate, fertility rate, and perinatal mortality rate. Specific rates are more precise indicators of risk than crude rates. It was noted that, to make comparisons across populations, adjusted rates also may be used. Two techniques were presented on how to adjust rates. Finally, the chapter gave illustrations of how the SMR (an example of indirect adjustment) is used.

Study Questions and Exercises

1. Define the following terms:
   a. crude death rate
   b. age-specific rate
   c. cause-specific rate
   d. proportional mortality ratio (PMR)
   e. maternal mortality rate
   f. infant mortality rate
   g. neonatal mortality rate
   h. fetal death rate and late fetal death rate
   i. fetal death ratio
   j. perinatal mortality rate
   k. postneonatal mortality rate
   l. crude birth rate
   m. general fertility rate
   n. age-adjusted (standardized) rate
   o. direct method of adjustment
   p. indirect method of adjustment
   q. standardized mortality ratio (SMR)
2. Using **Table 3A–1**, calculate age-specific death rates for the category of malignant neoplasms of trachea, bronchus, and lung. What inferences can be made from the age-specific death rates for malignant neoplasms of trachea, bronchus, and lung?

Table 3A–1 Malignant Neoplasms of Trachea, Bronchus, and Lung Deaths by Age Group, United States, 2003

Age (Years)	Population	Malignant Neoplasms of Trachea, Bronchus, and Lung* Deaths
25–34	39,872,598	154
35–44	44,370,594	2,478
45–54	40,804,599	12,374
55–64	27,899,736	30,956
65–74	18,337,044	49,386

* Includes ICD-10, 1992 codes C33–C34.

Sources: Data are from Hoyert DL, Heron MP, Murphy SL, Kung H. Deaths: Final Data for 2003, *National Vital Statistics Reports*, Vol 54, No 13, p. 30. Hyattsville, MD: National Center for Health Statistics; 2006; and from Heron MP, Smith BL. Deaths: Leading Causes for 2003, *National Vital Statistics Reports*, Vol 55, No 10, p. 92. Hyattsville, MD: National Center for Health Statistics; 2007.

3. Using **Table 3A–2**, calculate the following for the United States: the age-specific death rates and age- and sex-specific death rates per 100,000 (for age groups 20–24, 25–34, and 35–44 years). Note that there are nine calculations and answers. For example, the age- and sex-specific death rate for females aged 15–19 years is [(3,889/9,959,789) × 100,000].

Table 3A–2 Mortality by Selected Age Groups, Males and Females, United States, 2003

	Males		Females		Total	
Age (Years)	Population	Number of Deaths	Population	Number of Deaths	Population	Number of Deaths
15–19	10,518,680	9,706	9,959,789	3,889	20,478,469	13,595
20–24	10,663,922	14,964	10,063,772	5,009	20,727,694	19,973
25–34	20,222,486	28,602	19,650,112	12,698	39,872,598	41,300
35–44	22,133,659	56,435	22,236,935	33,026	44,370,594	89,461
45–54	20,043,656	110,682	20,760,943	66,099	40,804,599	176,781

Sources: Data are from Heron MP, Smith BL. Deaths: Leading Causes for 2003, *National Vital Statistics Reports*, Vol 55, No 10, p. 92. Hyattsville, MD: National Center for Health Statistics; 2007; and from Hoyert DL, Heron MP, Murphy SL, Kung H. Deaths: Final Data for 2003, *National Vital Statistics Reports*, Vol 54, No 13, p. 21. Hyattsville, MD: National Center for Health Statistics; 2006.

4. Refer to both Table 3A-2 and **Table 3A–3**. The total population in 2003 was 290,810,789 (males = 143,037,290; females = 147,773,499). For 2003, the total number of live births was 4,089,950.

Table 3A-3 Total Mortality from Selected Causes, Males and Females, United States, 2003

Cause of Death	Males	Females	Total
All Causes	1,201,964	1,246,324	2,448,288
Accidents	70,532	38,745	109,277
Malignant Neoplasms	287,990	268,912	556,902
Alzheimer's Disease	18,335	45,122	63,457
Infant Deaths	15,902	12,123	28,025
Maternal Deaths	NA	495	495

Sources: Data are from Heron MP, Smith BL. Deaths: Leading Causes for 2003, *National Vital Statistics Reports*, Vol 55, No 10, p. 7–8. Hyattsville, MD: National Center for Health Statistics; 2007; and Hoyert DL, Heron MP, Murphy SL, Kung H. Deaths: Final Data for 2003, *National Vital Statistics Reports*, Vol 54, No 13, p. 101–102. Hyattsville, MD: National Center for Health Statistics; 2006.

a. Calculate the crude death rates (per 100,000) and the cause-specific death rates (per 100,000) for accidents, malignant neoplasms, and Alzheimer's disease. Repeat these calculations for males and females separately.
b. What are the PMRs (percent) for accidents, malignant neoplasms, and Alzheimer's disease? Repeat these calculations for males and females separately.
c. Calculate the maternal mortality rate (per 100,000 live births).
d. Calculate the infant mortality rate (per 1,000 live births).
e. Calculate the crude birth rate (per 1,000 population).
f. Calculate the general fertility rate (per 1,000 women aged 15–44 years).

5. The population of Metroville was 3,187,463 on June 30, 2013. During the period January 1 through December 31, 2013, a total of 4,367 city residents were infected with HIV. During the same year, 768 new cases of HIV were reported. Calculate the prevalence per 100,000 population and incidence per 100,000 population.
6. Give definitions of the terms prevalence and incidence. What are appropriate uses of prevalence and incidence data? State the relationships among prevalence, incidence, and duration of a disease.
7. Suppose that "X" represents the name of a disease. An epidemiologist conducts a survey of disease "X" in a population. The prevalence of disease "X" among women is 40/1,000 and among men is 20/1,000. Assuming that the data have been age adjusted, is it correct to conclude that women have twice the risk of disease "X" as men? Explain.
8. The following data regarding alcohol drinking status among persons in the United States were reported for 2005:

	Number in thousands	
	All persons 18 years of age and older	**Current regular alcoholic beverage drinkers**
Male	104,919	59,300
Female	112,855	44,373

a. What is the sex ratio of male to female regular alcoholic beverage drinkers?
b. What proportion (percent) of regular alcoholic beverage drinkers are women?
c. What is the prevalence per 1,000 of regular alcoholic beverage drinking among men only, women only, and the total population aged 18 and older?

9. During 2005, the following statistics were reported regarding the frequency of diabetes, ulcers, kidney disease, and liver disease:

Diabetes	7% of adults had ever been told by their doctor that they had diabetes
Ulcers	7% had ever been told by their doctor that they had an ulcer
Kidney	2% had been told in the past 12 months that they had kidney disease
Liver	1% had been told in the past 12 months that they had liver disease

Which of the foregoing statistics were stated as incidence data and which as prevalence data?

a. Diabetes
b. Ulcers
c. Kidney disease
d. Liver disease

10. The National Health Interview Survey reported the percent of respondents with a hearing problem by age group during 2005:

Age (years)	Reporting a hearing problem, %
18–44	8.2
45–64	19.2
65–74	30.4
75+	48.1

Would it be correct to state that the risk of hearing loss increases with age? Be sure to explain and defend your answer.

11. During January 1 through December 31, 2008, epidemiologists conducted a prevalence survey of type 2 diabetes; 500,000 cases were detected in a population of 10,000,000 persons. It was known that the incidence of diabetes in this population was 10 per 1,000. Estimate the percentage of the prevalent cases that were newly identified during the year.

12. The sex ratio for the entire United States was less than 100, indicating that there were more females than males. The sex ratio at birth exceeded 1.0, denoting a greater number of male births to female births. How could one account for the difference between the sex ratio for the United States and sex ratio at birth?

References

1. Porta M, ed. *A Dictionary of Epidemiology*, 5th ed. New York, NY: Oxford University Press; 2008.
2. Hennekens CH, Buring JE. *Epidemiology in Medicine*. Boston, MA: Little, Brown; 1987.
3. Folsom AR, Kaye SA, Sellers TA, et al. Body fat distribution and 5-year risk of death in older women. *JAMA*. 1993;269:483–487.

4. Sellers TA, Kushi LH, Potter JD, et al. Effect of family history, body-fat distribution, and reproductive factors on the risk of postmenopausal breast cancer. *N Engl J Med*. 1992;326:1323–1329.
5. Link BG, Susser E, Stueve A, Phelan J, Moore RE, Struening E. Lifetime and five-year prevalence of homelessness in the United States. *Am J Public Health*. 1994;84:1907–1912.
6. Mausner JS, Kramer S. Mausner & Bahn *Epidemiology–An Introductory Text*, 2nd ed. Philadelphia, PA: Saunders; 1985.
7. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. Belmont, CA: Lifetime Learning; 1982.
8. Rothman KJ. *Modern Epidemiology*. Boston, MA: Little, Brown; 1986.
9. National Center for Health Statistics. *Health, United States, 2010. With Special Feature on Death and Dying*. Hyattsville, MD: National Center for Health Statistics; 2011.
10. MacDorman MF, Kirmeyer S. The challenge of fetal mortality. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. NCHS Data Brief No 16 April 2009.
11. U.S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau. *Child Health USA 2011*. Rockville, Maryland: U.S. Department of Health and Human Services; 2011.
12. World Health Organization (WHO). *World Health Statistics 2011*. Geneva, Switzerland: World Health Organization; 2011.
13. National Center for Health Statistics. *Health, United States, 2010. With Special Feature on Death and Dying*. Hyattsville, MD: National Center for Health Statistics; 2011.
14. Xu J, Kochanek KD, Murphy SL, Tejada-Vera B. Final data for 2008. *National Vital Statistics Reports*. 58(19). Hyattsville, MD: National Center for Health Statistics; 2010.
15. World Health Organization. Neonatal and perinatal mortality: Country, regional and global estimates. Geneva, Switzerland: World Health Organization; 2006.
16. Anderson RN, Rosenberg HM. Age standardization of death rates: Implementation of the Year 2000 Standard. *National Vital Statistics Reports*. 47(3). Hyattsville, MD: National Center for Health Statistics; 1998.