

3

MEASURES OF RISK

Lesson 2 described measures of central location and spread, which are useful for summarizing continuous variables. However, many variables used by field epidemiologists are categorical variables, some of which have only two categories — exposed yes/no, test positive/negative, case/control, and so on. These variables have to be summarized with frequency measures such as ratios, proportions, and rates. Incidence, prevalence, and mortality

rates are three frequency measures that are used to characterize the occurrence of health events in a population.

Objectives

After studying this lesson and answering the questions in the exercises, you will be able to:

- *Calculate and interpret the following epidemiologic measures:*
 - *Ratio*
 - *Proportion*
 - *Incidence proportion (attack rate)*
 - *Incidence rate*
 - *Prevalence*
 - *Mortality rate*
- *Choose and apply the appropriate measures of association and measures of public health impact*

Major Sections

Frequency Measures	3-2
Morbidity Frequency Measures	3-10
Mortality Frequency Measures	3-20
Natality (Birth) Measures	3-38
Measures of Association	3-38
Measures of Public Health Impact.....	3-47
Summary	3-50

Frequency Measures

Numerator = upper portion of a fraction

Denominator = lower portion of a fraction

A measure of central location provides a single value that summarizes an entire distribution of data. In contrast, a frequency measure characterizes only part of the distribution. Frequency measures compare one part of the distribution to another part of the distribution, or to the entire distribution. Common frequency measures are **ratios**, **proportions**, and **rates**. All three frequency measures have the same basic form:

$$\frac{\text{numerator}}{\text{denominator}} \times 10^n$$

Recall that:

$10^0 = 1$ (anything raised to the 0 power equals 1)

$10^1 = 10$ (anything raised to the 1st power is the value itself)

$10^2 = 10 \times 10 = 100$

$10^3 = 10 \times 10 \times 10 = 1,000$

So the fraction of (numerator/denominator) can be multiplied by 1, 10, 100, 1000, and so on. This multiplier varies by measure and will be addressed in each section.

Ratio

Definition of ratio

A ratio is the relative magnitude of two quantities or a comparison of any two values. It is calculated by dividing one interval- or ratio-scale variable by the other. The numerator and denominator need not be related. Therefore, one could compare apples with oranges or apples with number of physician visits.

Method for calculating a ratio

$$\frac{\text{Number or rate of events, items, persons, etc. in one group}}{\text{Number or rate of events, items, persons, etc. in another group}}$$

After the numerator is divided by the denominator, the result is often expressed as the result “to one” or written as the result “:1.”

Note that in certain ratios, the numerator and denominator are different categories of the same variable, such as males and females, or persons 20–29 years and 30–39 years of age. In other ratios, the numerator and denominator are completely different

variables, such as the number of hospitals in a city and the size of the population living in that city.

EXAMPLE: Calculating a Ratio — Different Categories of Same Variable

Between 1971 and 1975, as part of the National Health and Nutrition Examination Survey (NHANES), 7,381 persons ages 40–77 years were enrolled in a follow-up study.¹ At the time of enrollment, each study participant was classified as having or not having diabetes. During 1982–1984, enrollees were documented either to have died or were still alive. The results are summarized as follows.

	Original Enrollment (1971–1975)	Dead at Follow-Up (1982–1984)
Diabetic men	189	100
Nondiabetic men	3,151	811
Diabetic women	218	72
Nondiabetic women	3,823	511

Of the men enrolled in the NHANES follow-up study, 3,151 were nondiabetic and 189 were diabetic. Calculate the ratio of non-diabetic to diabetic men.

$$\text{Ratio} = 3,151 / 189 \times 1 = 16.7:1$$

Properties and uses of ratios

- Ratios are common descriptive measures, used in all fields. In epidemiology, ratios are used as both descriptive measures and as analytic tools. As a descriptive measure, ratios can describe the male-to-female ratio of participants in a study, or the ratio of controls to cases (e.g., two controls per case). As an analytic tool, ratios can be calculated for occurrence of illness, injury, or death between two groups. These ratio measures, including risk ratio (relative risk), rate ratio, and odds ratio, are described later in this lesson.
- As noted previously, the numerators and denominators of a ratio can be related or unrelated. In other words, you are free to use a ratio to compare the number of males in a population with the number of females, or to compare the number of residents in a population with the number of hospitals or dollars spent on over-the-counter medicines.
- Usually, the values of both the numerator and denominator of a ratio are divided by the value of one or the other so that either the numerator or the denominator equals 1.0. So the ratio of non-diabetics to diabetics cited in the previous example is more likely to be reported as 16.7:1 than 3,151:189.

EXAMPLES: Calculating Ratios for Different Variables

Example A: A city of 4,000,000 persons has 500 clinics. Calculate the ratio of clinics per person.

$$500 / 4,000,000 \times 10^n = 0.000125 \text{ clinics per person}$$

To get a more easily understood result, you could set $10^n = 10^4 = 10,000$. Then the ratio becomes:

$$0.000125 \times 10,000 = 1.25 \text{ clinics per } 10,000 \text{ persons}$$

You could also divide each value by 1.25, and express this ratio as 1 clinic for every 8,000 persons.

Example B: Delaware's infant mortality rate in 2001 was 10.7 per 1,000 live births.² New Hampshire's infant mortality rate in 2001 was 3.8 per 1,000 live births. Calculate the ratio of the infant mortality rate in Delaware to that in New Hampshire.

$$10.7 / 3.8 \times 1 = 2.8:1$$

Thus, Delaware's infant mortality rate was 2.8 times as high as New Hampshire's infant mortality rate in 2001.

A commonly used epidemiologic ratio: death-to-case ratio

Death-to-case ratio is the number of deaths attributed to a particular disease during a specified period divided by the number of new cases of that disease identified during the same period. It is used as a measure of the severity of illness: the death-to-case ratio for rabies is close to 1 (that is, almost everyone who develops rabies dies from it), whereas the death-to-case ratio for the common cold is close to 0.

For example, in the United States in 2002, a total of 15,075 new cases of tuberculosis were reported.³ During the same year, 802 deaths were attributed to tuberculosis. The tuberculosis death-to-case ratio for 2002 can be calculated as $802 / 15,075$. Dividing both numerator and denominator by the numerator yields 1 death per 18.8 new cases. Dividing both numerator and denominator by the denominator (and multiplying by $10^n = 100$) yields 5.3 deaths per 100 new cases. Both expressions are correct.

Note that, presumably, many of those who died had initially contracted tuberculosis years earlier. Thus many of the 802 in the numerator are not among the 15,075 in the denominator. Therefore, the death-to-case ratio is a ratio, but not a proportion.

Proportion

Definition of proportion

A proportion is the comparison of a part to the whole. It is a type

of ratio in which the numerator is included in the denominator. You might use a proportion to describe what fraction of clinic patients tested positive for HIV, or what percentage of the population is younger than 25 years of age. A proportion may be expressed as a decimal, a fraction, or a percentage.

Method for calculating a proportion

$$\frac{\text{Number of persons or events with a particular characteristic}}{\text{Total number of persons or events, of which the numerator is a subset}} \times 10^n$$

For a proportion, 10^n is usually 100 (or $n=2$) and is often expressed as a percentage.

EXAMPLE: Calculating a Proportion

Example A: *Calculate the proportion of men in the NHANES follow-up study who were diabetics.*

Numerator = 189 diabetic men

Denominator = Total number of men = 189 + 3,151 = 3,340

$$\text{Proportion} = (189 / 3,340) \times 100 = 5.66\%$$

Example B: *Calculate the proportion of deaths among men.*

Numerator = deaths in men

= 100 deaths in diabetic men + 811 deaths in nondiabetic men

= 911 deaths in men

Notice that the numerator (911 deaths in men) is a subset of the denominator.

Denominator = all deaths

= 911 deaths in men + 72 deaths in diabetic women + 511 deaths in nondiabetic women

= 1,494 deaths

$$\text{Proportion} = 911 / 1,494 = 60.98\% = 61\%$$

Your Turn: *What proportion of all study participants were men? (Answer = 45.25%)*

Properties and uses of proportions

- Proportions are common descriptive measures used in all fields. In epidemiology, proportions are used most often as descriptive measures. For example, one could calculate the proportion of persons enrolled in a study among all those eligible (“participation rate”), the proportion of children in a village vaccinated against measles, or the proportion of persons who developed illness among all passengers of a cruise ship.
- Proportions are also used to describe the amount of disease that can be attributed to a particular exposure. For example, on the

basis of studies of smoking and lung cancer, public health officials have estimated that greater than 90% of the lung cancer cases that occur are attributable to cigarette smoking.

- In a proportion, the numerator must be included in the denominator. Thus, the number of apples divided by the number of oranges is not a proportion, but the number of apples divided by the total number of fruits of all kinds is a proportion. Remember, the numerator is always a subset of the denominator.
- A proportion can be expressed as a fraction, a decimal, or a percentage. The statements “one fifth of the residents became ill” and “twenty percent of the residents became ill” are equivalent.
- Proportions can easily be converted to ratios. If the numerator is the number of women (179) who attended a clinic and the denominator is all the clinic attendees (341), the proportion of clinic attendees who are women is $179 / 341$, or 52% (a little more than half). To convert to a ratio, subtract the numerator from the denominator to get the number of clinic patients who are not women, i.e., the number of men ($341 - 179 = 162$ men.) Thus, ratio of women to men could be calculated from the proportion as:

$$\begin{aligned}\text{Ratio} &= 179 / (341 - 179) \times 1 \\ &= 179 / 162 \\ &= 1.1 \text{ to } 1 \text{ female-to-male ratio}\end{aligned}$$

Conversely, if a ratio's numerator and denominator together make up a whole population, the ratio can be converted to a proportion. You would add the ratio's numerator and denominator to form the denominator of the proportion, as illustrated in the NHANES follow-up study examples (provided earlier in this lesson).

A specific type of epidemiologic proportion: proportionate mortality

Proportionate mortality is the proportion of deaths in a specified population during a period of time that are attributable to different causes. Each cause is expressed as a percentage of all deaths, and the sum of the causes adds up to 100%. These proportions are not rates because the denominator is all deaths, not the size of the population in which the deaths occurred. Table 3.1 lists the primary causes of death in the United States in 2003 for persons of all ages and for persons aged 25–44 years, by number of deaths, proportionate mortality, and rank.

Table 3.1 Number, Proportionate Mortality, and Ranking of Deaths for Leading Causes of Death, All Ages and 25–44 Year Age Group, United States, 2003

	All Ages			Ages 25–44 Years		
	Number	Percentage	Rank	Number	Percentage	Rank
All causes	2,443,930	100.0		128,924	100.0	
Diseases of heart	684,462	28.0	1	16,283	12.6	3
Malignant neoplasms	554,643	22.7	2	19,041	14.8	2
Cerebrovascular disease	157,803	6.5	3	3,004	2.3	8
Chronic lower respiratory diseases	126,128	5.2	4	401	0.3	*
Accidents (unintentional injuries)	105,695	4.3	5	27,844	21.6	1
Diabetes mellitus	73,965	3.0	6	2,662	2.1	9
Influenza & pneumonia	64,847	2.6	7	1,337	1.0	10
Alzheimer's disease	63,343	2.6	8	0	0.0	*
Nephritis, nephrotic syndrome, nephrosis	33,615	1.4	9	305	0.2	*
Septicemia	34,243	1.4	10	328	0.2	*
Intentional self-harm (suicide)	30,642	1.3	11	11,251	8.7	4
Chronic liver disease and cirrhosis	27,201	1.1	12	3,288	2.6	7
Assault (homicide)	17,096	0.7	13	7,367	5.7	5
HIV disease	13,544	0.5	*	6,879	5.3	6
All other	456,703	18.7		29,480	22.9	

* Not among top ranked causes

Data Sources: Centers for Disease Control and Prevention. *Summary of notifiable diseases, United States, 2003. MMWR 2005;2(No. 54).*

Hoyert DL, Kung HC, Smith BL. Deaths: Preliminary data for 2003. National Vital Statistics Reports; vol. 53 no 15. Hyattsville, MD: National Center for Health Statistics 2005: p. 15, 27.

As illustrated in Table 3.1, the proportionate mortality for HIV was 0.5% among all age groups, and 5.3% among those aged 25–44 years. In other words, HIV infection accounted for 0.5% of all deaths, and 5.3% of deaths among 25–44 year olds.

Rate

Definition of rate

In epidemiology, a rate is a measure of the frequency with which an event occurs in a defined population over a specified period of time. Because rates put disease frequency in the perspective of the size of the population, rates are particularly useful for comparing disease frequency in different locations, at different times, or among different groups of persons with potentially different sized populations; that is, a rate is a measure of risk.

To a non-epidemiologist, rate means how fast something is happening or going. The speedometer of a car indicates the car's speed or rate of travel in miles or kilometers per hour. This rate is always reported per some unit of time. Some epidemiologists restrict use of the term rate to similar measures that are expressed per unit of time. For these epidemiologists, a rate describes how quickly disease occurs in a population, for example, 70 new cases of breast cancer per 1,000 women per year. This measure conveys

a sense of the speed with which disease occurs in a population, and seems to imply that this pattern has occurred and will continue to occur for the foreseeable future. This rate is an **incidence rate**, described in the next section, starting on page 3-13.

Other epidemiologists use the term rate more loosely, referring to proportions with case counts in the numerator and size of population in the denominator as rates. Thus, an **attack rate** is the proportion of the population that develops illness during an outbreak. For example, 20 of 130 persons developed diarrhea after attending a picnic. (An alternative and more accurate phrase for attack rate is **incidence proportion**.) A **prevalence rate** is the proportion of the population that has a health condition at a point in time. For example, 70 influenza case-patients in March 2005 reported in County A. A **case-fatality rate** is the proportion of persons with the disease who die from it. For example, one death due to meningitis among County A’s population. All of these measures are proportions, and none is expressed per units of time. Therefore, these measures are not considered “true” rates by some, although use of the terminology is widespread.

Table 3.2 summarizes some of the common epidemiologic measures as ratios, proportions, or rates.

Table 3.2 Epidemiologic Measures Categorized as Ratio, Proportion, or Rate

Condition	Ratio	Proportion	Rate
Morbidity (Disease)	Risk ratio (Relative risk)	Attack rate (Incidence proportion)	Person-time incidence rate
	Rate ratio	Secondary attack rate	
	Odds ratio	Point prevalence	
	Period prevalence	Attributable proportion	
Mortality (Death)	Death-to-case ratio	Proportionate mortality	Crude mortality rate
			Case-fatality rate
			Cause-specific mortality rate
			Age-specific mortality rate
			Maternal mortality rate
			Infant mortality rate
Natality (Birth)			Crude birth rate
			Crude fertility rate



Exercise 3.1

For each of the fractions shown below, indicate whether it is a ratio, a proportion, a rate, or none of the three.

- A. *Ratio*
- B. *Proportion*
- C. *Rate*
- D. *None of the above*

- _____ 1. $\frac{\text{number of women in State A who died from heart disease in 2004}}{\text{number of women in State A who died in 2004}}$
- _____ 2. $\frac{\text{number of women in State A who died from heart disease in 2004}}{\text{estimated number of women living in State A on July 1, 2004}}$
- _____ 3. $\frac{\text{number of women in State A who died from heart disease in 2004}}{\text{number of women in State A who died from cancer in 2004}}$
- _____ 4. $\frac{\text{number of women in State A who died from lung cancer in 2004}}{\text{number of women in State A who died from cancer (all types) in 2004}}$
- _____ 5. $\frac{\text{number of women in State A who died from lung cancer in 2004}}{\text{estimated revenue (in dollars) in State A from cigarette sales in 2004}}$



Check your answers on page 3-51

Morbidity Frequency Measures

Morbidity has been defined as any departure, subjective or objective, from a state of physiological or psychological well-being. In practice, morbidity encompasses disease, injury, and disability. In addition, although for this lesson the term refers to the number of persons who are ill, it can also be used to describe the periods of illness that these persons experienced, or the duration of these illnesses.⁴

Measures of morbidity frequency characterize the number of persons in a population who become ill (incidence) or are ill at a given time (prevalence). Commonly used measures are listed in Table 3.3.

Table 3.3 Frequently Used Measures of Morbidity

Measure	Numerator	Denominator
Incidence proportion (or attack rate or risk)	Number of new cases of disease during specified time interval	Population at start of time interval
Secondary attack rate	Number of new cases among contacts	Total number of contacts
Incidence rate (or person-time rate)	Number of new cases of disease during specified time interval	Summed person-years of observation or average population during time interval
Point prevalence	Number of current cases (new and preexisting) at a specified point in time	Population at the same specified point in time
Period prevalence	Number of current cases (new and preexisting) over a specified period of time	Average or mid-interval population

Incidence refers to the occurrence of new cases of disease or injury in a population over a specified period of time. Although some epidemiologists use incidence to mean the number of new cases in a community, others use incidence to mean the number of new cases per unit of population.

Two types of incidence are commonly used — **incidence proportion** and **incidence rate**.

Incidence proportion or risk

Definition of incidence proportion

Synonyms for incidence proportion

- Attack rate
 - Risk
 - Probability of developing disease
 - Cumulative incidence
-

Incidence proportion is the proportion of an initially disease-free population that develops disease, becomes injured, or dies during a specified (usually limited) period of time. Synonyms include attack rate, risk, probability of getting disease, and cumulative incidence. Incidence proportion is a proportion because the persons in the numerator, those who develop disease, are all included in the denominator (the entire population).

Method for calculating incidence proportion (risk)

$$\frac{\text{Number of new cases of disease or injury during specified period}}{\text{Size of population at start of period}}$$

EXAMPLES: Calculating Incidence Proportion (Risk)

Example A: *In the study of diabetics, 100 of the 189 diabetic men died during the 13-year follow-up period. Calculate the risk of death for these men.*

Numerator = 100 deaths among the diabetic men
Denominator = 189 diabetic men
 $10^0 = 10^2 = 100$

$$\text{Risk} = (100 / 189) \times 100 = 52.9\%$$

Example B: *In an outbreak of gastroenteritis among attendees of a corporate picnic, 99 persons ate potato salad, 30 of whom developed gastroenteritis. Calculate the risk of illness among persons who ate potato salad.*

Numerator = 30 persons who ate potato salad and developed gastroenteritis
Denominator = 99 persons who ate potato salad
 $10^0 = 10^2 = 100$

$$\text{Risk} = \text{"Food-specific attack rate"} = (30 / 99) \times 100 = 0.303 \times 100 = 30.3\%$$

Properties and uses of incidence proportions

- Incidence proportion is a measure of the risk of disease or the probability of developing the disease during the specified period. As a measure of incidence, it includes only new cases of disease in the numerator. The denominator is the number of persons in the population at the start of the observation period. Because all of the persons with new cases of disease (numerator) are also represented in the denominator, a risk is also a proportion.

More About Denominators

The denominator of an incidence proportion is the number of persons at the start of the observation period. The denominator should be limited to the “population at risk” for developing disease, i.e., persons who have the potential to get the disease and be included in the numerator. For example, if the numerator represents new cases of cancer of the ovaries, the denominator should be restricted to women, because men do not have ovaries. This is easily accomplished because census data by sex are readily available. In fact, ideally the denominator should be restricted to women with ovaries, excluding women who have had their ovaries removed surgically (often done in conjunction with a hysterectomy), but this is not usually practical. This is an example of field epidemiologists doing the best they can with the data they have.

- In the outbreak setting, the term **attack rate** is often used as a synonym for risk. It is the risk of getting the disease during a specified period, such as the duration of an outbreak. A variety of attack rates can be calculated.

Overall attack rate is the total number of new cases divided by the total population.

A **food-specific attack rate** is the number of persons who ate a specified food and became ill divided by the total number of persons who ate that food, as illustrated in the previous potato salad example.

A **secondary attack rate** is sometimes calculated to document the difference between community transmission of illness versus transmission of illness in a household, barracks, or other closed population. It is calculated as:

$$\frac{\frac{\text{Number of cases among contacts of primary cases}}{\text{Total number of contacts}}}{\text{Total number of contacts}} \times 10^n$$

Often, the total number of contacts in the denominator is calculated as the total population in the households of the primary cases, minus the number of primary cases. For a secondary attack rate, 10^n usually is 100%.

EXAMPLE: Calculating Secondary Attack Rates

Consider an outbreak of shigellosis in which 18 persons in 18 different households all became ill. If the population of the community was 1,000, then the overall attack rate was $18 / 1,000 \times 100\% = 1.8\%$. One incubation period later, 17 persons in the same households as these “primary” cases developed shigellosis. If the 18 households included 86 persons, calculate the secondary attack rate.

$$\text{Secondary attack rate} = (17 / (86 - 18)) \times 100\% = (17 / 68) \times 100\% = 25.0\%$$

Incidence rate or person-time rate

Definition of incidence rate

Incidence rate or person-time rate is a measure of incidence that incorporates time directly into the denominator. A person-time rate is generally calculated from a long-term cohort follow-up study, wherein enrollees are followed over time and the occurrence of new cases of disease is documented. Typically, each person is observed from an established starting time until one of four “end points” is reached: onset of disease, death, migration out of the study (“lost to follow-up”), or the end of the study. Similar to the incidence proportion, the numerator of the incidence rate is the number of new cases identified during the period of observation. However, the denominator differs. The denominator is the sum of the time each person was observed, totaled for all persons. This denominator represents the total time the population was at risk of and being watched for disease. Thus, the incidence rate is the ratio of the number of cases to the total time the population is at risk of disease.

Method for calculating incidence rate

$$\frac{\text{Number of new cases of disease or injury during specified period}}{\text{Time each person was observed, totaled for all persons}}$$

In a long-term follow-up study of morbidity, each study participant may be followed or observed for several years. One person followed for 5 years without developing disease is said to contribute 5 person-years of follow-up.

What about a person followed for one year before being lost to follow-up at year 2? Many researchers assume that persons lost to follow-up were, on average, disease-free for half the year, and thus contribute $\frac{1}{2}$ year to the denominator. Therefore, the person followed for one year before being lost to follow-up contributes 1.5 person-years. The same assumption is made for participants diagnosed with the disease at the year 2 examination — some may have developed illness in month 1, and others in months 2 through 12. So, on average, they developed illness halfway through the year. As a result, persons diagnosed with the disease contribute $\frac{1}{2}$ year of follow-up during the year of diagnosis.

The denominator of the person-time rate is the sum of all of the person-years for each study participant. So, someone lost to

follow-up in year 3, and someone diagnosed with the disease in year 3, each contributes 2.5 years of disease-free follow-up to the denominator.

Properties and uses of incidence rates

- An incidence rate describes how quickly disease occurs in a population. It is based on person-time, so it has some advantages over an incidence proportion. Because person-time is calculated for each subject, it can accommodate persons coming into and leaving the study. As noted in the previous example, the denominator accounts for study participants who are lost to follow-up or who die during the study period. In addition, it allows enrollees to enter the study at different times. In the NHANES follow-up study, some participants were enrolled in 1971, others in 1972, 1973, 1974, and 1975.
- Person-time has one important drawback. Person-time assumes that the probability of disease during the study period is constant, so that 10 persons followed for one year equals one person followed for 10 years. Because the risk of many chronic diseases increases with age, this assumption is often not valid.
- Long-term cohort studies of the type described here are not very common. However, epidemiologists far more commonly calculate incidence rates based on a numerator of cases observed or reported, and a denominator based on the mid-year population. This type of incident rate turns out to be comparable to a person-time rate.
- Finally, if you report the incidence rate of, say, the heart disease study as 2.5 per 1,000 person-years, epidemiologists might understand, but most others will not. Person-time is epidemiologic jargon. To convert this jargon to something understandable, simply replace “person-years” with “persons per year.” Reporting the results as 2.5 new cases of heart disease per 1,000 persons per year sounds like English rather than jargon. It also conveys the sense of the incidence rate as a dynamic process, the speed at which new cases of disease occur in the population.

EXAMPLES: Calculating Incidence Rates

Example A: Investigators enrolled 2,100 women in a study and followed them annually for four years to determine the incidence rate of heart disease. After one year, none had a new diagnosis of heart disease, but 100 had been lost to follow-up. After two years, one had a new diagnosis of heart disease, and another 99 had been lost to follow-up. After three years, another seven had new diagnoses of heart disease, and 793 had been lost to follow-up. After four years, another 8 had new diagnoses with heart disease, and 392 more had been lost to follow-up.

The study results could also be described as follows: No heart disease was diagnosed at the first year. Heart disease was diagnosed in one woman at the second year, in seven women at the third year, and in eight women at the fourth year of follow-up. One hundred women were lost to follow-up by the first year, another 99 were lost to follow-up after two years, another 793 were lost to follow-up after three years, and another 392 women were lost to follow-up after 4 years, leaving 700 women who were followed for four years and remained disease free.

Calculate the incidence rate of heart disease among this cohort. Assume that persons with new diagnoses of heart disease and those lost to follow-up were disease-free for half the year, and thus contribute $\frac{1}{2}$ year to the denominator.

Numerator	= number of new cases of heart disease = $0 + 1 + 7 + 8 = 16$
Denominator	= person-years of observation = $(2,000 + \frac{1}{2} \times 100) + (1,900 + \frac{1}{2} \times 1 + \frac{1}{2} \times 99) + (1,100 + \frac{1}{2} \times 7 + \frac{1}{2} \times 793) + (700 + \frac{1}{2} \times 8 + \frac{1}{2} \times 392)$ = 6,400 person-years of follow-up
	or
Denominator	= person-years of observation = $(1 \times 1.5) + (7 \times 2.5) + (8 \times 3.5) + (100 \times 0.5) + (99 \times 1.5) + (793 \times 2.5) + (392 \times 3.5) + (700 \times 4)$ = 6,400 person-years of follow-up
Person-time rate	= $\frac{\text{Number of new cases of disease or injury during specified period}}{\text{Time each person was observed, totaled for all persons}}$ = $16 / 6,400$ = .0025 cases per person-year = 2.5 cases per 1,000 person-years

In contrast, the incidence proportion can be calculated as $16 / 2,100 = 7.6$ cases per 1,000 population during the four-year period, or an average of 1.9 cases per 1,000 per year (7.6 divided by 4 years). The incidence proportion underestimates the true rate because it ignores persons lost to follow-up, and assumes that they remained disease-free for all four years.

Example B: The diabetes follow-up study included 218 diabetic women and 3,823 nondiabetic women. By the end of the study, 72 of the diabetic women and 511 of the nondiabetic women had died. The diabetic women were observed for a total of 1,862 person-years; the nondiabetic women were observed for a total of 36,653 person-years. Calculate the incidence rates of death for the diabetic and non-diabetic women.

For diabetic women, numerator = 72 and denominator = 1,862

Person-time rate	= $72 / 1,862$ = 0.0386 deaths per person-year = 38.6 deaths per 1,000 person-years
------------------	---

For nondiabetic women, numerator = 511 and denominator = 36,653

Person-time rate	= $511 / 36,653 = 0.0139$ deaths per person-year = 13.9 deaths per 1,000 person-years
------------------	--

Example C: In 2003, 44,232 new cases of acquired immunodeficiency syndrome (AIDS) were reported in the United States.⁵ The estimated mid-year population of the U.S. in 2003 was approximately 290,809,777.⁶ Calculate the incidence rate of AIDS in 2003.

$$\begin{aligned} \text{Numerator} &= 44,232 \text{ new cases of AIDS} \\ \text{Denominator} &= 290,809,777 \text{ estimated mid-year population} \\ 10^n &= 100,000 \\ \text{Incidence rate} &= (44,232 / 290,809,777) \times 100,000 \\ &= 15.21 \text{ new cases of AIDS per } 100,000 \text{ population} \end{aligned}$$

Prevalence

Definition of prevalence

Prevalence, sometimes referred to as **prevalence rate**, is the proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time. Prevalence differs from incidence in that prevalence includes all cases, both new and preexisting, in the population at the specified time, whereas incidence is limited to new cases only.

Point prevalence refers to the prevalence measured at a particular point in time. It is the proportion of persons with a particular disease or attribute on a particular date.

Period prevalence refers to prevalence measured over an interval of time. It is the proportion of persons with a particular disease or attribute at any time during the interval.

Method for calculating prevalence of disease

$$\frac{\text{All new and pre-existing cases during a given time period}}{\text{Population during the same time period}} \times 10^n$$

Method for calculating prevalence of an attribute

$$\frac{\text{Persons having a particular attribute during a given time period}}{\text{Population during the same time period}} \times 10^n$$

The value of 10^n is usually 1 or 100 for common attributes. The value of 10^n might be 1,000, 100,000, or even 1,000,000 for rare attributes and for most diseases.

EXAMPLE: Calculating Prevalence

In a survey of 1,150 women who gave birth in Maine in 2000, a total of 468 reported taking a multivitamin at least 4 times a week during the month before becoming pregnant.⁷ Calculate the prevalence of frequent multivitamin use in this group.

Numerator = 468 multivitamin users
Denominator = 1,150 women

$$\text{Prevalence} = (468 / 1,150) \times 100 = 0.407 \times 100 = 40.7\%$$

Properties and uses of prevalence

- Prevalence and incidence are frequently confused. Prevalence refers to proportion of persons who *have* a condition at or during a particular time period, whereas incidence refers to the proportion or rate of persons who *develop* a condition during a particular time period. So prevalence and incidence are similar, but prevalence includes new and pre-existing cases whereas incidence includes new cases only. The key difference is in their numerators.

Numerator of incidence = new cases that occurred during a given time period

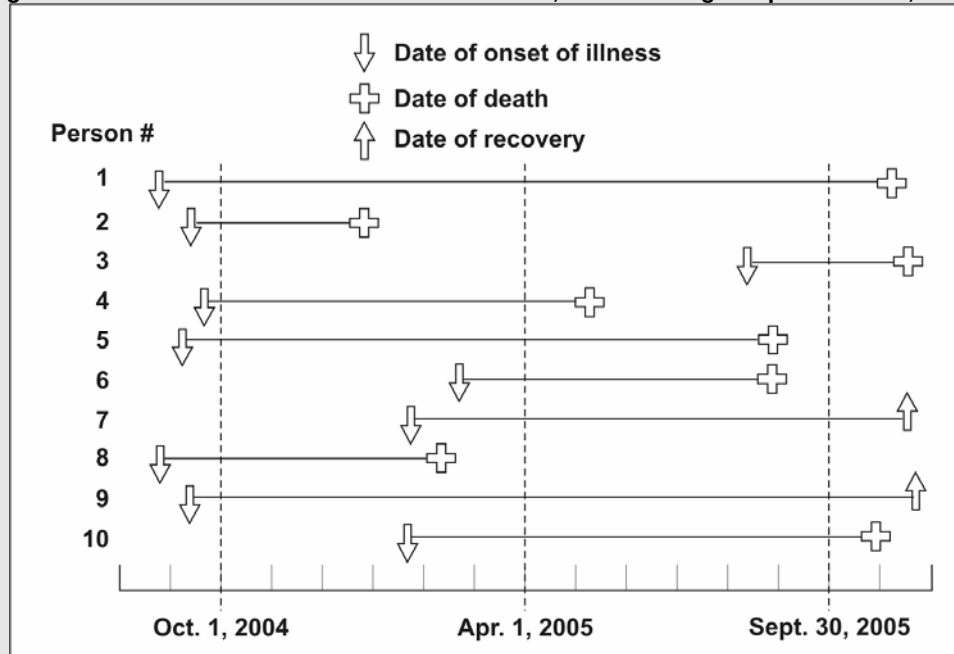
Numerator of prevalence = all cases present during a given time period

- The numerator of an incidence proportion or rate consists only of persons whose illness began during the specified interval. The numerator for prevalence includes all persons ill from a specified cause during the specified interval **regardless of when the illness began**. It includes not only new cases, but also preexisting cases representing persons who remained ill during some portion of the specified interval.
- Prevalence is based on both incidence and duration of illness. High prevalence of a disease within a population might reflect high incidence or prolonged survival without cure or both. Conversely, low prevalence might indicate low incidence, a rapidly fatal process, or rapid recovery.
- Prevalence rather than incidence is often measured for chronic diseases such as diabetes or osteoarthritis which have long duration and dates of onset that are difficult to pinpoint.

EXAMPLES: Incidence versus Prevalence

Figure 3.1 represents 10 new cases of illness over about 15 months in a population of 20 persons. Each horizontal line represents one person. The down arrow indicates the date of onset of illness. The solid line represents the duration of illness. The up arrow and the cross represent the date of recovery and date of death, respectively.

Figure 3.1 New Cases of Illness from October 1, 2004 through September 30, 2005



Example A: Calculate the incidence rate from October 1, 2004, to September 30, 2005, using the midpoint population (population alive on April 1, 2005) as the denominator. Express the rate per 100 population.

Incidence rate numerator	=	number of new cases between October 1 and September 30
	=	4 (the other 6 all had onsets before October 1, and are not included)
Incidence rate denominator	=	April 1 population
	=	18 (persons 2 and 8 died before April 1)
Incidence rate	=	$(4 / 18) \times 100$
	=	22 new cases per 100 population

Example B: Calculate the point prevalence on April 1, 2005. Point prevalence is the number of persons ill on the date divided by the population on that date. On April 1, seven persons (persons 1, 4, 5, 7, 9, and 10) were ill.

Point prevalence	=	$(7 / 18) \times 100$
	=	38.89%

Example C: Calculate the period prevalence from October 1, 2004, to September 30, 2005. The numerator of period prevalence includes anyone who was ill any time during the period. In Figure 3.1, the first 10 persons were all ill at some time during the period.

Period prevalence	=	$(10 / 20) \times 100$
	=	50.0%



Exercise 3.2

For each of the fractions shown below, indicate whether it is an incidence proportion, incidence rate, prevalence, or none of the three.

- A. Incidence proportion
- B. Incidence rate
- C. Prevalence
- D. None of the above

- _____ 1.
$$\frac{\text{number of women in Framingham Study who have died through last year from heart disease}}{\text{number of women initially enrolled in Framingham Study}}$$
- _____ 2.
$$\frac{\text{number of women in Framingham Study who have died through last year from heart disease}}{\text{number of person-years contributed through last year by women initially enrolled in Framingham Study}}$$
- _____ 3.
$$\frac{\text{number of women in town of Framingham who reported having heart disease in recent health survey}}{\text{estimated number of women residents of Framingham during same period}}$$
- _____ 4.
$$\frac{\text{number of women in Framingham Study newly diagnosed with heart disease last year}}{\text{number of women in Framingham Study without heart disease at beginning of same year}}$$
- _____ 5.
$$\frac{\text{number of women in State A newly diagnosed with heart disease in 2004}}{\text{estimated number of women living in State A on July 1, 2004}}$$
- _____ 6.
$$\frac{\text{estimated number of women smokers in State A according to 2004 Behavioral Risk Factor Survey}}{\text{estimated number of women living in State A on July 1, 2004}}$$
- _____ 7.
$$\frac{\text{number of women in State A who reported having heart disease in 2004 health survey}}{\text{estimated number of women smokers in State A according to 2004 Behavioral Risk Factor Survey}}$$



Check your answers on page 3-51

Mortality Frequency Measures

Mortality rate

A mortality rate is a measure of the frequency of occurrence of death in a defined population during a specified interval. Morbidity and mortality measures are often the same mathematically; it's just a matter of what you choose to measure, illness or death. The formula for the mortality of a defined population, over a specified period of time, is:

$$\frac{\text{Deaths occurring during a given time period}}{\text{Size of the population among which the deaths occurred}} \times 10^n$$

When mortality rates are based on vital statistics (e.g., counts of death certificates), the denominator most commonly used is the size of the population at the middle of the time period. In the United States, values of 1,000 and 100,000 are both used for 10^n for most types of mortality rates. Table 3.4 summarizes the formulas of frequently used mortality measures.

Table 3.4 Frequently Used Measures of Mortality

Measure	Numerator	Denominator	10^n
Crude death rate	Total number of deaths during a given time interval	Mid-interval population	1,000 or 100,000
Cause-specific death rate	Number of deaths assigned to a specific cause during a given time interval	Mid-interval population	100,000
Proportionate mortality	Number of deaths assigned to a specific cause during a given time interval	Total number of deaths from all causes during the same time interval	100 or 1,000
Death-to-case ratio	Number of deaths assigned to a specific cause during a given time interval	Number of new cases of same disease reported during the same time interval	100
Neonatal mortality rate	Number of deaths among children < 28 days of age during a given time interval	Number of live births during the same time interval	1,000
Postneonatal mortality rate	Number of deaths among children 28–364 days of age during a given time interval	Number of live births during the same time interval	1,000
Infant mortality rate	Number of deaths among children < 1 year of age during a given time interval	Number of live births during the same time interval	1,000
Maternal mortality rate	Number of deaths assigned to pregnancy-related causes during a given time interval	Number of live births during the same time interval	100,000

Crude mortality rate (crude death rate)

The crude mortality rate is the mortality rate from all causes of death for a population. In the United States in 2003, a total of 2,419,921 deaths occurred. The estimated population was 290,809,777. The crude mortality rate in 2003 was, therefore, $(2,419,921 / 290,809,777) \times 100,000$, or 832.1 deaths per 100,000 population.⁸

Cause-specific mortality rate

The cause-specific mortality rate is the mortality rate from a specified cause for a population. The numerator is the number of deaths attributed to a specific cause. The denominator remains the size of the population at the midpoint of the time period. The fraction is usually expressed per 100,000 population. In the United States in 2003, a total of 108,256 deaths were attributed to accidents (unintentional injuries), yielding a cause-specific mortality rate of 37.2 per 100,000 population.⁸

Age-specific mortality rate

An age-specific mortality rate is a mortality rate limited to a particular age group. The numerator is the number of deaths in that age group; the denominator is the number of persons in that age group in the population. In the United States in 2003, a total of 130,761 deaths occurred among persons aged 25–44 years, or an age-specific mortality rate of 153.0 per 100,000 25–44 year olds.⁸ Some specific types of age-specific mortality rates are neonatal, postneonatal, and infant mortality rates, as described in the following sections.

Infant mortality rate

The infant mortality rate is perhaps the most commonly used measure for comparing health status among nations. It is calculated as follows:

$$\frac{\text{Number of deaths among children < 1 year of age reported during a given time period}}{\text{Number of live births reported during the same time period}} \times 1,000$$

The infant mortality rate is generally calculated on an annual basis. It is a widely used measure of health status because it reflects the health of the mother and infant during pregnancy and the year thereafter. The health of the mother and infant, in turn, reflects a wide variety of factors, including access to prenatal care, prevalence of prenatal maternal health behaviors (such as alcohol

or tobacco use and proper nutrition during pregnancy, etc.), postnatal care and behaviors (including childhood immunizations and proper nutrition), sanitation, and infection control.

Is the infant mortality rate a ratio? Yes. Is it a proportion? No, because some of the deaths in the numerator were among children born the previous year. Consider the infant mortality rate in 2003. That year, 28,025 infants died and 4,089,950 children were born, for an infant mortality rate of 6.951 per 1,000.⁸ Undoubtedly, some of the deaths in 2003 occurred among children born in 2002, but the denominator includes only children born in 2003.

Is the infant mortality rate truly a rate? No, because the denominator is not the size of the mid-year population of children < 1 year of age in 2003. In fact, the age-specific death rate for children < 1 year of age for 2003 was 694.7 per 100,000.⁸ Obviously the infant mortality rate and the age-specific death rate for infants are very similar (695.1 versus 694.7 per 100,000) and close enough for most purposes. They are not exactly the same, however, because the estimated number of infants residing in the United States on July 1, 2003 was slightly larger than the number of children born in the United States in 2002, presumably because of immigration.

Neonatal mortality rate

The neonatal period covers birth up to but not including 28 days. The numerator of the neonatal mortality rate therefore is the number of deaths among children under 28 days of age during a given time period. The denominator of the neonatal mortality rate, like that of the infant mortality rate, is the number of live births reported during the same time period. The neonatal mortality rate is usually expressed per 1,000 live births. In 2003, the neonatal mortality rate in the United States was 4.7 per 1,000 live births.⁸

Postneonatal mortality rate

The postneonatal period is defined as the period from 28 days of age up to but not including 1 year of age. The numerator of the postneonatal mortality rate therefore is the number of deaths among children from 28 days up to but not including 1 year of age during a given time period. The denominator is the number of live births reported during the same time period. The postneonatal mortality rate is usually expressed per 1,000 live births. In 2003, the postneonatal mortality rate in the United States was 2.3 per 1,000 live births.⁸

Maternal mortality rate

The maternal mortality rate is really a ratio used to measure mortality associated with pregnancy. The numerator is the number of deaths during a given time period among women while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes. The denominator is the number of live births reported during the same time period. Maternal mortality rate is usually expressed per 100,000 live births. In 2003, the U.S. maternal mortality rate was 8.9 per 100,000 live births.⁸

Sex-specific mortality rate

A sex-specific mortality rate is a mortality rate among either males or females. Both numerator and denominator are limited to the one sex.

Race-specific mortality rate

A race-specific mortality rate is a mortality rate related to a specified racial group. Both numerator and denominator are limited to the specified race.

Combinations of specific mortality rates

Mortality rates can be further stratified by combinations of cause, age, sex, and/or race. For example, in 2002, the death rate from diseases of the heart among women ages 45–54 years was 50.6 per 100,000.⁹ The death rate from diseases of the heart among men in the same age group was 138.4 per 100,000, or more than 2.5 times as high as the comparable rate for women. These rates are a cause-, age-, and sex-specific rates, because they refer to one cause (diseases of the heart), one age group (45–54 years), and one sex (female or male).

EXAMPLE: Calculating Mortality Rates

Table 3.5 provides the number of deaths from all causes and from accidents (unintentional injuries) by age group in the United States in 2002. Review the following rates. Determine what to call each one, then calculate it using the data provided in Table 3.5.

a. Unintentional-injury-specific mortality rate for the entire population

This is a cause-specific mortality rate.

$$\begin{aligned}\text{Rate} &= \frac{\text{number of unintentional injury deaths in the entire population}}{\text{estimated midyear population}} \times 100,000 \\ &= (106,742 / 288,357,000) \times 100,000 \\ &= 37.0 \text{ unintentional-injury-related deaths per } 100,000 \text{ population}\end{aligned}$$

b. All-cause mortality rate for 25–34 year olds

This is an age-specific mortality rate.

$$\begin{aligned}\text{Rate} &= \frac{\text{number of deaths from all causes among 25–34 year olds}}{\text{estimated midyear population of 25–34 year olds}} \times 100,000 \\ &= (41,355 / 39,928,000) \times 100,000 \\ &= 103.6 \text{ deaths per } 100,000 \text{ 25–34 year olds}\end{aligned}$$

c. All-cause mortality among males

This is a sex-specific mortality rate.

$$\begin{aligned}\text{Rate} &= \frac{\text{number of deaths from all causes among males}}{\text{estimated midyear population of males}} \times 100,000 \\ &= (1,199,264 / 141,656,000) \times 100,000 \\ &= 846.6 \text{ deaths per } 100,000 \text{ males}\end{aligned}$$

d. Unintentional-injury-specific mortality among 25- to 34-year-old males

This is a cause-specific, age-specific, and sex-specific mortality rate

$$\begin{aligned}\text{Rate} &= \frac{\text{number of unintentional injury deaths among 25–34 year old males}}{\text{estimated midyear population of 25–34 year old males}} \times 100,000 \\ &= (9,635 / 20,203,000) \times 100,000 \\ &= 47.7 \text{ unintentional-injury-related deaths per } 100,000 \text{ 25–34 year olds}\end{aligned}$$

Table 3.5 All-Cause and Unintentional Injury Mortality and Estimated Population by Age Group, For Both Sexes and For Males Alone, United States, 2002

Age group (years)	All Races, Both Sexes			All Races, Males		
	All Causes	Unintentional Injuries	Estimated Pop. (x 1000)	All Causes	Unintentional Injuries	Estimated Pop. (x 1000)
0–4	32,892	2,587	19,597	18,523	1,577	10,020
5–14	7,150	2,718	41,037	4,198	1713	21,013
15–24	33,046	15,412	40,590	24,416	11,438	20,821
25–34	41,355	12,569	39,928	28,736	9,635	20,203
35–44	91,140	16,710	44,917	57,593	12,012	22,367
45–54	172,385	14,675	40,084	107,722	10,492	19,676
55–64	253,342	8,345	26,602	151,363	5,781	12,784
65+	1,811,720	33,641	35,602	806,431	16,535	14,772
Not stated	357	85	0	282	74	0
Total	2,443,387	106,742	288,357	1,199,264	69,257	141,656

Data Source: Web-based Injury Statistics Query and Reporting System (WISQARS) [online database] Atlanta; National Center for Injury Prevention and Control. Available from: <http://www.cdc.gov/ncipc/wisqars>.



Exercise 3.3

In 2001, a total of 15,555 homicide deaths occurred among males and 4,753 homicide deaths occurred among females. The estimated 2001 midyear populations for males and females were 139,813,000 and 144,984,000, respectively.

1. Calculate the homicide-related death rates for males and for females.
2. What type(s) of mortality rates did you calculate in Question 1?
3. Calculate the ratio of homicide-mortality rates for males compared to females.
4. Interpret the rate you calculated in Question 3 as if you were presenting information to a policymaker.



Check your answers on page 3-51

Age-adjusted mortality rate: a mortality rate statistically modified to eliminate the effect of different age distributions in the different populations.

Age-adjusted mortality rates

Mortality rates can be used to compare the rates in one area with the rates in another area, or to compare rates over time. However, because mortality rates obviously increase with age, a higher mortality rate among one population than among another might simply reflect the fact that the first population is older than the second.

Consider that the mortality rates in 2002 for the states of Alaska and Florida were 472.2 and 1,005.7 per 100,000, respectively (see Table 3.6). Should everyone from Florida move to Alaska to reduce their risk of death? No, the reason that Alaska's mortality rate is so much lower than Florida's is that Alaska's population is considerably younger. Indeed, for seven age groups, the age-specific mortality rates in Alaska are actually higher than Florida's.

To eliminate the distortion caused by different underlying age distributions in different populations, statistical techniques are used to adjust or standardize the rates among the populations to be compared. These techniques take a weighted average of the age-specific mortality rates, and eliminate the effect of different age distributions among the different populations. Mortality rates computed with these techniques are **age-adjusted** or **age-standardized mortality rates**. Alaska's 2002 age-adjusted mortality rate (794.1 per 100,000) was higher than Florida's (787.8 per 100,000), which is not surprising given that 7 of 13 age-specific mortality rates were higher in Alaska than Florida.

Death-to-case ratio

Definition of death-to-case ratio

The death-to-case ratio is the number of deaths attributed to a particular disease during a specified time period divided by the number of new cases of that disease identified during the same time period. The death-to-case ratio is a ratio but not necessarily a proportion, because some of the deaths that are counted in the numerator might have occurred among persons who developed disease in an earlier period, and are therefore not counted in the denominator.

Table 3.6 All-Cause Mortality by Age Group, Alaska and Florida, 2002

Age Group (years)	ALASKA			FLORIDA		
	Population	Deaths	Death Rate (per 100,000)	Population	Deaths	Death Rate (per 100,000)
<1	9,938	55	553.4	205,579	1,548	753.0
1–4	38,503	12	31.2	816,570	296	36.2
5–9	50,400	6	11.9	1,046,504	141	13.5
10–14	57,216	24	41.9	1,131,068	219	19.4
15–19	56,634	43	75.9	1,073,470	734	68.4
20–24	42,929	63	146.8	1,020,856	1,146	112.3
25–34	84,112	120	142.7	2,090,312	2,627	125.7
35–44	107,305	280	260.9	2,516,004	5,993	238.2
45–54	103,039	427	414.4	2,225,957	10,730	482.0
55–64	52,543	480	913.5	1,694,574	16,137	952.3
65–74	24,096	502	2,083.3	1,450,843	28,959	1,996.0
65–84	11,784	645	5,473.5	1,056,275	50,755	4,805.1
85+	3,117	373	11,966.6	359,056	48,486	13,503.7
Unknown	NA	0	NA	NA	43	NA
Total	3,030	3,030	472.2	16,687,068	167,814	1,005.7
Age-adjusted rate:			794.1			787.8

Data Source: Web-based Injury Statistics Query and Reporting System (WISQARS) [online database] Atlanta; National Center for Injury Prevention and Control. Available from: <http://www.cdc.gov/ncipc/wisqars>.

Method for calculating death-to-case ratio

$$\frac{\text{Number of deaths attributed to a particular disease during specified period}}{\text{Number of new cases of the disease identified during the specified period}} \times 10^n$$

EXAMPLE: Calculating Death-to-Case Ratios

Between 1940 and 1949, a total of 143,497 incident cases of diphtheria were reported. During the same decade, 11,228 deaths were attributed to diphtheria. Calculate the death-to-case ratio.

$$\text{Death-to-case ratio} = 11,228 / 143,497 \times 1 = 0.0783$$

or

$$= 11,228 / 143,497 \times 100 = 7.83 \text{ per } 100$$



Exercise 3.4

Table 3.7 provides the number of reported cases of diphtheria and the number of diphtheria-associated deaths in the United States by decade. Calculate the death-to-case ratio by decade. Describe the data in Table 3.7, including your results.

Table 3.7 Number of Cases and Deaths from Diphtheria by Decade, United States, 1940–1999

Decade	Number of New Cases	Number of Deaths	Death-to-case Ratio (x 100)
1940–1949	143,497	11,228	7.82
1950–1959	23,750	1,710	
1960–1969	3,679	390	
1970–1979	1,956	90	
1980–1989	27	3	
1990–1999	22	5	

Data Sources: Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 2001. MMWR 2001;50(No. 53).

Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 1998. MMWR 1998;47(No. 53).

Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 1989. MMWR 1989;38 (No. 53).



Check your answers on page 3-52

Case-fatality rate

The case-fatality rate is the proportion of persons with a particular condition (cases) who die from that condition. It is a measure of the severity of the condition. The formula is:

$$\frac{\text{Number of cause-specific deaths among the incident cases}}{\text{Number of incident cases}} \times 10^n$$

The case-fatality rate is a proportion, so the numerator is restricted to deaths among people included in the denominator. The time periods for the numerator and the denominator do not need to be the same; the denominator could be cases of HIV/AIDS diagnosed during the calendar year 1990, and the numerator, deaths among those diagnosed with HIV in 1990, could be from 1990 to the present.

EXAMPLE: Calculating Case-Fatality Rates

In an epidemic of hepatitis A traced to green onions from a restaurant, 555 cases were identified. Three of the case-patients died as a result of their infections. Calculate the case-fatality rate.

$$\text{Case-fatality rate} = (3 / 555) \times 100 = 0.5\%$$

The case-fatality rate is a proportion, not a true rate. As a result, some epidemiologists prefer the term **case-fatality ratio**.

The concept behind the case-fatality rate and the death-to-case ratio is similar, but the formulations are different. The death-to-case ratio is simply the number of cause-specific deaths that occurred during a specified time divided by the number of new cases of that disease that occurred during the same time. The deaths included in the numerator of the death-to-case ratio are not restricted to the new cases in the denominator; in fact, for many diseases, the deaths are among persons whose onset of disease was years earlier. In contrast, in the case-fatality rate, the deaths included in the numerator are restricted to the cases in the denominator.

Proportionate mortality

Definition of proportionate mortality

Proportionate mortality describes the proportion of deaths in a specified population over a period of time attributable to different causes. Each cause is expressed as a percentage of all deaths, and

the sum of the causes must add to 100%. These proportions are not mortality rates, because the denominator is all deaths rather than the population in which the deaths occurred.

Method for calculating proportionate mortality

For a specified population over a specified period,

$$\frac{\text{Deaths caused by a particular cause}}{\text{Deaths from all causes}} \times 100$$

The distribution of primary causes of death in the United States in 2003 for the entire population (all ages) and for persons ages 25–44 years are provided in Table 3.1. As illustrated in that table, accidents (unintentional injuries) accounted for 4.3% of all deaths, but 21.6% of deaths among 25–44 year olds.⁸

Sometimes, particularly in occupational epidemiology, proportionate mortality is used to compare deaths in a population of interest (say, a workplace) with the proportionate mortality in the broader population. This comparison of two proportionate mortalities is called a **proportionate mortality ratio**, or PMR for short. A PMR greater than 1.0 indicates that a particular cause accounts for a greater proportion of deaths in the population of interest than you might expect. For example, construction workers may be more likely to die of injuries than the general population.

However, PMRs can be misleading, because they are not based on mortality rates. A low cause-specific mortality rate in the population of interest can elevate the proportionate mortalities for all of the other causes, because they must add up to 100%. Those workers with a high injury-related proportionate mortality very likely have lower proportionate mortalities for chronic or disabling conditions that keep people out of the workforce. In other words, people who work are more likely to be healthier than the population as a whole — this is known as the healthy worker effect.



Exercise 3.5

Using the data in Table 3.8, calculate the missing proportionate mortalities for persons ages 25–44 years for diseases of the heart and assaults (homicide).

Table 3.8 Number, Proportion (Percentage), and Ranking of Deaths for Leading Causes of Death, All Ages and 25–44 Year Age Group, United States, 2003

	Number	All Ages		Ages 25–44 Years		
		Percentage	Rank	Number	Percentage	Rank
All causes	2,443,930	100.0		128,924	100.0	
Diseases of heart	684,462	28.0	1	16,283		3
Malignant neoplasms	554,643	22.7	2	19,041	14.8	2
Cerebrovascular disease	157,803	6.5	3	3,004	2.3	8
Chronic lower respiratory diseases	126,128	5.2	4	401	0.3	*
Accidents (unintentional injuries)	105,695	4.3	5	27,844	21.6	1
Diabetes mellitus	73,965	3.0	6	2,662	2.1	9
Influenza & pneumonia	64,847	2.6	7	1,337	1.0	10
Alzheimer's disease	63,343	2.6	8	0	0.0	*
Nephritis, nephrotic syndrome, nephrosis	33,615	1.4	9	305	0.2	*
Septicemia	34,243	1.4	10	328	0.2	*
Intentional self-harm (suicide)	30,642	1.3	11	11,251	8.7	4
Chronic liver disease and cirrhosis	27,201	1.1	12	3,288	2.6	7
Assault (homicide)	17,096	0.7	13	7,367		5
HIV disease	13,544	0.5	*	6,879	5.3	6
All other	456,703	18.7		29,480	22.9	

* Not among top ranked causes

Data Sources: CDC. Summary of notifiable diseases, United States, 2003. MMWR 2005;2(No. 54).

Hoyert DL, Kung HC, Smith BL. Deaths: Preliminary data for 2003. National Vital Statistics Reports; vol. 53 no 15. Hyattsville, MD: National Center for Health Statistics 2005: 15, 27.



Check your answers on page 3-52