4

The Occurrence of Disease: II. Mortality and Other Measures of Disease Impact

You do not die from being born, nor from having lived, nor from old age. You die from something. ... There is no such thing as a natural death: Nothing that happens to a man is ever natural, since his presence calls the world into question. All men must die: but for every man his death is an accident and, even if he knows it and consents to it, an unjustifiable violation.

***— Simone de Beauvoir, writing of her mother's death, in A Very Easy Death*[1]**

LEARNING OBJECTIVES

- To compare different measures of mortality, including mortality rates, case-fatality, proportionate mortality, and years of potential life lost.
- To show when mortality can approximate the risk of disease.
- To introduce issues that arise in comparing mortality across two or more populations.
- To define, calculate, and interpret direct and indirect age-adjusted mortality rates.
- To introduce other measures of disease impact.

Mortality is of great interest for several reasons. First of all, death is the ultimate experience that every human being is destined to experience. Death is clearly of tremendous importance to each person, including questions of when and how death will occur and whether there is any way to delay it. From the standpoint of studying disease occurrence, expressing mortality in quantitative terms can pinpoint differences in the risk of dying from a disease between people in different geographic areas and subgroups in the population. Mortality rates can serve as measures of disease severity and can help us determine whether the treatments for a disease have become more effective over time. In addition, given the problem that often arises in identifying new cases of a disease, mortality rates may serve as surrogates for incidence rates when the disease being studied is a severe and lethal one (that is, of short duration between detection and death). This chapter will address the quantitative expression of mortality and the uses of such measures in epidemiologic studies.

MEASURES OF MORTALITY

Fig. 4.1 shows the number of cancer deaths from 1999 to 2019 in the United States. Clearly, the absolute *number* of people dying from cancer is seen increasing significantly through the year 2019, but from this graph, we cannot say that the *risk* of dying from cancer is increasing, because the only data that we have in this graph are numbers of deaths (numerators); we do not have denominators (populations at risk). If, for example, the size of the US population is also increasing at the same rate, the risk of dying from cancer does not change.

For this reason, if we wish to address the risk of dying, we must deal with rates. Fig. 4.2 shows mortality rates for several types of cancer in men from 1930 to 2018. Note that the rates are expressed as deaths per 100,000 population. The most dramatic increase is in deaths from lung cancer. This increase is clearly of epidemic proportions and, tragically, lung cancer is a

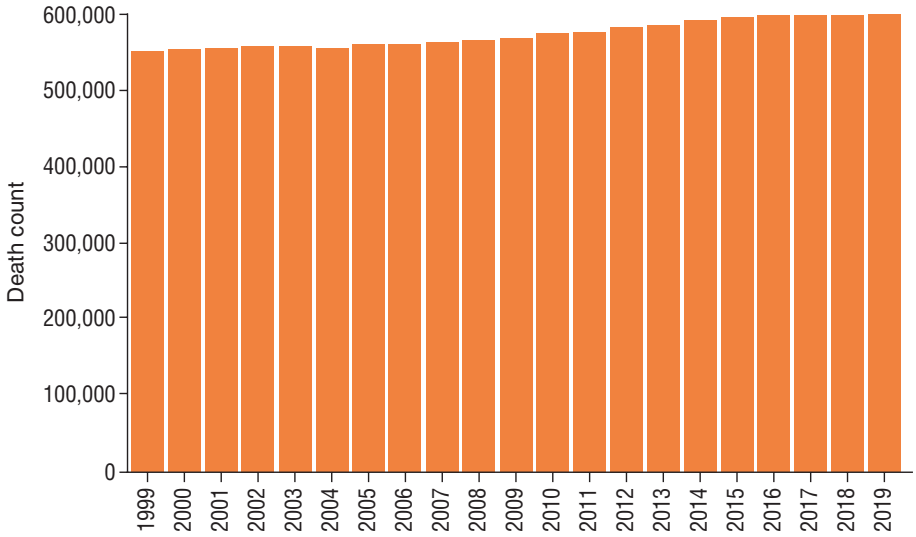
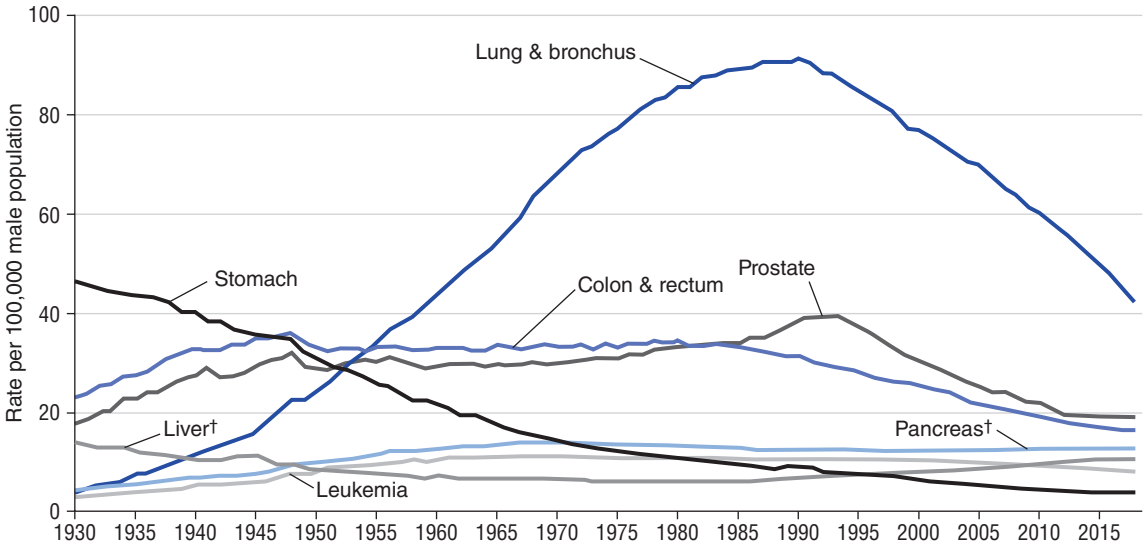


Fig. 4.1 Annual number of cancer deaths, 1999–2019. (From Centers for Disease Control and Prevention. United States Cancer Statistics: Data Visualizations. URL: https://gis.cdc.gov/Cancer/USCS/#/Trends/).



*Per 100,000, age adjusted to the 2000 US standard population.
†Mortality rates for pancreatic and liver cancers are increasing.
Note: Due to changes in ICD coding, numerator information has changed over time for cancers of the liver, lung and bronchus, and colon and rectum.

Fig. 4.2 Trends in age-adjusted cancer death rates* by site, males, US, 1930–2018. *ICD,* International Classification of Diseases. (From US Mortality Volumes 1930 to 1959, US Mortality Data 1960 to 2018, National Center for Health Statistics, Centers for Disease Control and Prevention. © 2021, American Cancer Society, Inc., Surveillance Research.)

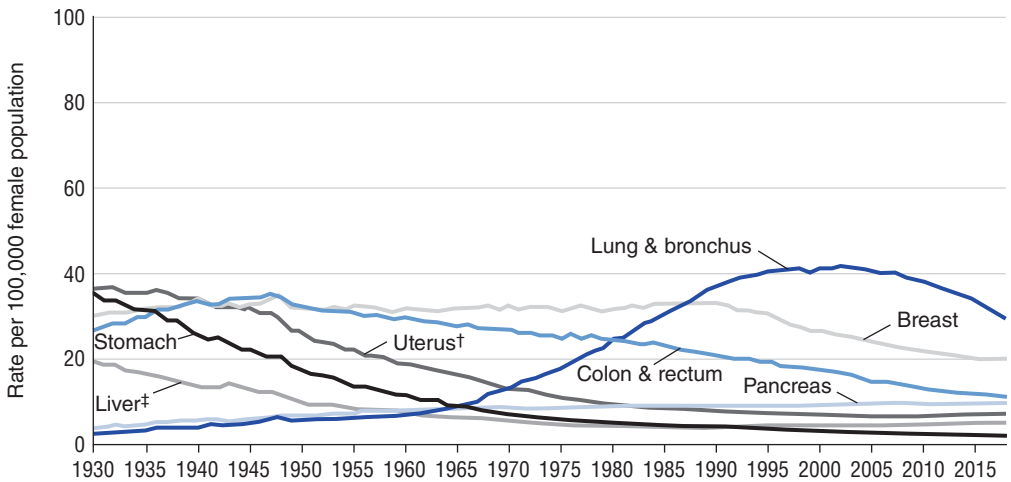
preventable cause of death. Fortunately, since the mid-1990s, lung cancer mortality has declined, paralleling earlier decreases in rates of smoking among men. Other cancers are also of interest. Age-adjusted mortality from prostate cancer also peaked in the mid-1990s and has declined since. Cancers of the colon and rectum have declined over many years (no doubt attributable to the expansion of screening to detect these cancers at an earlier stage). The rate of death from stomach cancer has declined dramatically since 1930, although the precise explanation is not known. It has been suggested that the decline may be the result of the increased availability of refrigeration, which decreased the need to smoke foods and thereby decreased human exposure to carcinogens produced in the smoking process. Another possible cause is improved hygiene, which may have reduced the incidence of *Helicobacter pylori* infections that have been implicated in the etiology (or cause) of stomach cancer.

Fig. 4.3 shows a similar presentation for cancer mortality in women for the period 1930 to 2018. Breast cancer mortality remained at essentially the same level for many years but has declined since the early 1990s until 2018. It would be desirable to study changes in the incidence of breast cancer. Such a study is difficult, however, because with aggressive public education campaigns encouraging women to have mammograms and perform breast self-examination, many breast cancers may be detected today at much earlier stages that might have gone undetected years ago. Nevertheless, available evidence suggests that the true incidence of breast cancer in women may have increased for many years but then decreased from 2001 to 2018.

Uterine cancer mortality has declined, perhaps because of earlier detection and diagnosis. Lung cancer mortality in women has increased, and lung cancer has exceeded breast cancer as a cause of death in women. Lung cancer is almost completely preventable, being mostly due to a lifestyle habit, cigarette smoking, which has been voluntarily adopted by many women; today it is the leading cause of cancer death in women in the United States.

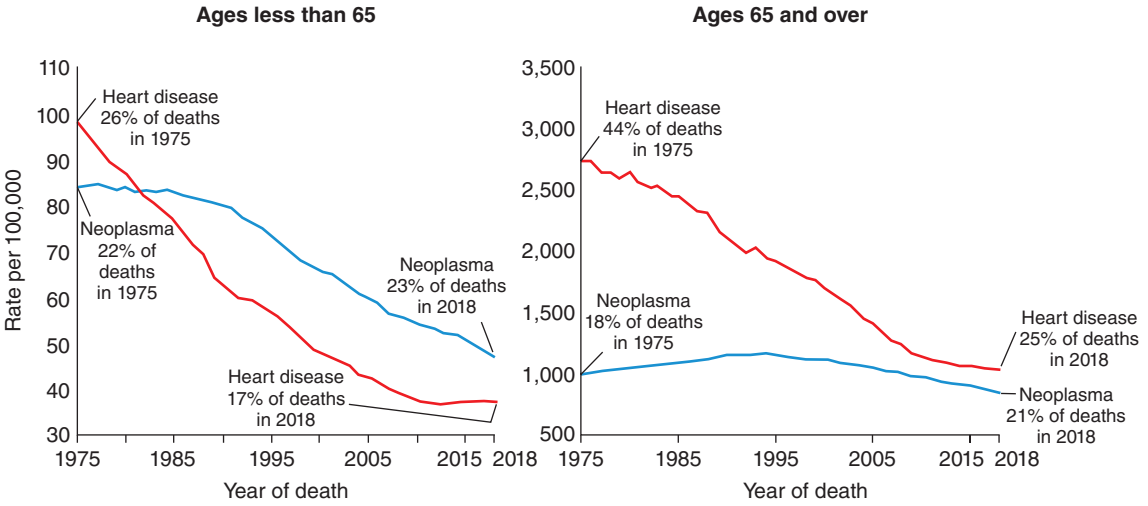
We may be particularly interested in mortality relating to age. Fig. 4.4 shows death rates from cancer and from heart disease for people younger than 65 and for those 65 or older. Cancer is the leading cause of death in men and women younger than 65 years, but above age 65, heart disease clearly exceeds cancer as a cause of death.

Fig. 4.5 shows the causes of death worldwide for children younger than 5 years in 2019. The leading



*Per 100,000, age adjusted to the 2000 US standard population. Rate exclude deaths in Puerto Rico and other US territories.
†Uterus refers to uterine cervix and uterine corpus combined.
‡The mortality rates for liver cancer is increasing.
Note: Due to changes in ICD coding, numerator information has changed for cancers of the liver, lung and bronchus, colon and rectum, and uterus

Fig. 4.3 Trends in age-adjusted cancer death rates* by site, females, US, 1930–2018. *ICD,* International Classification of Diseases. (From US Mortality Volumes 1930 to 1959, US Mortality Data 1960 to 2018, National Center for Health Statistics, Centers for Disease Control and Prevention. © 2021 American Cancer Society, Inc., Surveillance Research.)



Rates are per 100,000 and age-adjusted to the 2000 US Std population (19 age groups - Census P25-1103).

Fig. 4.4 US death rates, 1975–2018 heart disease compared to neoplasms, by age at death. (From US Mortality Files, National Center for Health Statistics, Centers for Disease Control and Prevention.)

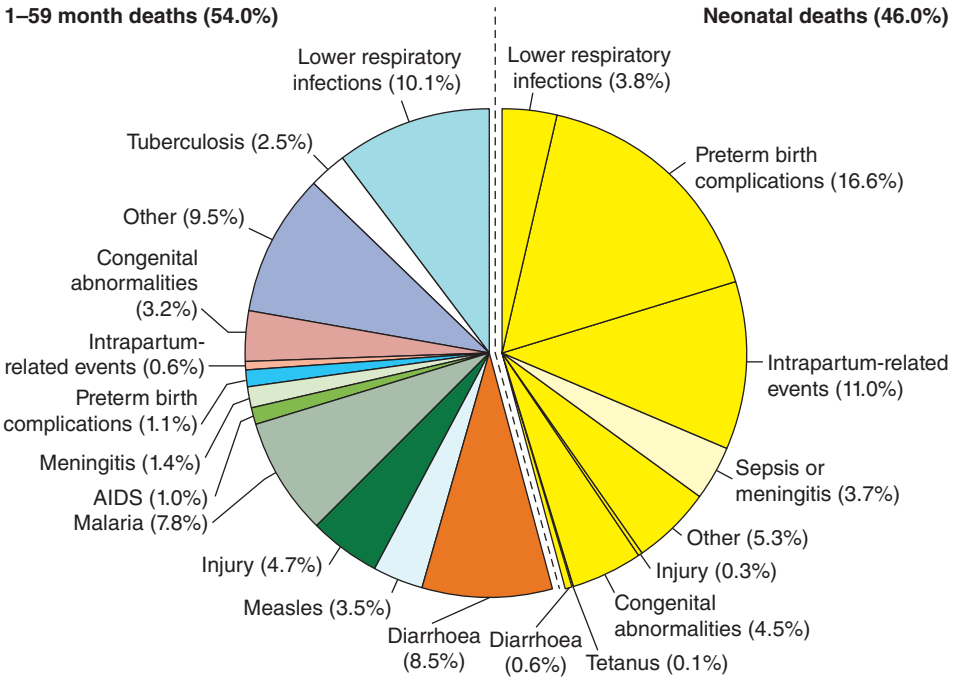


Fig. 4.5 Global causes of under-5 deaths in 2019. Deaths of neonates (aged 0–27 days) are on the right-hand side and deaths of children aged 1–59 months are on the left-hand side. (From Perin J, Mulick A, Yeung D, et al. Global, regional, and national causes of under-5 mortality in 2000-19: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet Child Adolesc Health.* 2022;6(2):106–115.)

causes of death among children under 5 years of age in 2019 were preterm birth complications, lower respiratory infections, intrapartum-related complications, diarrhea, and congenital abnormalities. Neonatal deaths accounted for 46% of under-5 deaths in 2019. Infectious diseases accounted for over half of the 5.9 million deaths of children under age 5, with the largest percentages due to lower respiratory infections, diarrhea, and malaria.

Mortality Rates

How is mortality expressed in quantitative terms? Let us examine some types of mortality rates. The first is the annual death rate, or mortality rate, from all causes combined:

$$\text{Annual mortality rate for all causes (per 100,000 population)} = \frac{\text{Total no. of deaths from all causes in 1 year}}{\text{No. of persons in the population at midyear}} \times 100{,}000$$

Note that because the population changes over time, the number of persons in the population at midyear is generally used as an approximation of average population in that year.

The same principles mentioned in the discussion of morbidity apply to mortality; for a mortality rate to make sense, anyone in the group represented by the denominator must have the potential to enter the group represented by the numerator.

We may not always be interested in a rate for the entire population; perhaps we are interested only in a certain age group, in men or in women, or in one ethnic group. Thus, if we are interested in mortality in children younger than 10 years, we can calculate a rate specifically for that group:

$$\text{Annual mortality rate from all causes for children younger than 10 years of age (per 1,000 population)} = \frac{\text{No. of deaths from all causes in 1 year in children younger than 10 years of age}}{\text{No of children in the population younger than 10 years of age at midyear}} \times 1{,}000$$

In putting a restriction on age, for example, the same restriction must apply to *both* the numerator and the denominator, so that every person in the denominator group will be at risk for entering the numerator group. When such a restriction is placed on a rate, it is called a *specific rate*. The above rate, then, is an *age-specific mortality rate*.

We could also place a restriction on a rate by specifying a diagnosis, and thus limit the rate to deaths from a certain disease, that is, a *disease-specific* or a *cause-specific rate*. For example, if we are interested in mortality from lung cancer, we would calculate it in the following manner:

$$\text{Annual mortality rate from lung cancer (per 1,000 population)} = \frac{\text{No. of deaths from lung cancer in 1 year}}{\text{No. of persons in the population at midyear}} \times 1{,}000$$

We can also place restrictions on more than one characteristic simultaneously, for example, age and cause of death, as follows:

$$\text{Annual mortality rate from leukemia in children} < 10 \text{ years of age (per 1,000 population)} = \frac{\text{No. of deaths from leukemia in 1 year in children} < 10 \text{ years of age}}{\text{No. of children in the population} < 10 \text{ years of age at midyear}} \times 1{,}000$$

Time must also be specified in any mortality rate, in the case above, in one year (which is usually included in the title, e.g., 2022). Mortality can be calculated over 1 year, 5 years, or longer. The period selected is arbitrary, but it must be specified precisely.

Case-Fatality

We must distinguish between a *mortality rate* and *case-fatality*. Case-fatality is calculated as follows:

$$\text{Case-fatality (\%)} = \frac{\text{No. of individuals dying during a specified period of time after disease onset or diagnosis}}{\text{No. of individuals with the specified disease}} \times 100$$

In other words, what percentage of people *who have a certain disease* die within a certain time after their disease was diagnosed? Ideally, we would like to use the date of disease onset as the beginning of the time period specified in the numerator. However, date of disease

onset is often hard to standardize since many diseases develop insidiously (without symptoms) over a long period of time. As a result, in many chronic diseases, it may be difficult to determine precisely when the disease process began. For example, many patients with arthritis cannot recall when their joint pain first began. In practice therefore we often use date of diagnosis as a surrogate measure for date of disease onset because the exact date of diagnosis can generally be documented from available medical records (note, however, that the medical record is only as reliable as the person entering the data). If the information is to be obtained from respondents, it is worth noting that if the disease in question is a serious one, the date on which the diagnosis was given may well have been a life-changing date for the patient and not easily forgotten.

What is the difference between case-fatality and a mortality rate? In a mortality rate, the denominator represents the entire population at risk of dying from the disease, including both those who have the disease and those who do not have the disease (but who are *at risk* of developing the disease). In case-fatality, however, the denominator is limited to those who *already have the disease*. Thus, case-fatality is a measure of the *severity* or *fatality* of the disease. It can also be used to measure any benefits of a new therapy; as therapy improves, case-fatality would be expected to decline.

The numerator of case-fatality should ideally be restricted to deaths *from that disease*. However, it is not always easy to distinguish between deaths from that disease and deaths from other causes. For example, a person with alcohol-use disorder may die in a car accident; however, the death may or may not be related to alcohol intake.

Let us look at a hypothetical example to clarify the difference between mortality and case-fatality (Box 4.1).

BOX 4.1 Comparison of Mortality Rate With Case-Fatality in the Same Year

Assume a population of 100,000 people of whom 20 are sick with disease X, and in 1 year, 18 of the 20 die from disease X

$$\text{Mortality rate from disease } X = \frac{18}{100{,}000} = 0.00018,$$

or 0.18%

$$\text{Case-fatality from disease } X = \frac{18}{20} = 0.9, \text{ or } 90\%$$

Assume that in a population of 100,000 persons, 20 have disease X. In 1 year, 18 people die from that disease. The mortality is very low (0.018%) because the disease is rare; however, once a person has the disease, their chances of dying are great (90%).

Since case-fatality traditionally does not include a time unit, it is considered a proportion, but commonly *mistakenly* reported as *rate*.

Proportionate Mortality

Another commonly used measure of mortality is proportionate mortality, which is not a rate. The proportionate mortality from cardiovascular disease in the United States in 2014 is defined as follows:

$$\text{Proportionate mortality from cardiovascular diseases in the US in 2014 (\%)} = \frac{\text{No. of deaths from cardiovascular diseases in the US in 2014}}{\text{Total deaths in the US in 2014}} \times 100$$

In other words, of all deaths in the United States, what *proportion* was caused by cardiovascular disease? Fig. 4.6 shows proportionate mortality from heart disease in the total population and by race. In each age group, the full bar represents all deaths (100%), and deaths from heart disease are indicated by the dark blue portion. We see that the *proportion* of deaths from heart disease increases with age. However, this does not tell us that the *risk* of death from heart disease is also increasing. This is demonstrated in the following examples.

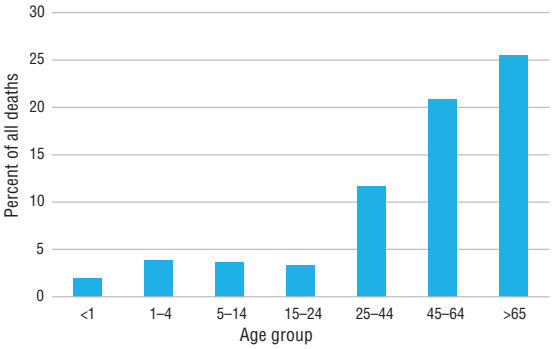


Fig. 4.6 Deaths from heart disease as a percentage of deaths from all causes, by age group, United States, 2014. (From National Center for Health Statistics [NCHS]. *Data from Health, United States, 2015, With Special Feature on Racial and Ethnic Health Disparities*. Hyattsville, MD: NCHS; 2016.)

Table 4.1 shows all deaths and deaths from heart disease in two communities, A and B. All-cause mortality in community A is twice that in community B. When we look at proportionate mortality, we find that 10% of the deaths in community A and 20% of the deaths in community B are due to heart disease. Does this tell us that the risk of dying from heart disease is twice as high in community B as it is in A? The answer is no. When the mortality rates from heart disease are calculated for the two communities (10% of 30/1,000 and 20% of 15/1,000), we find that the mortality rates are identical.

If we observe a change in proportionate mortality from a certain disease over time, the change may be due not to changes in mortality from that disease, but to changes in the mortality of some other disease. Let us consider a hypothetical example: in Table 4.2, we see mortality rates from heart disease, cancer, and other causes in a population in an early period and a later period. First, compare the mortality rates in the two time periods: mortality from heart disease doubled over time (from 40/1,000 to 80/1,000), but mortality rates from cancer and from all other causes (20/1,000) did not change. However, if we now examine the proportionate mortality from each cause, we see that the proportionate mortality from cancer and from other causes has decreased in the population, but only because the proportionate mortality from heart disease has increased. Thus, if the proportion of one segment of the mortality "pie" increases, there will necessarily be a decrease in the proportion of some other segment (Fig. 4.7). Another view of this is seen in Fig. 4.8.

TABLE 4.1 Comparison of Mortality Rate and Proportionate Mortality: I. Deaths From Heart Disease in Two Communities

	Community A	Community B
Mortality rate from all causes	30/1,000	15/1,000
Proportionate mortality from heart disease	10%	20%
Mortality rate from heart disease	3/1,000	3/1,000

TABLE 4.2 Hypothetical Example of Mortality Rates and Proportionate Mortality in Two Periods

	EARLY PERIOD		LATER PERIOD	
Cause of Death	Mortality Rate	Proportionate Mortality	Mortality Rate	Proportionate Mortality
Heart disease	40/1,000	50%	80/1,000	66.7%
Cancer	20/1,000	25%	20/1,000	16.7%
All other causes	20/1,000	25%	20/1,000	16.7%
All deaths	80/1,000	100%	120/1,000	100.0%

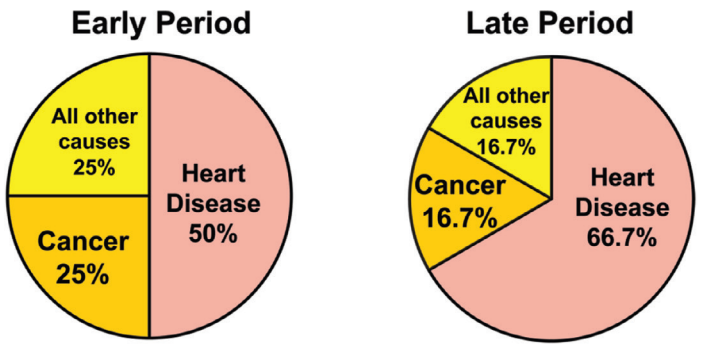


Fig. 4.7 Hypothetical example of proportionate mortality: changes in proportionate mortality from heart disease, cancer, and other causes from the early period to the late period.

Fig. 4.8 Understanding proportionate mortality. (Family Circus © 2002 Bill Keane, Inc. Distributed by King Features Syndicate, Inc.)

TABLE 4.3 Comparison of Mortality Rate and Proportionate Mortality: II. Deaths From Heart Disease in Two Communities

	Community A	Community B
Mortality rate from all causes	20/1,000	10/1,000
Proportionate mortality from heart disease	30%	30%
Mortality rate from heart disease	6/1,000	3/1,000

As seen in the example in Table 4.3, if all-cause mortality rates differ, cause-specific mortality rates can differ significantly, even when the proportionate mortality is the same. Thus, these examples show that, although proportionate mortality can give us a quick look at the major causes of death, it cannot tell us the risk of dying from a disease. For that, we need a mortality rate.

Years of Potential Life Lost

Another mortality index, years of potential life lost (YPLL), has been increasingly used for setting health priorities. YPLL is a measure of premature mortality, or early death. YPLL recognizes that death occurring in a person at a younger age clearly involves a greater loss of future productive years than death occurring at an older age. Two steps are involved in this calculation: in the first step, for each cause, each deceased person's age at death is subtracted from a predetermined (or "average") age at death. In the United States, this predetermined "standard" age is usually 75 years. Thus, an infant dying at 1 year of age has lost 74 years of life (75 to 1), but a person dying at 50 years of age has lost 25 years of life (75 to 50). Thus, the younger the age at which death occurs, the more years of potential life are lost. In the second step, the "*years of potential life lost*" for each individual are then added together to yield the total YPLL for the specific cause of death. When looking at reports that use YPLL, it is important to note what assumptions the author has made, including what predetermined standard age has been selected.

Fig. 4.9 shows the YPLL in the United States before age 75 years in 2020. The top bar shows the total YPLL from all causes (100%), and the bars below show the individual YPLL from each leading cause of death, with the percentage of YPLL from all causes for which it accounts. We see that the greatest single source of YPLL was unintentional injuries, which, in the same year, was the fourth leading cause of death by its mortality rate (see Fig. 1.2 and Table 1.1). In 2020, the ranking of malignant neoplasms by its mortality rate was second, while its ranking by YPLL was second. This discrepancy results from the fact that injury is the leading cause of death up to age 34 years, and therefore it accounts for a large proportion of YPLL.

Fig. 4.10 shows YPLL from unintentional injuries before age 75 years. We see that the YPLL from motor vehicle traffic accidents accounts for over one fourth of the YPLL. Thus, if we want to have an impact on YPLL, we should address this specific cause of injury related to motor vehicles.

Table 4.4 shows a ranking of causes of death in the United States for 2020 by YPLL, together with crude number of deaths for each cause. By cause-specific number of deaths, suicide is ranked seventh, but by YPLL, it ranked fourth. This reflects the fact that a large proportion of suicide-related deaths occur in young persons.

YPLL can assist in three important public health functions: establishing research and resource priorities,

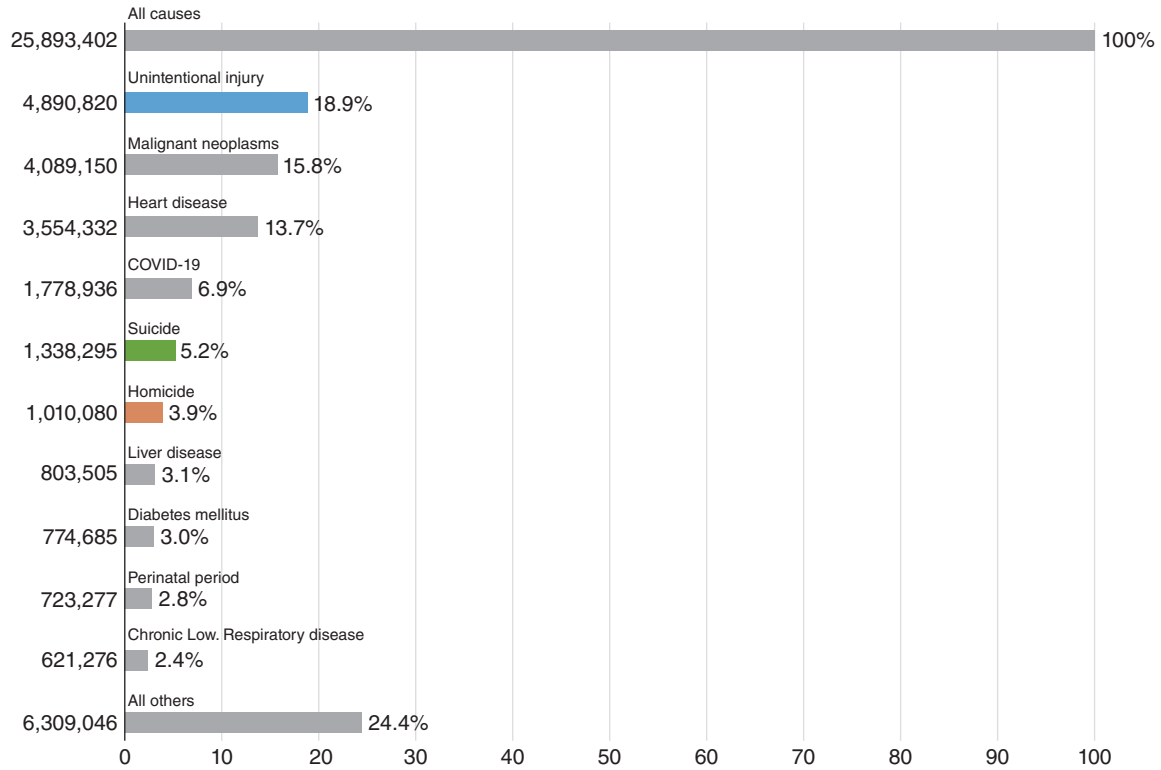


Fig. 4.9 Years of potential life lost (YPLL) before age 75 for 10 leading causes of death 2020, United States, both sexes, all races. (From Centers for Disease Control and Prevention. WISQARS Leading Causes of Death Visualization Tool. URL: https://wisqars.cdc.gov/data/lcd/home)

surveillance of temporal trends in premature mortality, and evaluating the effectiveness of program interventions.[2]

Why Look at Mortality?

Mortality is clearly an index of the severity of a disease from both clinical and public health standpoints, but mortality can also be used as an index of the risk of disease, as shown in Figs. 4.2 and 4.3. In general, mortality data are easier to obtain than incidence data for a given disease, and it therefore may be more feasible to use mortality data as a proxy indicator for incidence. In addition, and unless you are interested in cause-specific mortality, mortality data does not need adjudication or standardization, since all individuals being studied can either be classified as alive or dead. However, when a disease is mild and not fatal, mortality may not be a good index of incidence. A mortality rate is a good reflection of the incidence rate under two conditions: first, when the case-fatality is high (as in untreated rabies), and second, when the duration of disease (survival) is short (as in COVID-19). Under these conditions, mortality is a good measure of incidence, and thus a measure of the risk of disease. For example, cancer of the pancreas is a highly lethal disease: death generally occurs within a few months of diagnosis, and long-term survival is rare. Thus, unfortunately, mortality from pancreatic cancer is a good surrogate for incidence of the disease.

Fig. 4.11 shows mortality trends in the United States from 2000 to 2017 by race. It is evident that the mortality rates for Black and White individuals have gone down, but yet there is a clear disparity between the two races, as shown by the consistent gap between the two curves. Fig. 4.12 shows mortality trends in the United States from 1955 to 2019 by gender and age group. In both panels for males and females, we can see that there

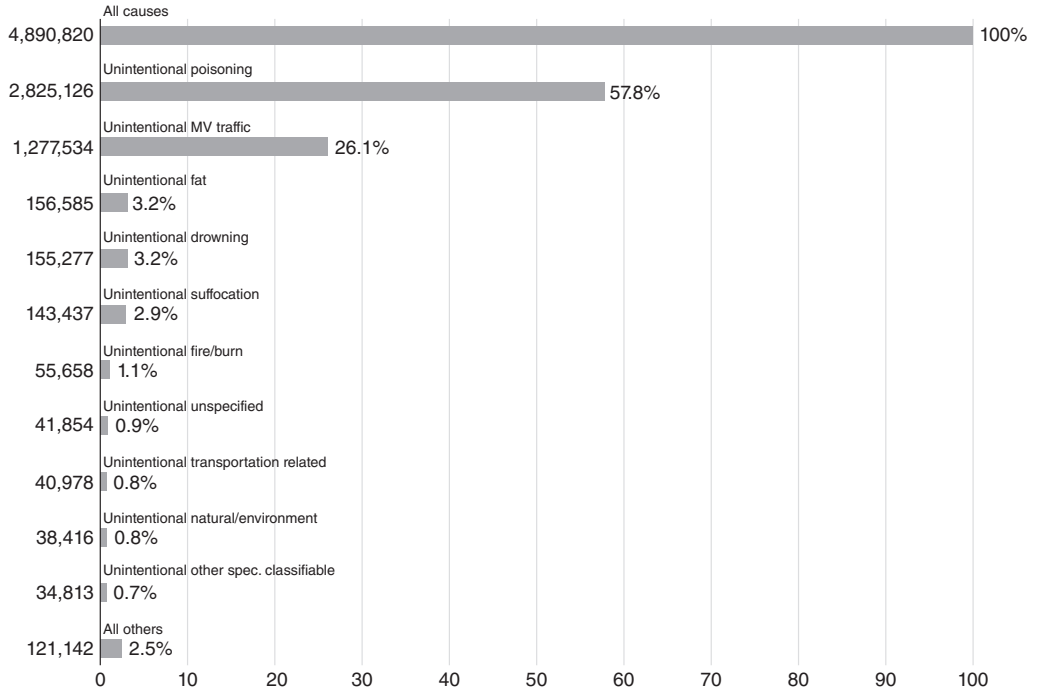


Fig. 4.10 Years of potential life lost (YPLL) before age 75 for 10 leading causes of unintentional injury death 2020, United States, both sexes, all races. (From Centers for Disease Control and Prevention. WISQARS Leading Causes of Death Visualization Tool. URL: https://wisqars.cdc.gov/data/lcd/home)

TABLE 4.4 Estimated YPLL Before Age 75 Years and Age-Adjusted Mortality Rates per 100,000 Persons, by Cause of Death, United States, 2020

Cause Category	Number of Deaths in 2020	YPLL
Heart Disease	696,962	1,443,729
Malignant Neoplasms	602,350	1,542,327
Covid-19	350,831	708,789
Unintentional Injury	200,955	3,403,047
Cerebrovascular	160,264	244,796
Chronic Low. Respiratory Disease	152,657	188,007
Diabetes Mellitus	102,188	330,236
Influenza & Pneumonia	53,544	143,931
Nephritis	52,547	103,656
Liver Disease	51,642	409,897
Suicide	45,979	942,431
Hypertension	41,907	71,390
Homicide	24,576	771,608
Benign Neoplasms	16,229	36,157

YPLL, Years of potential life lost.
Data from Centers for Disease Control and Prevention. WISQARS™ Web-based Injury Statistics Query and Reporting System. www.cdc.gov/injury/wisqars/

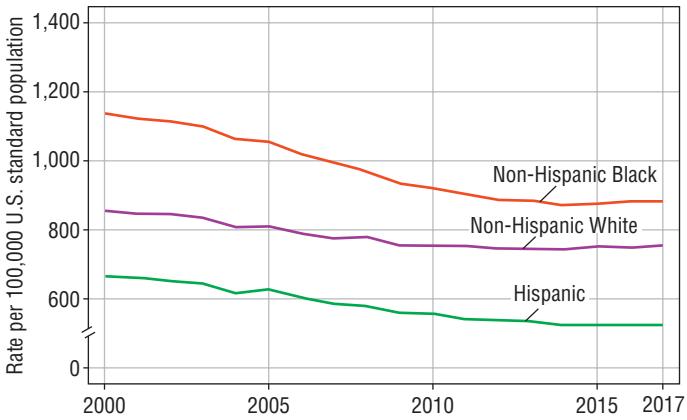
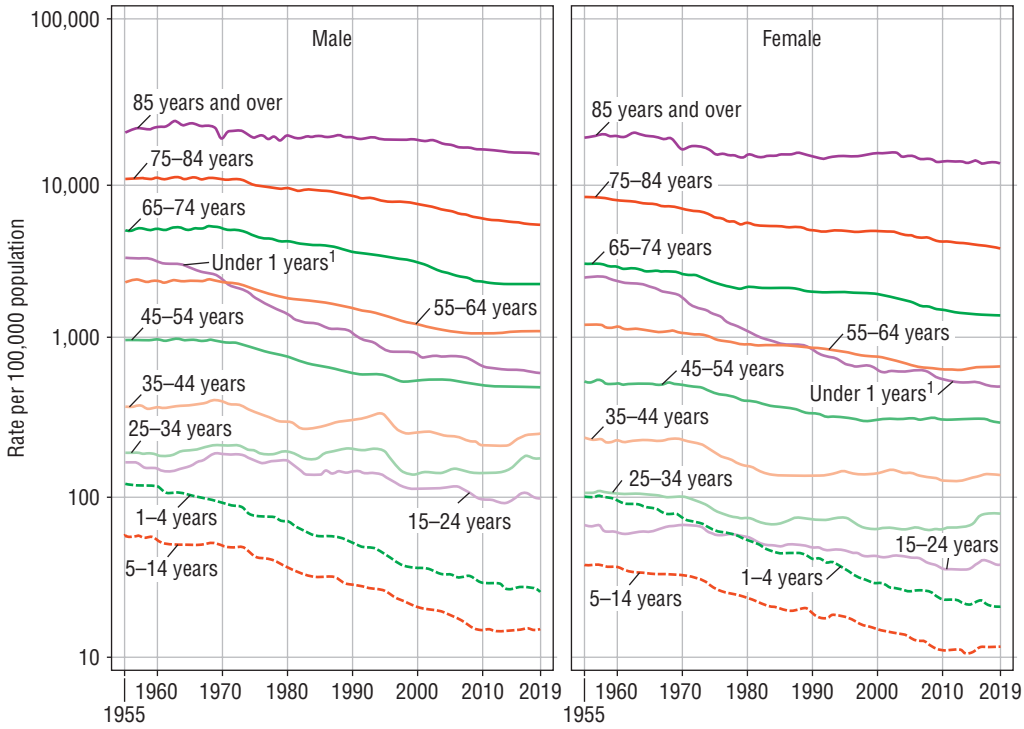


Fig. 4.11 Age-adjusted death rates, by race and Hispanic origin: United States, 2000–2017. (From Kochanek KD, Murphy SL, Xu J, Arias E. Deaths: Final data for 2017. National Vital Statistics Reports 68 (9). US Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics. National Vital Statistics Systems; June 24, 2019.)



[1]Rates are based on population estimates which differ from infant mortality rates (based on live births), see Figure 5 in this report for infant mortality rates and Technical Notes in this report for more discussion of the difference.

Fig. 4.12 Death rates, by age and sex: United States, 1955–2019. (From Xu J, Murphy SL, Kochanek, KD, Arias E. Deaths: Final data for 2019. National Vital Statistics Reports 70 (8). US Department of Health and Human Services. Center for Disease Control and Prevention. National Center for Health Statistics. National Vital Statistics Systems. July 26, 2021.)

is a steady decline in the death rate throughout the years, particularly in the age groups less than 14 years. This could be potentially attributed to the widespread coverage of childhood vaccinations. On the other hand, the decline was more modest in the age groups 45 to 64 years due to improvements in the early detection of cardiovascular diseases and cancer, and the evolving new effective treatments. If we look at the left panel for males, we see an increase in the mortality rate for age groups 25 to 44 years in the 1980s, followed by a sharp decline in the early 1990s. This can be explained by the then-emerging human immunodeficiency virus (HIV) disease, and followed by the newly introduced, highly active antiretroviral therapy in the mid-1990s, as well as lifestyle changes resulting from public health education.

A comparison of mortality and incidence is seen in Figs. 4.13 and 4.14. Fig. 4.13 shows breast cancer rates by year in selected European countries from 1975 to 2020. During this period, the age-standardized rates per 100,000 increased in all countries shown in the figure. This increase has been attributed to early detection and improved diagnostic modalities. As seen in Fig. 4.14, however, death rates from breast cancer in selected countries decreased markedly during the 1990s onward, perhaps as a result of earlier detection and increasingly prompt medical and surgical intervention.

Fig. 4.15 presents recent data on time trends in incidence and mortality from breast cancer in Black women and White women in the United States. Compare the time trends in incidence and mortality. What do these curves tell us about new cases of breast cancer over time and survival from breast cancer? Compare the experiences of Black women and White women with regard to both incidence and mortality. How can we describe the differences, and what could be some of the possible explanations?

A final example relates to reports in recent years that the incidence of thyroid cancer in the United States has been increasing. One of two possible explanations is likely. The first explanation is that these reports reflect a

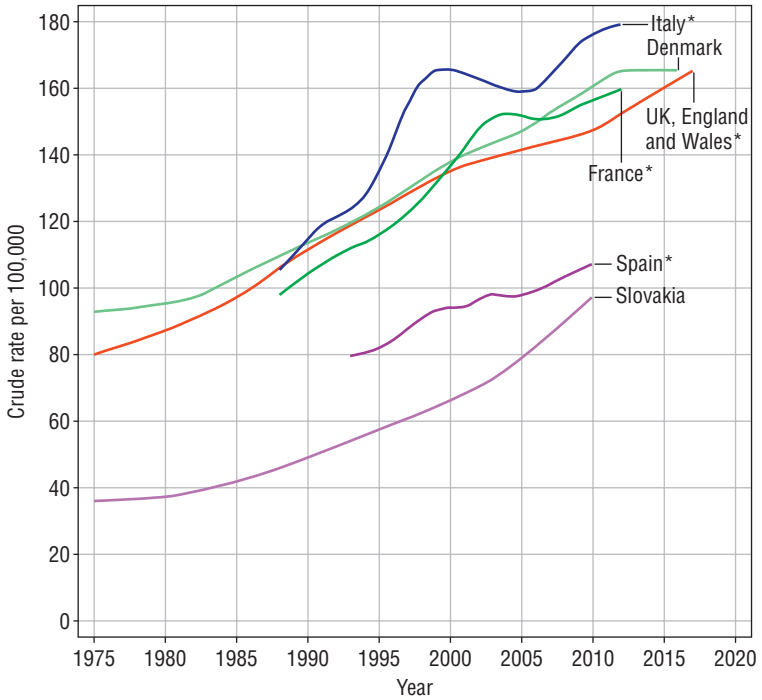
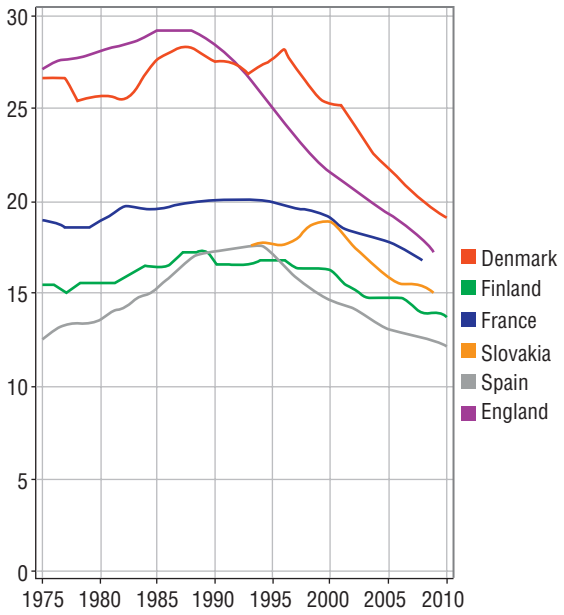


Fig. 4.13 Breast Cancer rate per 100,000, incidence. (From Cancer Over Time. International Agency for Research on Cancer. World Health Organization https://gco.iarc.fr/overtime/en/dataviz © 2022.)



WHO (www.who.int/healthinfo/en/)

Fig. 4.14 Breast Cancer per 10,000, mortality. (From Cancer Over Time. International Agency for Research on Cancer. World Health Organization https://gco.iarc.fr/overtime/en/dataviz © 2022.)

true increase in incidence that has resulted from increases in prevalence of risk factors for the disease. The second explanation is that the reported increased incidence is only an increase in *apparent* incidence. It does not reflect any true increase in new cases but rather an increase in the early detection and diagnosis of subclinical cases, because new diagnostic methods permit us to identify small and asymptomatic thyroid cancers that could not be detected previously.

In order to distinguish between these two possible explanations, Lim et al.[3] studied changes in incidence and mortality from thyroid cancer in the United States from 1974 to 2013.

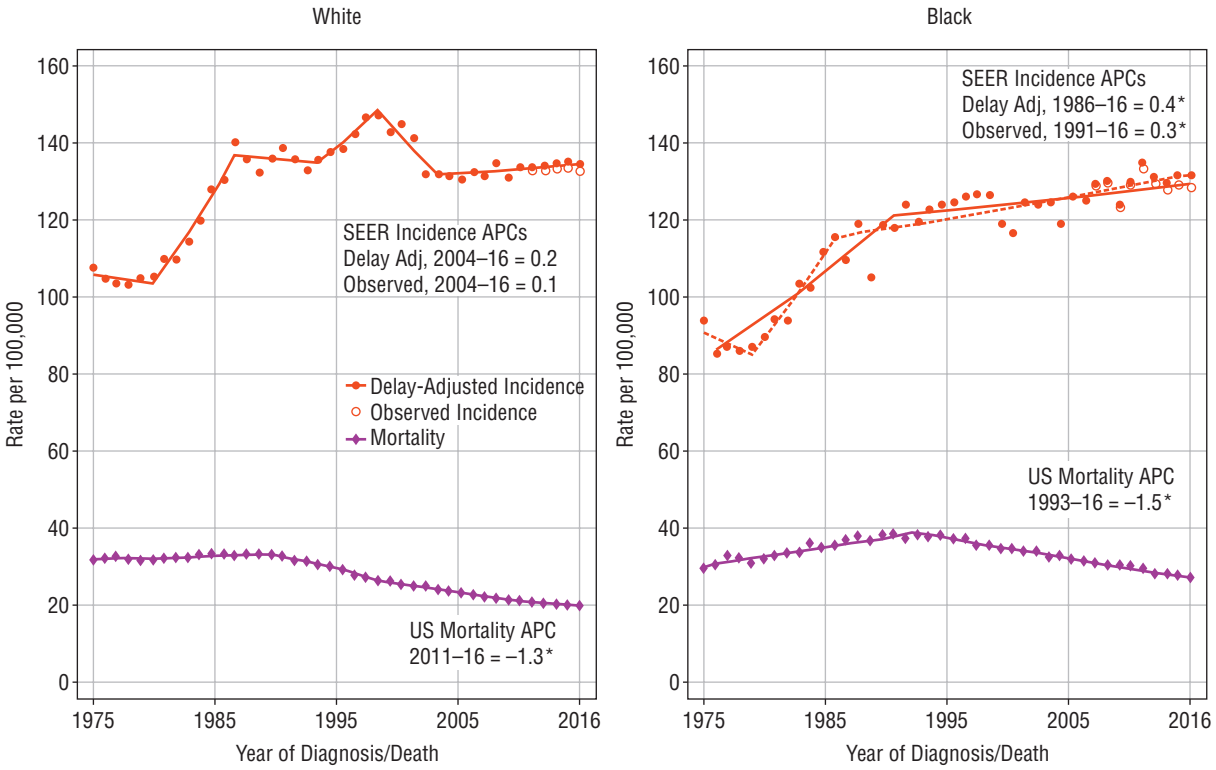
Thyroid cancer is a malignant growth of the cells that make up the thyroid glands. It involves several types according to the which cells are malignantly growing and this is identified after a small thyroid tissue (*biopsy*) is examined under the microscope to identify the histological type of the thyroid cancer. As seen in Fig. 4.16, at one extreme, papillary carcinoma has the best prognosis and at the opposite extreme, poorly differentiated types—medullary and anaplastic—are generally the most aggressive with poorest prognoses. The authors found that the increase in incidence of thyroid cancer was almost entirely due to an increase in the incidence of papillary cancer. Within the papillary cancers, most of the increase in this incidence was accounted for by the smallest-sized tumors. Thus, the authors found that 87% of the increase in thyroid cancer incidence over a 30-year period was accounted for by an increase in the smallest-sized papillary cancers, tumors that have the best prognosis. A number of earlier studies have shown a high prevalence of previously unrecognized, asymptomatic small papillary cancers at autopsy. If the increased incidence was due to the availability of more refined diagnostic methods, we would expect to see an increase in the incidence of small tumors, which is exactly what the authors found in their study.

Problems With Mortality Data

Most of our information about deaths comes from death certificates. A death certificate is shown in Fig. 4.17. By international agreement, deaths are coded according to the *underlying cause*. The underlying cause of death is defined as "*the disease or injury which initiated the train of morbid events leading directly or indirectly to death or the circumstances of the accident or violence which produced the fatal injury*."[4] Thus, the death certificate from which Fig. 4.18 is taken would be coded as a death from chronic ischemic heart disease, the underlying cause, which is always found on the lowest line used in part I of item 32 of the certificate. The underlying cause of death therefore "excludes information pertaining to the immediate cause of death, contributory causes, and those causes that intervene between the underlying and immediate causes of death."[5] As pointed out by Savage and coworkers,[6] the total contribution of a given cause of death may not be reflected in the mortality data as generally reported; this may apply to a greater extent in some diseases than in others.

Countries and regions vary greatly in the quality of the data provided on their death certificates. Studies of validity of death certificates compared with hospital and autopsy records generally find higher validity for certain diseases, such as cancers, than for others.

Deaths are coded according to the International Classification of Diseases (ICD), now in its 11th revision (2022). Because coding categories and regulations change from one revision to another, any study of time trends in mortality that spans more than one revision



[a]Rates are age-adjusted to the 2000 US Std population (19 age groups - Census P25-1103).
Regression lines and APCs are calculated using the Jointpoint Regression Program Version 4.7. February 2019, National Cancer Institute.
The APC is the Annual Percent Change for the regression line segments. The APC shown on the graph is for the most recent trend.
*The APC is significantly different from zero ($p < 0.05$).

Fig. 4.15 SEER observed incidence, SEER delay adjusted incidence and US death rates[a] cancer of the female breast, by race. *SEER,* Surveillance, Epidemiology, and End Results. (From National Cancer Institute: Surveillance, Epidemiology, and End Results Program. URL: https://seer.cancer.gov/archive/csr/1975_2016/browse_csr.php?sectionSEL=4&pageSEL=sect_04_zfig.01)

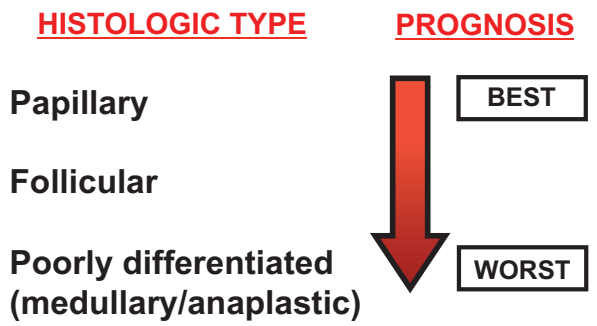
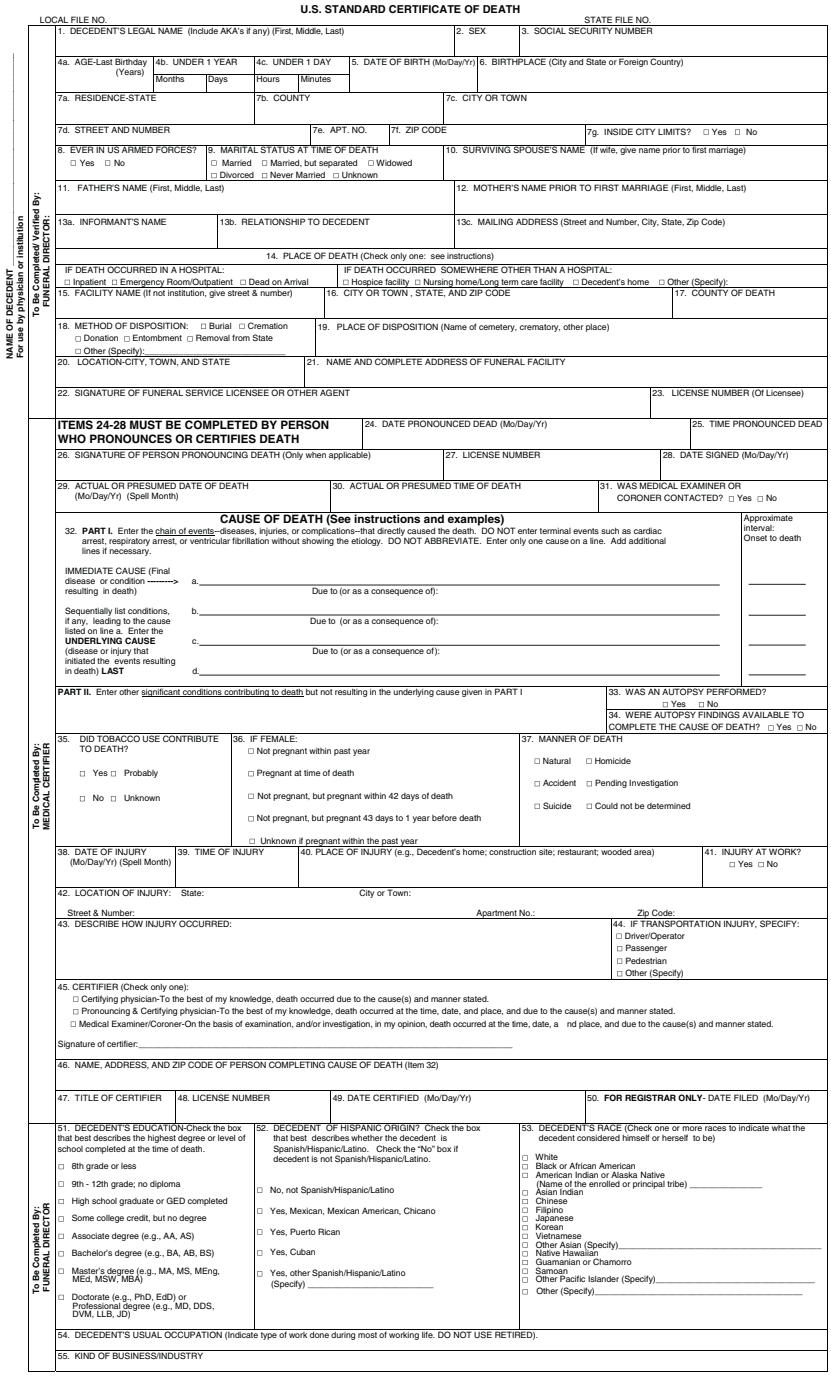


Fig. 4.16 Histologic types of thyroid cancer and their prognoses.

must examine the possibility that observed changes could be due entirely or in part to changes in the ICD. In 1949, mortality rates from diabetes showed a dramatic decline in both men and women (Fig. 4.19). However, any euphoria that these data might have caused was short-lived; analysis of this drop indicated that it occurred at a time of change from the seventh revision to the eighth revision of the ICD. Prior to 1949, the policy was that any death certificate that included mention of diabetes anywhere be coded as a death from diabetes. After 1949, only death certificates on which the underlying cause of death was listed as diabetes were



U.S. STANDARD CERTIFICATE OF DEATH

LOCAL FILE NO. STATE FILE NO.

NAME OF DECEDENT
For use by physician or institution
To Be Completed/ Verified By: FUNERAL DIRECTOR:

1. DECEDENT'S LEGAL NAME (Include AKA's if any) (First, Middle, Last) | 2. SEX | 3. SOCIAL SECURITY NUMBER

4a. AGE-Last Birthday (Years) | 4b. UNDER 1 YEAR Months Days | 4c. UNDER 1 DAY Hours Minutes | 5. DATE OF BIRTH (Mo/Day/Yr) | 6. BIRTHPLACE (City and State or Foreign Country)

7a. RESIDENCE-STATE | 7b. COUNTY | 7c. CITY OR TOWN

7d. STREET AND NUMBER | 7e. APT. NO. | 7f. ZIP CODE | 7g. INSIDE CITY LIMITS? ☐ Yes ☐ No

8. EVER IN US ARMED FORCES? ☐ Yes ☐ No | 9. MARITAL STATUS AT TIME OF DEATH ☐ Married ☐ Married, but separated ☐ Widowed ☐ Divorced ☐ Never Married ☐ Unknown | 10. SURVIVING SPOUSE'S NAME (If wife, give name prior to first marriage)

11. FATHER'S NAME (First, Middle, Last) | 12. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last)

13a. INFORMANT'S NAME | 13b. RELATIONSHIP TO DECEDENT | 13c. MAILING ADDRESS (Street and Number, City, State, Zip Code)

14. PLACE OF DEATH (Check only one: see instructions)

IF DEATH OCCURRED IN A HOSPITAL: ☐ Inpatient ☐ Emergency Room/Outpatient ☐ Dead on Arrival
IF DEATH OCCURRED SOMEWHERE OTHER THAN A HOSPITAL: ☐ Hospice facility ☐ Nursing home/Long term care facility ☐ Decedent's home ☐ Other (Specify):

15. FACILITY NAME (If not institution, give street & number) | 16. CITY OR TOWN , STATE, AND ZIP CODE | 17. COUNTY OF DEATH

18. METHOD OF DISPOSITION: ☐ Burial ☐ Cremation ☐ Donation ☐ Entombment ☐ Removal from State ☐ Other (Specify):_____ | 19. PLACE OF DISPOSITION (Name of cemetery, crematory, other place)

20. LOCATION-CITY, TOWN, AND STATE | 21. NAME AND COMPLETE ADDRESS OF FUNERAL FACILITY

22. SIGNATURE OF FUNERAL SERVICE LICENSEE OR OTHER AGENT | 23. LICENSE NUMBER (Of Licensee)

To Be Completed By: MEDICAL CERTIFIER

ITEMS 24-28 MUST BE COMPLETED BY PERSON WHO PRONOUNCES OR CERTIFIES DEATH

24. DATE PRONOUNCED DEAD (Mo/Day/Yr) | 25. TIME PRONOUNCED DEAD

26. SIGNATURE OF PERSON PRONOUNCING DEATH (Only when applicable) | 27. LICENSE NUMBER | 28. DATE SIGNED (Mo/Day/Yr)

29. ACTUAL OR PRESUMED DATE OF DEATH (Mo/Day/Yr) (Spell Month) | 30. ACTUAL OR PRESUMED TIME OF DEATH | 31. WAS MEDICAL EXAMINER OR CORONER CONTACTED? ☐ Yes ☐ No

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

Approximate interval: Onset to death

IMMEDIATE CAUSE (Final disease or condition ---------> resulting in death) a._____ Due to (or as a consequence of):

Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST** b._____ Due to (or as a consequence of): c._____ Due to (or as a consequence of): d._____

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I

33. WAS AN AUTOPSY PERFORMED? ☐ Yes ☐ No

34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ☐ Yes ☐ No

35. DID TOBACCO USE CONTRIBUTE TO DEATH? ☐ Yes ☐ Probably ☐ No ☐ Unknown

36. IF FEMALE: ☐ Not pregnant within past year ☐ Pregnant at time of death ☐ Not pregnant, but pregnant within 42 days of death ☐ Not pregnant, but pregnant 43 days to 1 year before death ☐ Unknown if pregnant within the past year

37. MANNER OF DEATH ☐ Natural ☐ Homicide ☐ Accident ☐ Pending Investigation ☐ Suicide ☐ Could not be determined

38. DATE OF INJURY (Mo/Day/Yr) (Spell Month) | 39. TIME OF INJURY | 40. PLACE OF INJURY (e.g., Decedent's home; construction site; restaurant; wooded area) | 41. INJURY AT WORK? ☐ Yes ☐ No

42. LOCATION OF INJURY: State: City or Town: Street & Number: Apartment No.: Zip Code:

43. DESCRIBE HOW INJURY OCCURRED: | 44. IF TRANSPORTATION INJURY, SPECIFY: ☐ Driver/Operator ☐ Passenger ☐ Pedestrian ☐ Other (Specify)

45. CERTIFIER (Check only one):
☐ Certifying physician-To the best of my knowledge, death occurred due to the cause(s) and manner stated.
☐ Pronouncing & Certifying physician-To the best of my knowledge, death occurred at the time, date, and place, and due to the cause(s) and manner stated.
☐ Medical Examiner/Coroner-On the basis of examination, and/or investigation, in my opinion, death occurred at the time, date, a nd place, and due to the cause(s) and manner stated.
Signature of certifier:_____

46. NAME, ADDRESS, AND ZIP CODE OF PERSON COMPLETING CAUSE OF DEATH (Item 32)

47. TITLE OF CERTIFIER | 48. LICENSE NUMBER | 49. DATE CERTIFIED (Mo/Day/Yr) | 50. **FOR REGISTRAR ONLY**- DATE FILED (Mo/Day/Yr)

To Be Completed By: FUNERAL DIRECTOR

51. DECEDENT'S EDUCATION-Check the box that best describes the highest degree or level of school completed at the time of death.
☐ 8th grade or less
☐ 9th - 12th grade; no diploma
☐ High school graduate or GED completed
☐ Some college credit, but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

52. DECEDENT OF HISPANIC ORIGIN? Check the box that best describes whether the decedent is Spanish/Hispanic/Latino. Check the "No" box if decedent is not Spanish/Hispanic/Latino.
☐ No, not Spanish/Hispanic/Latino
☐ Yes, Mexican, Mexican American, Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino (Specify) _____

53. DECEDENT'S RACE (Check one or more races to indicate what the decedent considered himself or herself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native (Name of the enrolled or principal tribe) _____
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)_____
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)_____
☐ Other (Specify)_____

54. DECEDENT'S USUAL OCCUPATION (Indicate type of work done during most of working life. DO NOT USE RETIRED).

55. KIND OF BUSINESS/INDUSTRY

REV. 11/2003

Fig. 4.17 US standard certificate of death. (From Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/data/dvs/death11-03final-acc.pdf. Accessed June 7, 2017.)

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

Approximate interval: Onset to death

IMMEDIATE CAUSE (Final disease or condition ----------> resulting in death)	a.	Rupture of myocardium	Mins
		Due to (or as a consequence of):	
Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST**	b.	Acute myocardial infarction	6 days
		Due to (or as a consequence of):	
	c.	Chronic ischemic heart disease	5 years
		Due to (or as a consequence of):	
	d.		

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I

Diabetes, Chronic obstructive pulmonary disease, smoking

33. WAS AN AUTOPSY PERFORMED? ☐ Yes ☐ No

34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ☐ Yes ☐ No

Fig. 4.18 Example of a completed cause-of-death section on a death certificate, including immediate and underlying causes and other significant conditions.

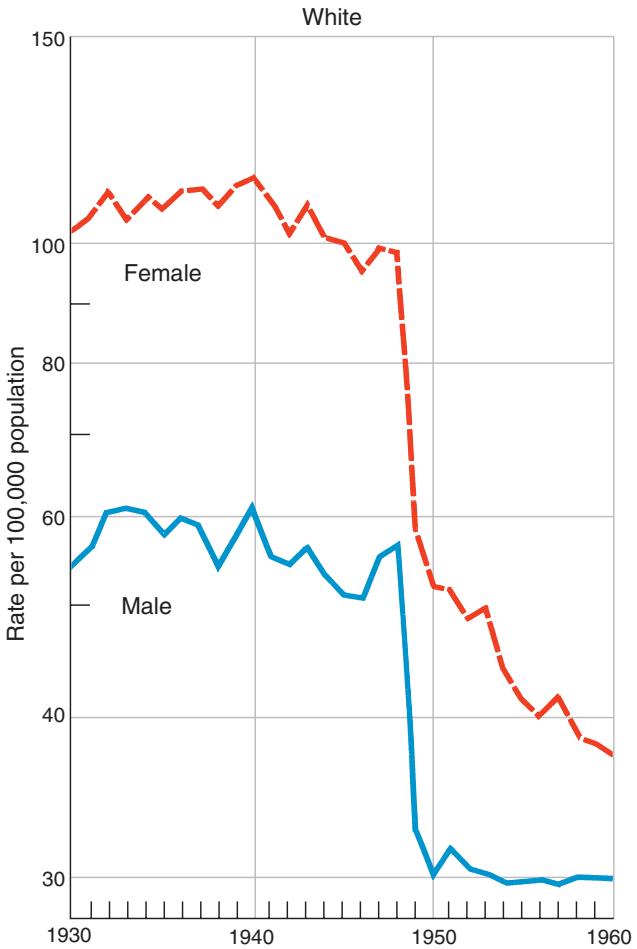


Fig. 4.19 Drop in death rates for diabetes among 55- to 64-year-old men and women, United States, 1930–60, due to changes in International Classification of Diseases coding. (From *US Public Health Service Publication No. 1000, Series 3, No. 1.* Washington, DC: US Government Printing Office; 1964.)

coded as a death from diabetes. Hence, the decline seen in Fig. 4.19 was an artifact of the change in coding. Whenever we see a time trend of an increase or a decrease in mortality, the first question we must ask is, "Is it real?" Specifically, when we look at trends in mortality over time, we must ask whether any changes took place in how death certificates were coded during the period being examined and whether these changes could have contributed to changes observed in mortality during the same period.

Changes in the definition of disease can also have a significant effect on the number of cases of the disease that are reported or that are subsequently classified as meeting the diagnostic criteria for the disease. In early 1993, a new definition of acquired immunodeficiency syndrome (AIDS) was introduced; as shown in Fig. 4.20, this change resulted in a rapid rise in the number of reported cases. With the new definition, even after the initial peak, the number of reported cases remained higher than it had been for several years.

In discussing morbidity in Chapter 3, we said that everyone in the group represented by the denominator must be at risk to enter the group represented by the numerator, and we looked at cervical cancer incidence rates as an example. The same principle regarding numerator and denominator applies to mortality rates. Fig. 4.21 shows a similar set of observations for mortality rates from cervical cancers. Once again, correcting for hysterectomy reduces the number of women in the denominator and thus increases the mortality rate. In a lighter vein, Box 4.2 lists some causes of death that were listed on death certificates early in the 20th century.

COMPARING MORTALITY IN DIFFERENT POPULATIONS

An important use of mortality data is to compare two or more populations, or one population in different time periods. Such populations may differ with regard to many characteristics that affect mortality, of which the age distribution is the most important. In fact, age is the single most important predictor of mortality. Therefore, methods have been developed for comparing mortality in such populations while effectively holding constant characteristics such as age.

Table 4.5 shows data that exemplify this problem. Deaths and death rates by race for residents of the State of Maryland between 1940 and 2019 are given. The data may seem surprising because we would expect rates to have been higher for Black residents, given the problems

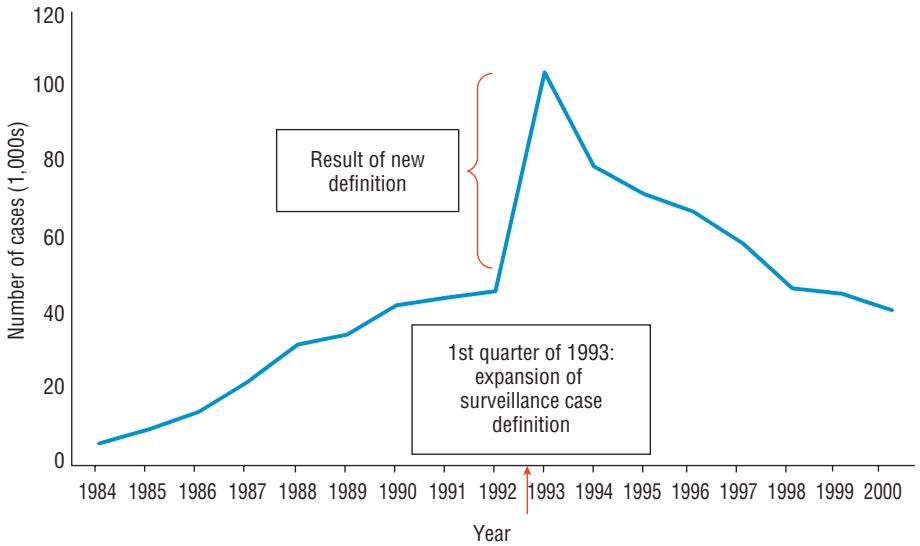


Fig. 4.20 Acquired immunodeficiency syndrome cases by quarter year of report, United States, 1984–2000. (From Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 2000. *MMWR*. 2000;49:86; and Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 1993. *MMWR*. 1993;45:68.)

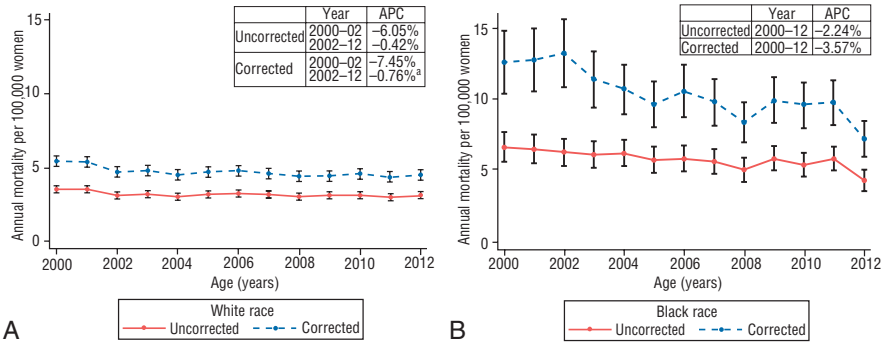


Fig. 4.21 Trends in age-standardized cervical cancer mortality rates, uncorrected and corrected for the prevalence of hysterectomy, from 2000–2012 for (A) White and (B) Black women. *APC,* Annual percentage change. (From Beavis AL, Gravitt PE, Rositch AF. Hysterectomy-corrected cervical cancer mortality rates reveal a larger racial disparity in the United States. *Cancer.* 2017;123:1044–1050.)

BOX 4.2 Some Causes of Death That Were Reported on Death Certificates in the Early 1900s

"Died suddenly without the aid of a physician"
"A mother died in infancy"
"Deceased had never been fatally sick"
"Died suddenly, nothing serious"
"Went to bed feeling well, but woke up dead"

TABLE 4.5 Deaths and Death Rates by Race, Maryland: Selected Years, 1940–2019

	NUMBER OF DEATHS			DEATH RATES[a]		
Year	All Races	White	Black	All Races	White	Black
1940	21,883	16,943	–	11.9	11.1	–
1950	22,450	17,811	–	9.0	8.7	–
1960	27,992	22,447	–	9.0	8.7	–
1970	32,790	26,290	6,448	8.3	8.2	9.2
1980	34,025	26,828	7,056	8.1	8.5	7.3
1990	38,384	28,897	9,172	8.0	8.4	7.6
2000	43,602	31,603	11,349	8.2	8.9	7.5
2010	43,255	30,488	11,761	7.5	8.5	6.6
2011	43,650	30,947	11,645	7.5	8.5	6.5
2012	44,110	31,173	11,884	7.5	8.6	6.5
2013	45,444	31,990	12,358	7.7	8.8	6.7
2014	45,688	31,979	12,559	7.6	8.7	6.7
2015	47,235	32,560	13,434	7.9	8.9	7.1
2016	48,884	33,176	14,113	8.1	9.1	7.4
2017	50,009	33,839	14,507	8.3	9.3	7.5
2018	50,668	33,837	15,034	8.4	9.3	7.8
2019	50,873	33,654	15,290	8.4	9.3	7.9

– Data arc not available
[a] Per 1,000 population
(From Maryland Vital Statistics Annual Report 2019. Maryland Department of Health Vital Statistics Administration. 135. https://health.maryland.gov/vsa/Documents/Reports%20and%20Data/Annual%20Reports/2019Annual.pdf.)

TABLE 4.6 Death Rates by Age and Race, State of Maryland, 2015

	DEATH RATES BY AGE PER 1,000 POPULATION[A]											
Race	All Ages	<1 Year	1–4 Years	5–14 Years	15–24 Years	25–34 Years	35–44 Years	45–54 Years	55–64 Years	65–74 Years	75–84 Years	>85 Years
White	9.95	4.06	0.21	0.11	0.64	1.29	1.73	3.62	7.68	16.45	45.39	138.7
Black	7.35	11.25	0.43	0.18	1.14	1.74	2.23	5.09	11.14	21.55	49.49	124.45

[a]Age-adjusted to the 2000 US population.

From Maryland Vital Statistics Annual Report; 2015. https://health.maryland.gov/vsa/Documents/15annual.pdf. Accessed June 8, 2017. Certain data were provided by the Vital Statistics Administration, Maryland Department of Health, Baltimore, Maryland. The Department disclaims responsibility for any analyses, interpretations, or conclusions.

associated with poorer living conditions and less access to medical care. When we look at Table 4.6, we see the data from Table 4.5 on the left, but now we have added data for each age-specific stratum (layer) of the population. Interestingly, although in each age-specific group, mortality is higher in Black than in White residents, the overall mortality (also called *crude* or *unadjusted mortality*) is higher in White than in Black residents. Why is this so? This is a reflection of the fact that in both White and Black residents, mortality increases markedly in the oldest age groups; older age is the major contributor to mortality. However, the White population in this example is older than the Black population, and in 2015, there were few Black people in the oldest age groups. Thus, in White people, the overall mortality is heavily weighted by high rates in the oldest age groups. The overall (or crude) mortality rate in the White population is increased by the greater number of deaths in the large subgroup of older White people, but the overall mortality rate in the Black population has not increased as much because there are so many fewer deaths in the small number of Black people in the older age groups. Clearly, the crude mortality reflects both differences in the force of mortality and differences in the age composition of the population. Let us look at two approaches for dealing with this problem: direct and indirect age adjustment.

Direct Age Adjustment

Tables 4.7 through 4.9 show a hypothetical example of direct age adjustment. Table 4.7 shows mortality in a population in two different time periods. The mortality rate is considerably higher in the later period. These data are supplemented with age-specific data in Table 4.8.

TABLE 4.7 Hypothetical Example of Direct Age Adjustment: I. Comparison of Total Death Rates in a Population at Two Different Times

EARLY PERIOD			LATER PERIOD		
Population	No. of Deaths	Death Rate per 100,000	Population	No. of Deaths	Death Rate per 100,000
900,000	862	96	900,000	1,130	126

TABLE 4.8 Hypothetical Example of Direct Age Adjustment: II. Comparison of Age-Specific Death Rates in Two Different Time Periods

	EARLY PERIOD			LATER PERIOD		
Age Group (years)	Population	No. of Deaths	Death Rates per 100,000	Population	No. of Deaths	Death Rates per 100,000
All ages	900,000	862	96	900,000	1,130	126
30–49	500,000	60	12	300,000	30	10
50–69	300,000	396	132	400,000	400	100
70+	100,000	406	406	200,000	700	350

TABLE 4.9 Hypothetical Example of Direct Age Adjustment: III. Carrying Out an Age Adjustment Using the Total of the Two Populations as the Standard

Age Group (years)	Standard Population	"Early" Age-Specific Mortality Rates per 100,000	Expected No. of Deaths Using "Early" Rates	"Later" Age-Specific Mortality Rates per 100,000	Expected No. of Deaths Using "Later" Rates
All ages	1,800,000				
30–49	800,000	12	96	10	80
50–69	700,000	132	924	100	700
70+	300,000	406	1,218	350	1,050
Total no. of deaths expected in the standard population:		2,238		1,830	
Age-adjusted rates:		"Early" $= \frac{2{,}238}{1{,}800{,}000} = 124.3$		"Later" $= \frac{1{,}830}{1{,}800{,}000} = 101.7$	

Here, we see three age groups, and age-specific mortality for the later period is lower in each group. How, then, is it possible to account for the higher overall mortality in the later period in this example?

The answer lies in the changing age structure of the population. Mortality is highest in the oldest age groups, and during the later period, the size of the oldest group doubled from 100,000 to 200,000, whereas the number of young people declined substantially, from 500,000 to 300,000. We would like to eliminate this age difference and, in effect, ask: if the age composition of the populations were the same, would there be any differences in mortality between the early period and the later period?

In *direct age adjustment*, a standard population is used in order to eliminate the effects of any differences in age between two or more populations being compared (see Table 4.9). A hypothetical "standard" population is created to which we apply both the age-specific mortality rates from the early period and the age-specific mortality rates from the later period. By applying mortality rates from both periods to a single standard population, we eliminate any possibility that observed differences could be a result of age differences in the population. (In this example, we have created a standard by combining the populations from the early and the later periods, but any population could have been used.)

By applying each age-specific mortality rate to the population in each age group of the standard population, we derive the expected number of deaths that would have occurred had those rates been applied. We can then calculate the total number of deaths expected in the standard population had the age-specific rates of the early period applied and the total number of deaths expected in the standard population had the age-specific rates of the later period applied. Dividing each of these two total expected numbers of deaths by the total standard population, we can calculate an expected mortality rate in the standard population if it had had the mortality experience of the early period and the expected mortality rate for the standard population if it had had the mortality experience for the later period. These are called *age-adjusted rates*, and they appropriately reflect the decline seen in the age-specific rates. Differences in age-composition of the population are no longer a factor.

In this example the rates have been adjusted for age, but adjustment can also be carried out for any characteristic of interest such as sex, socioeconomic status, or race, and techniques are also available to adjust for multiple variables simultaneously.

Although age-adjusted rates can be very useful in making comparisons, the first step in examining and analyzing comparative mortality data should always be to carefully examine the *age-specific* rates for any interesting differences or changes. These differences may be hidden by the *age-adjusted* rates and may be lost if we proceed immediately to age adjustment without first examining the age-specific rates.

Age-adjusted rates are *hypothetical* because they involve applying actual age-specific rates to a hypothetical standard population. They do not reflect the true mortality risk of a "real" population because the numerical value of an age-adjusted death rate depends on the standard population used. Selection of such a population is somewhat arbitrary because there is no "correct" standard population, but it is generally accepted that the "standard" should not be markedly different from the populations that are being compared with regard to age or whatever the variable is for which the adjustment is being made. In the United States, for more than 50 years, the 1940 US population was regularly used as the standard population for age adjustment for most purposes, but in recent years, this population was increasingly considered outdated and incompatible with the older age structure of the US population. Beginning with 1999 mortality statistics, the US population in the year 2000 replaced the 1940 population as the standard population for adjustment.

The change in standard population to the year 2000 US population has had some significant effects, as illustrated with a comparison of cause-specific mortality rates using data through 1995.[7] These include increases in age-adjusted mortality rates that were observed for causes in which risk increases significantly with age. For example, age-adjusted death from cerebrovascular diseases (stroke) is 26.7 deaths per 100,000 using the 1940 standard, but it is 63.9 per 100,000 using the 2000 standard. Cancer mortality increased using the 2000 population standard compared to when an earlier population was used as a standard because more people are surviving into older ages, when many of the leading types of cancer are more common. Rates for heart disease, chronic obstructive lung disease, diabetes, kidney disease, and Alzheimer disease were similarly affected because age-specific death rates for all these conditions are higher in older age groups.

Age-adjusted rates of cancer are higher in Black people compared to White people in the United States, but the differential between Black and White people is less with the 2000 population standard than with the earlier standard population. Thus, the change to the year 2000 US population as the standard complicates comparisons of age-adjusted rates before and after 1999, because many of the rates before 1999 were calculated using the 1940 standard population. However, the rates from 1999 forward are being calculated using the year 2000 population as the new standard.

In summary, the goal of direct adjustment is to compare rates in at least two different populations when we wish to eliminate the possible effect of a given factor, such as age, on the rates we are comparing. It is important to keep in mind that adjusted rates are not "real" rates in the populations being compared, because they depend on the choice of the standard population used in carrying out the adjustment. Nevertheless, direct adjustment is a very useful tool for making such comparisons and in fact, comparison of rates in different populations almost always utilizes direct adjustment, such as adjustment for age. Note that adjustment is based on replacing each population with a common set of weights (the standard population) in order to estimate weighted averages—that is, the adjusted rates.

Indirect Age Adjustment (Standardized Mortality Ratios)

Indirect age adjustment is often used when numbers of deaths for each age-specific stratum are not available. It is also used to study mortality in an occupationally exposed population: Do people who work in a certain industry, such as mining or construction, have a higher mortality than people of the same age in the general population? Is an additional risk associated with that occupation?

To answer the question of whether a population of workers has a higher mortality than we would expect in a similar population that is not engaged in the occupation being observed, the age-specific rates for a known population, such as all men of the same age, are applied to each age group in the population of interest. This will yield the number of deaths expected in each age group in the population of interest, if this population had had the mortality experience of the known population. Thus, for each age group, the number of deaths *expected* is calculated, and these numbers are totaled. The numbers of deaths that were actually *observed* in that population are also calculated and totaled. The ratio of the total number of deaths actually observed to the total number of deaths expected, if the population of interest had the mortality experience of the known population, is then calculated. This ratio is called the *standardized mortality ratio* (*SMR*).

TABLE 4.10 Hypothetical Computation of a Standardized Mortality Ratio (SMR) for Disease X for White Workers Ages 20–59 Years, 2023

	Estimated Population for White Workers	Death Rate (per 100,000) for Disease X in Males in the General Population	Expected Deaths From Disease X in White Workers If They Had Same Risk as General Population	Observed Deaths From Disease in White Workers
Age (years)	1	2	3 = 1 × 2	4
20–24	62,253	8.9	5.5	5
25–29	72,732	12.7	9.3	15
30–34	68,500	18.1	12.4	17
35–44	136,525	30.6	41.7	93
45–54	90,304	53.4	48.2	169
55–59	30,149	71.8	21.7	107
Totals	460,463		138.8	406

$$\text{SMR (for 20- to 59-year-olds)} = \frac{406}{138.8} \times 100 = 292$$

The SMR is defined as follows:

$$\text{SMR} = \frac{\text{Observed no. of deaths per year}}{\text{Expected no. of deaths per year}}$$

Let us look at the example in Table 4.10. In a hypothetical population of 460,463 White male workers, 406 deaths from disease X occurred in 2023. The question we are interested in is whether this mortality experience from disease X is greater than, less than, or about the same as that expected in White men of the same ages in the general population (most of whom are not included in this classification of workers). To help address this question, we may calculate the expected number of deaths for White workers in each age group by applying the known age-specific mortality rate from the general population to the number of workers in each age group. By doing so, we ask, "How many deaths would we expect in these White workers if they had the same mortality experience as White men in the same age group in the general population?" These data are listed in column 3. Column 4 shows the number of deaths observed in the workers.

The SMR is calculated by totaling the observed number of deaths (406) and dividing it by the expected number of deaths (138.8), which yields a result of 2.92. Multiplication by 100 is often done to yield results without decimals. If this were done in this case, the SMR would be 292. An SMR of 100 indicates that the observed number of deaths is the same as the expected number of deaths. An SMR greater than 100 indicates that the observed number of deaths exceeds the expected number, and an SMR less than 100 indicates that the observed number of deaths is less than the expected number.

The Cohort Effect

Table 4.11 shows age-specific obesity prevalence (%) from 1971 to 2006 in the United States using data from the National Center for Health Statistics. (For this discussion, we will ignore the data for age groups 2 to 19 years, since childhood obesity is a somewhat different phenomenon.) If, for example, we then read *down* the column in the table (the data for a given National Health and Nutrition Examination Survey [NHANES] cycle) for 1971–1975, it appears that obesity prevalence peaks in the age group 55 to 59 years and then declines with advancing age. Such a view of the data, by year, is called a *cross-sectional view*.

Actually, however, the picture of obesity prevalence is somewhat different (Table 4.12). A person who was 20 to 24 years of age in 1971 was 25 to 29 years of age in 1976. In other words, persons who were born in a certain year are moving through time together. We can now examine the obesity prevalence over time of the same cohort (i.e., a group of people who share the

TABLE 4.11 Age-Period Contingency Table for Obesity Prevalence by Age (Rows) and Period (Columns) in the United States, 1971–2006 (N = 91,755)

	NHANES I	NHANES II	NHANES III, Phase 1	NHANES III, Phase 2	NHANES 99–00	NHANES 01–02	NHANES 03–04	NHANES 05–06
	1971–75	1976–80	1988–91	1991–94	1999–2000	2001–02	2003–04	2005–06
2–4	3.1	3.25	3.48	4.34	7.02	6.29	8.72	8.54
5–9	5.48	7.17	8.75	13.12	17.45	16.92	20.22	16.25
10–14	6.88	7.9	8.93	13.57	18.97	18.72	22.85	21.81
15–19	6.64	5.5	8.31	13.55	18.03	17.8	19.94	18.43
20–24	6.08	7.14	9.87	13.81	20.59	26.67	26.59	24.98
25–29	10.34	10.49	11.97	18.74	27.69	26.55	26.47	35.9
30–34	13.64	13.49	18.02	20.07	31.64	24.82	30.19	36.6
35–39	14.34	14.73	17.24	23.3	29.08	30.19	36.54	33.3
40–44	16.76	15.82	19.27	24.63	32.68	32.85	39.68	42.69
45–49	15.26	18.05	18.85	30.75	31.93	35.83	35.79	38.5
50–54	17.18	17.46	22.37	35.42	40.55	31.69	39.32	38.73
55–59	19.5	19.62	26.55	32.46	35.7	38	38.62	46.9
60–64	18.68	17.57	20.82	30.67	41.37	44.28	34.49	42.67
65–69	16.83	18.51	21.26	27.79	41.23	35.43	38	40.64
70–74	17.15	16.31	18.68	25.03	29.34	34.87	32.48	31.45

NHANES, National Health and Nutrition Examination Survey.
From Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971–2006. *Soc Sci Med.* 2010;70(7):1100–1108.

same experience), born in the same 5-year period. Looking at persons who were 20 to 24 years of age in the 1971–1975 cycle and following them over time, as indicated by the bold black boxes in the table, it is apparent that obesity prevalence for this *cohort* has been increasing throughout the years and did not decline later on, as we have seen in the cross-sectional view of the data. When we examine changes in prevalence over time, we should always ask whether any apparent changes that are observed could be the result of such a cohort effect.

Interpreting Observed Changes in Mortality

If we find a difference in mortality over time or between populations—either an increase or a decrease—it may be an artifact or it may be real. If it is an artifact, the artifact could result from problems with either the numerator or the denominator (Table 4.13). However, if we conclude that the change is real, what could be the possible explanation? Some possibilities are seen in Box 4.3.

OTHER MEASURES OF THE IMPACT OF DISEASE

Quality of Life

Most diseases have a major impact on the afflicted individuals above and beyond mortality. Diseases that may not be lethal may be associated with considerable physical and emotional suffering, resulting from disability associated with the illness. It is therefore important to consider the total impact of a disease as measured by its effect on a person's quality of life, even though such measures are not, in fact, measures of disease occurrence. For example, it is possible to examine the extent to which patients with arthritis are compromised by the illness in carrying out activities of daily living. In addition, regulatory authorities such as the US Food and Drug Administration review and evaluate quality of life as part of the patient-reported outcome measures that can be used to support claims of improving patients' quality of life in approved drug labels.

TABLE 4.12 Age-Period Contingency Table for Obesity Prevalence by Age (Rows) and Period (Columns) in the United States, 1971–2006 (N = 91,755)

Age (years)	NHANES I 1971–75	NHANES II 1976–80	NHANES III, Phase 1 1988–91	NHANES III, Phase II 1991–94	NHANES 99–00 1999–2000	NHANES 01–02 2001–02	NHANES 03–04 2003–04	NHANES 05–06 2005–06
2–4	3.1	3.25	3.48	4.34	7.02	6.29	8.72	8.54
5–9	5.48	7.17	8.75	13.12	17.45	16.92	20.22	16.25
10–14	6.88	7.9	8.93	13.57	18.97	18.72	22.85	21.81
15–19	6.64	5.5	8.31	13.55	18.03	17.8	19.94	18.43
20–24	**6.08**	7.14	9.87	13.81	20.59	26.67	26.59	24.98
25–29	10.34	**10.49**	11.97	18.74	27.69	26.55	26.47	35.9
30–34	13.64	13.49	**18.02**	20.07	31.64	24.82	30.19	36.6
35–39	14.34	14.73	17.24	**23.3**	29.08	30.19	36.54	33.3
40–44	16.76	15.82	19.27	24.63	**32.68**	32.85	39.68	42.69
45–49	15.26	18.05	18.85	30.75	31.93	**35.83**	35.79	38.5
50–54	17.18	17.46	22.37	35.42	40.55	31.69	**39.32**	38.73
55–59	19.5	19.62	26.55	32.46	35.7	38	38.62	**46.9**
60–64	18.68	17.57	20.82	30.67	41.37	44.28	34.49	42.67
65–69	16.83	18.51	21.26	27.79	41.23	35.43	38	40.64
70–74	17.15	16.31	18.68	25.03	29.34	34.87	32.48	31.45

Bold black boxes denote persons who were 20 to 24 years of age during the 1971–1975 cycle and were followed over time, forming a cohort.
NHANES, National Health and Nutrition Examination Survey.

From Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971–2006. *Soc Sci Med.* 2010;70(7):1100–1108.
Note that the age and calendar time intervals do not correspond entirely but, unless there is a marked period effect during the "follow-up" of the cohort, the pattern shown in the table should approximate what would be seen with perfect alignment between time and age intervals.

TABLE 4.13 Possible Explanations of Trends or Differences in Mortality: I. Artifactual

1. Numerator	Errors in diagnosis Errors in age Changes in coding rules Changes in classification
2. Denominator	Errors in counting population Errors in classifying by demographic characteristics (e.g., age, race, sex) Differences in percentages of populations at risk

BOX 4.3 Possible Explanations of Trends or Differences in Mortality: II. Real

Change in diagnostic modalities or management strategies
Change in survivorship without change in incidence
Change in incidence
Change in age composition of the population(s)
A combination of the above factors

Although considerable controversy exists about which quality-of-life measures are most appropriate and valid, there is general agreement that such measures can be reasonably used to plan short-term treatment programs for groups of patients. Such patients can be evaluated over a period of months to determine the effects of the treatment on their self-reported quality of life. Quality-of-life measures have also been used for establishing priorities in situations of scarce health care resources. Although prioritizing health care resources is often primarily based on mortality data, quality of life must also be taken into account for this purpose, because many diseases are chronic and non–life-threatening but may be associated with many years of disability, having both physical and mental health consequences. Patients may place different weights on different quality-of-life measures depending on differences in

their occupations and other activities, personalities, cultural backgrounds, education, and moral and ethical values. As a result, measuring quality of life and developing valid indices that are useful for obtaining comparative data in different patients and in different populations remain major challenges.

Projecting the Future Burden of Disease

An interesting and valuable use of current data to predict the future impact of disease was a comprehensive assessment of current mortality and disability from diseases, injuries, and risk factors for all regions of the world in 1990, which was projected to the year 2020. The study, titled the Global Burden of Disease, attempted to quantify not only deaths but also the impact of premature death and disability on a population and to combine these into a single index to express the overall "burden of disease."[8] The index that was developed for this study is the disability-adjusted life year (DALY), which is the number of years of life lost to premature death and years lived with a disability of specified severity and duration. Thus, a DALY is 1 lost year of healthy life.

The results showed that 5 of the 10 leading causes of disability in 1990 were psychiatric conditions; psychiatric and neurologic conditions accounted for 28% of all years lived with disability of known severity and duration, compared with 1.4% of all deaths and 1.1% of years of life lost. Fig. 4.22 shows the burden of disease among the world's poorest billion people vs. high-income regions.[9,10] Again, the importance of ischemic heart disease in high-income countries and lower respiratory tract infections in low-income countries is dramatically evident.

In 2019 the disease burden was not equitably distributed. As seen in Table 4.14, the top 20 causes of disease burden were responsible for 57.2% of all DALYs. Five of them primarily affect children younger than 5 years of age. Three of the top ten (ischemic heart disease, stroke, and depression) are chronic conditions. This table shows the value of using a measure such as DALYs to assess the burden of disease, a measure that is not limited to either morbidity or mortality but is weighted by both.

With the aging of the population worldwide and advances in economic development, particularly in low- and middle-income countries, an "epidemiologic transition" is taking place so that, by 2020, noncommunicable diseases accounted for 70% of all deaths in developing countries. As projected in Fig. 4.23, by 2020,

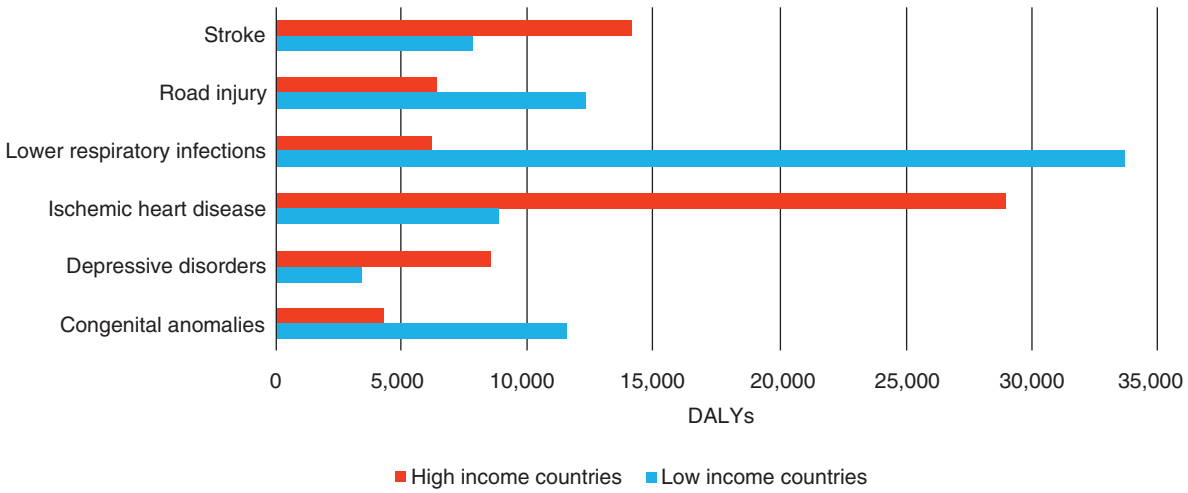


Fig. 4.22 Selected causes of disease burden by low- versus high-income countries, 2015. *DALYs,* Disability-adjusted life years. (From Global Health Estimates 2015. *Disease Burden by Cause, Age, Sex, by Country and by Region, 2000–2015.* Geneva, Switzerland: World Health Organization; 2016.)

TABLE 4.14 Global Health Estimates 2019: 20 Leading Causes of DALYs Globally, 2019

Rank	Cause	DALYs (000s)	% DALYs	DALYs per 100,000 Population
0	**All Causes**	**2,531,710**	**100.0**	**32,844**
1	Neonatal conditions	201,821	8.0	2,618
2	Ischemic heart disease	180,847	7.1	2,346
3	Stroke	139,429	5.5	1,809
4	Lower respiratory infections	105,652	4.2	1,371
5	Diarrheal diseases	79,311	3.1	1,029
6	Road injury	79116	3.1	1,026
7	Chronic obstructive pulmonary disease	73,981	2.9	960
8	Diabetes mellitus	70,411	2.8	913
9	Tuberculosis	66,024	2.6	857
10	Congenital anomalies	51,797	2.0	672
11	Back and neck pain	46532	1.8	604
12	Depressive disorders	46,359	1.8	601
13	Cirrhosis of the liver	42,798	1.7	555
14	Trachea, bronchus, lung cancers	41,378	1.6	537
15	Kidney diseases	40,571	1.6	526
16	HIV/AIDS	40,147	1.6	521
17	Other hearing loss	39,477	1.6	512
18	Falls	38,216	1.5	496
19	Malaria	33,398	1.3	433
20	Uncorrected refractive errors	31,981	1.3	415

AIDS, Acquired immunodeficiency syndrome; *DALYs,* disability-adjusted life years; *HIV,* human immunodeficiency virus.
Data from Global Health Estimates 2019: Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2019. Geneva, World Health Organization; 2020. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates

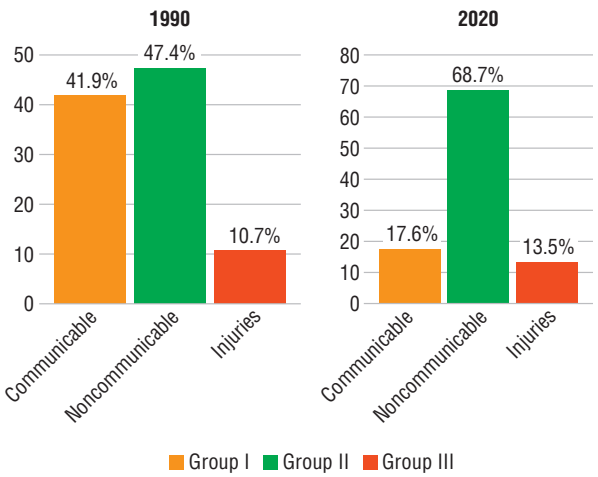


Fig. 4.23 The "epidemiologic transition": distribution of deaths from communicable and noncommunicable causes in developing countries, 1990 and projected into 2020. (From Murray CJL, Lopez AD. *The Global Burden of Disease: a Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020.* Cambridge: Harvard University Press on behalf of the World Health Organization and the World Bank; 1996.)

the disease burden due to communicable diseases, maternal and perinatal conditions, and nutritional deficiencies decreased dramatically. The burden due to noncommunicable diseases (group II) is expected to increase sharply, as will the burden from injuries (group III). The burden of disease attributable to tobacco exceeded that caused by any single disease—clearly a strong call for public health action. Although there is no universal agreement on the methodology or applicability of a single measure of disease burden such as the DALY, this study is an excellent demonstration of an attempt at worldwide surveillance designed to develop such a measure to permit valid regional comparisons and future projections so that appropriate interventions can be developed.

CONCLUSION

Chapters 3 and 4 have reviewed important approaches to quantitatively measuring and expressing human morbidity and mortality. The concepts reviewed in these chapters may at first seem overwhelming but, as

we shall see in later chapters, they are critical to understanding how epidemiology helps us to elucidate the measurement of disease risk, the determination of disease causation, and the evaluation of the effectiveness of intervening to modify the disease process.

In Chapter 5, we will turn to questions about the numerators of morbidity rates: how do we identify those people who have a disease and distinguish them from those who do not, and how do we evaluate the quality of the diagnostic and screening tests that are used to separate these individuals and populations? A discussion of the use of screening tests in public health programs is presented in Chapter 18.

REFERENCES

1. De Beauvoir S. *A Very Easy Death.* Translated by Patrick O'Brian. New York: Pantheon Books; 1965.
2. Centers for Disease Control and Prevention. *Premature mortality in the United States: public health issues in the use of years of potential life lost. MMWR Suppl.* 1986;35:1S–11S.
3. Lim H, Devesa SS, Sosa JA, et al. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA.* 2017;317(13):1338–1348.
4. National Center for Health Statistics (NCHS). *Instructions for Classifying the Underlying Cause of Death.* Hyattsville, MD: NCHS; 1983.
5. Chamblee RF, Evans MC. *TRANSAX: The NCHS System for Producing Multiple Cause-of-Death Statistics, 1968–1978. Vital and Health Statistics, Series 1, No. 20, DHHS Publication No. (PHS) 86–1322. Washington, DC: Bureau of Vital and Health Statistics; June 1986.*
6. Savage G, Rohde FC, Grant B, et al. *Liver Cirrhosis Mortality in the United States, 1970–90: Surveillance Report No. 29.* Bethesda, MD: Department of Health and Human Services; December 1993.
7. Anderson RN, Rosenberg HM. *Age Standardization of Death Rates: Implementation of the Year 2000 Standard. National Vital Statistics Reports,* Vol. 47, No. 3, pp. 1–16. Hyattsville, MD: National Center for Health Statistics; October 7, 1998.
8. Murray CJL, Lopez AD. *The Global Burden of Disease.* Cambridge, MA: Harvard University Press; 1996.
9. Coates MM, Ezzati M, Robles Aguilar G, et al. *Burden of disease among the world's poorest billion people: An expert-informed secondary analysis of Global Burden of Disease estimates. PLoS ONE.* 2021;16(8):e0253073. https://journals.plos.org/plosone/article/figure?id=10.1371/journal.pone.0253073.g003
10. World Health Organization (WHO). *The Global Burden of Disease: 2015 Update.* Geneva, Switzerland: WHO; 2015.

REVIEW QUESTIONS FOR CHAPTER 4

Questions 1 and 2 are based on the information given below:

In an Asian country with a population of 6 million people, 60,000 deaths occurred during the year ending December 31, 2010. These included 30,000 deaths from cholera in 100,000 people who were sick with cholera.

1. *What was the cause-specific mortality rate from cholera in 2010?* _____

2. *What was the case-fatality from cholera in 2010?* _____

3. *Age-adjusted death rates are used to:*
- **a.** Correct death rates for errors in the statement of age
- **b.** Determine the actual number of deaths that occurred in specific age groups in a population
- **c.** Correct death rates for missing age information
- **d.** Compare deaths in persons of the same age group
- **e.** Eliminate the effects of differences in the age distributions of populations in comparing death rates

4. *The mortality rate from disease X in city A is 75/100,000 in persons 65 to 69 years old. The mortality rate from the same disease in city B is 150/100,000 in persons 65 to 69 years old. The inference that disease X is two times more prevalent in persons 65 to 69 years old in city B than it is in persons 65 to 69 years old in city A is:*
- **a.** Correct
- **b.** Incorrect, because of failure to distinguish between prevalence and mortality
- **c.** Incorrect, because of failure to adjust for differences in age distributions
- **d.** Incorrect, because of failure to distinguish between period and point prevalence
- **e.** Incorrect, because a proportion is used when a rate is required to support the inference

5. *The incidence rate of a disease is five times greater in women than in men, but the prevalence rates show no sex difference. The best explanation is that:*
- **a.** The crude all-cause mortality rate is greater in women
- **b.** The case-fatality from this disease is greater in women

c. The case-fatality from this disease is lower in women
d. The duration of this disease is shorter in men
e. Risk factors for the disease are more common in women

6. *For a disease such as pancreatic cancer, which is highly fatal and of short duration:*
a. Incidence rates and mortality rates will be similar
b. Mortality rates will be much higher than incidence rates
c. Incidence rates will be much higher than mortality rates
d. Incidence rates will be unrelated to mortality rates
e. None of the above

7. *In 1990, there were 4,500 deaths due to lung diseases in miners aged 20 to 64 years. The expected number of deaths in this occupational group, based on age-specific death rates from lung diseases in all males aged 20 to 64 years, was 1,800 during 1990. What was the standardized mortality ratio (SMR) for lung diseases in miners?* _____

Question 8 is based on the information given below:

Annual Cancer Deaths in White Male Workers in Two Industries

	INDUSTRY A		INDUSTRY B	
Cancer Site	**No. of Deaths**	**% of All Cancer Deaths**	**No. of Deaths**	**% of All Cancer Deaths**
Respiratory system	180	33	248	45
Digestive system	160	29	160	29
Genitourinary	80	15	82	15
All other sites	130	23	60	11
Totals	550	100	550	100

Based on the preceding information, it was concluded that workers in industry B are at higher risk of death from respiratory system cancer than workers in industry A. (Assume that the age distributions of the workers in the two industries are nearly identical.)

8. *Which of the following statements is true?*
a. The conclusion reached is correct
b. The conclusion reached may be incorrect because proportionate mortality rates were used when age-specific mortality rates were needed
c. The conclusion reached may be incorrect because there was no comparison group
d. The conclusion reached may be incorrect because proportionate mortality was used when cause-specific mortality rates were needed
e. None of the above

9. *A program manager from an international health funding agency needs to identify regions that would benefit from an intervention aimed at reducing premature disability. The program manager asks a health care consultant to develop a proposal using an index that would help her make this decision. Which of the following would best serve this purpose?*
a. Case-fatality
b. Crude mortality rate
c. Disability-adjusted life-years
d. Standardized mortality ratio

10. *The following are standardized mortality ratios (SMRs) for lung cancer in England:*

	STANDARDIZED MORTALITY RATIOS	
Occupation	**1949–60**	**1968–79**
Carpenters	209	135
Bricklayers	142	118

Based on these SMRs alone, it is possible to conclude that:
a. The number of deaths from lung cancer in carpenters in 1949–60 was greater than the number of deaths from lung cancer in bricklayers during the same period
b. The proportionate mortality from lung cancer in bricklayers in 1949–60 was greater than the proportionate mortality from lung cancer in the same occupational group in 1968–79
c. The age-adjusted rate of death from lung cancer in bricklayers was greater in 1949–60 than it was in 1968–79
d. The rate of death from lung cancer in carpenters in 1968–79 was greater than would have been expected for a group of men of similar ages in all occupations
e. The proportionate mortality rate from lung cancer in carpenters in 1968–79 was 1.35 times

greater than would have been expected for a group of men of similar ages in all occupations

Questions 11, 12, and 13 are based on the information given below:

Numbers of People and Deaths from Disease Z by Age Group in Communities X and Y

	COMMUNITY X		COMMUNITY Y	
Age Group	No. of People	No. of Deaths From Disease Z	No. of People	No. of Deaths From Disease Z
Young	8,000	69	5,000	48
Old	11,000	115	3,000	60

Calculate the age-adjusted death rate for disease Z in communities X and Y by the direct method, using the total of both communities as the standard population.

11. *The age-adjusted death rate from disease Z for community X is: _____*

12. *The proportionate mortality from disease Z for community Y is: _____*

a. 9.6/1,000
b. 13.5/1,000
c. 20.0/1,000
d. 10.8/1,000
e. None of the above

13. *Which of the following statements regarding direct adjustment is TRUE?*

a. The age-adjusted mortality rate of community X is still higher than the mortality rate of community Y, as compared to the crude mortality rate
b. Age-adjusted mortality rates for community X should be used to make decisions regarding allocation of funding for hospital care of the dying in community X
c. For direct age-adjustment, the weight for a given age category is the percentage of deaths for that age group
d. For direct age-adjustment, the weight for a given age category is the number of individuals in the standard population for that age group
e. The difference in the adjusted mortality rates between community X and community Y is always attributable to differences in age composition between the two populations

14. *Surveillance data indicate that the prevalence of chronic liver disease in the United States increased 104% between the years 1990 and 2008. While chronic liver disease occurs in persons of all ages, the highest mortality rate occurs in people 65 years old or older. The United States has proportionately more people 65 years or older than Country X. What would happen if crude mortality rates in the United States were age standardized to the population of Country X in order to compare the risk of dying of chronic liver disease in the two populations?*

a. The age-standardized mortality rate for the United States would be less than the crude mortality rate for the United States
b. The age-standardized mortality rate for the United States would be greater than the crude mortality rate for the United States
c. The age-standardized mortality rate for the United States would be the same as the crude mortality rate for the United States
d. The age-standardized mortality rate for the United States cannot be used for this comparison
e. The age-standardized mortality rate for the United States would be the same as the proportionate mortality rate

15. *Among workers in a fish processing plant, 30% of all deaths were due to myocardial infarction. Among workers in a brewery, 10% of all deaths were due to myocardial infarction. Investigators concluded that workers in the fish processing plant had a greater risk of death due to myocardial infarction than workers in the brewery. This conclusion:*

a. Is correct
b. May be incorrect because it is based on proportionate mortality
c. May be incorrect because it assumes the same case fatality for myocardial infarction in both work sites
d. May be incorrect because consumed fish oil is protective against death due to myocardial infarction
e. May be incorrect because the prevalence of myocardial infarction in the two groups is not known

TABLES IN CHAPTER 6

that were updated after the first printing, with corrected dates, follow this page.

patient's 5-year survival, the outcome of the second scenario is any better than that of the first, because no change in the natural history of the disease has occurred, as reflected by the year of death. Indeed, the only change that has taken place is that when the diagnosis was made 3 years earlier (2010 vs. 2013), the patient received medical care for breast cancer, with all its attendant difficulties, for an additional 3 years. Thus, when screening is performed, a higher 5-year survival may be observed, not because people live longer, but only because an earlier diagnosis has been made. This type of potential bias (known as *lead time bias*) must be taken into account in evaluating any screening program before it can be concluded that the screening is beneficial in extending survival.

Another problem with 5-year survival is that if we want to look at the survival experience of a group of patients who were diagnosed less than 5 years ago, we clearly cannot use this criterion, because 5 years of observation are necessary in these patients to calculate 5-year survival. Therefore, if we want to assess a therapy that was introduced less than 5 years ago, 5-year survival is not an appropriate measure.

A final issue relating to 5-year survival is shown in Fig. 6.9. Here we see survival curves for two populations, A and B. Five-year survival is about 10%. However, the curves leading to the same 5-year survival are quite different. Although survival at 5 years is the same in both groups, most of the deaths in group A did not occur until the fifth year, whereas most of the deaths in group B occurred in the first year since they generally had a shorter time to event (death) compared with group A. Thus, despite the identical 5-year survivals, survival during the 5 years is clearly better for those in group A.

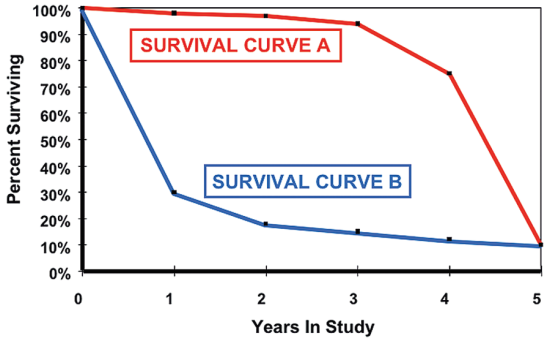


Fig. 6.9 Five-year survival curves in two hypothetical populations.

OBSERVED SURVIVAL

Rationale for the Life Table

Another approach to quantifying prognosis is to use the actual observed survival of patients followed over time, based on knowing the interval within which the event has occurred. For this purpose, we use a *life table*. Life tables have been used by actuaries to estimate risk in populations for centuries when there were no data on individuals. Actuarial methods and models have been applied in a large number of situations, including property/casualty, life insurance, pensions and health insurance, among others. Actuaries are credentialed, with a foundation of statistics and probability, stochastic processes, and actuarial methods and models.

Let's examine the conceptual framework underlying the calculation of survival rates using a life table, especially when the exact event time is not known, but rather we use the interval within which the event took place.

Table 6.1 shows a hypothetical study of treatment results in patients who were treated from 2019 to 2023 and followed to 2024. (By just glancing at this table, you can tell that the example is hypothetical, because the title indicates that no patients were lost to follow-up! However, there may have been administrative censoring in which the researcher removed some participants from further analysis.)

For each calendar year of treatment, the table shows the number of patients enrolled in treatment and the number of patients alive at each calendar year after the initiation of that treatment. For example, of 84 patients enrolled in treatment in 2019, 44 were alive in 2020, a year after beginning treatment; 21 were alive in 2021; and so on.

The results in Table 6.1 include all the data available for assessing the treatment. If we want to describe the prognosis in these treated patients using all of the data in the table, obviously we cannot use 5-year survival, because the entire group of 375 patients has not been observed for 5 years. We could calculate 5-year survival using only the first 84 patients who were enrolled in 2019 and observed until 2024, because they were the only ones observed for 5 years. However, this would require us to discard the rest of the data, which would be unfortunate, given the effort and expense involved in obtaining the data, and also given the additional light that the survival experience of those patients

TABLE 6.1 Hypothetical Study of Treatment Results in Patients Treated From 2019–2023 and Followed to 2024 (None Lost to Follow-Up)

		NUMBER ALIVE ON ANNIVERSARY OF TREATMENT				
Year of Treatment	No. of Patients Treated	2020	2021	2022	2023	2024
2019	84	44	21	13	10	8
2020	62		31	14	10	6
2021	93			50	20	13
2022	60				29	16
2023	76					43

TABLE 6.2 Rearrangement of Data in Table 6.1, Showing Survival Tabulated by Years Since Enrollment in Treatment (None Lost to Follow-Up)

		NUMBER ALIVE AT END OF YEAR				
Year of Treatment	No. of Patients Treated	1st year	2nd year	3rd year	4th year	5th year
2019	84	44	21	13	10	8
2020	62	31	14	10	6	
2021	93	50	20	13		
2022	60	29	16			
2023	76	43				

would cast on the effectiveness of the treatment. The question is: How can we use *all* of the information in Table 6.1 to describe the survival experience of the patients in this study?

To use all of the data, we rearrange the data from Table 6.1 as shown in Table 6.2. In this table, the data show the number of patients who started treatment each calendar year and the number of those who remained alive on each anniversary of the initiation of treatment. The patients who started treatment in 2023 were observed for only 1 year, because the study ended in 2024.

With the data in this format, how do we use the table? First we ask, "What is the probability of surviving for 1 year after the beginning of treatment?" To answer this, we divide the total number of patients who were alive 1 year after the initiation of treatment (197) by the total number of patients who started treatment (375; Table 6.3).

The probability of surviving the first year (P_1) is:

$$P_1 = \frac{197}{375} = 0.525$$

Next, we ask, "What is the probability that, having survived the first year after beginning treatment, the patients will survive the second year?" We see in Table 6.4 that 197 people survived the first year, but for 43 of them (the ones who were enrolled in 2023), we have no further information because they were observed for only 1 year. Because 71 survived the second year, we calculate the probability of surviving the second year, if the patient survived the first year (P_2), as:

$$P_2 = \frac{71}{197 - 43} = 0.461$$

In the denominator we subtract the 43 patients for whom we have no data for the second year.

Following this approach, we next ask, "Given that a person has survived to the end of the second year, what is the probability, on average, that he or she will survive to the end of the third year?"

In Table 6.5, we see that 36 survived the third year. Although 71 had survived the second year, we have no further information on survival for 16 of them because they

TABLE 6.3 Analysis of Survival in Patients Treated From 2019–2023 and Followed to 2024 (None Lost to Follow-Up): I

		NUMBER ALIVE AT END OF YEAR				
Year of Treatment	No. of Patients Treated	1st year	2nd year	3rd year	4th year	5th year
2019	84	44	21	13	10	8
2020	62	31	14	10	6	
2021	93	50	20	13		
2022	60	29	16			
2023	76	43				
Totals	375	197				

$$P_1 = \text{Probability of surviving the 1st year} = \frac{197}{375} = 0.525$$

TABLE 6.4 Analysis of Survival in Patients Treated From 2019–2023 and Followed to 2024 (None Lost to Follow-Up): II

		NUMBER ALIVE AT END OF YEAR				
Year of Treatment	No. of Patients Treated	1st year	2nd year	3rd year	4th year	5th year
2019	84	44	21	13	10	8
2020	62	31	14	10	6	
2021	93	50	20	13		
2022	60	29	16			
2023	76	43				
Totals		197	71			

$$P_2 = \text{Probability of surviving the 2nd year} = \frac{71}{197 - 43} = 0.461$$

TABLE 6.5 Analysis of Survival in Patients Treated From 2019–2023 and Followed to 2024 (None Lost to Follow-Up): III

		NUMBER ALIVE AT END OF YEAR				
Year of Treatment	No. of Patients Treated	1st year	2nd year	3rd year	4th year	5th year
2019	84	44	21	13	10	8
2020	62	31	14	10	6	
2021	93	50	20	13		
2022	60	29	16			
2023	76	43				
Totals			71	36		

$$P_3 = \text{Probability of surviving the 3rd year} = \frac{36}{71 - 16} = 0.655$$

TABLE 6.6 Analysis of Survival in Patients Treated From 2019–2023 and Followed to 2024 (None Lost to Follow-Up): IV

		NUMBER ALIVE AT END OF YEAR				
Year of Treatment	No. of Patients Treated	1st year	2nd year	3rd year	4th year	5th year
2019	84	44	21	13	10	8
2020	62	31	14	10	6	
2021	93	50	20	13		
2022	60	29	16			
2023	76	43				
Totals				36	16	

$$P_4 = \text{Probability of surviving the 4th year} = \frac{16}{36 - 13} = 0.696$$

TABLE 6.7 Analysis of Survival in Patients Treated From 2019–2023 and Followed to 2024 (None Lost to Follow-Up): V

		NUMBER ALIVE AT END OF YEAR				
Year of Treatment	No. of Patients Treated	1st year	2nd year	3rd year	4th year	5th year
2019	84	44	21	13	10	8
2020	62	31	14	10	6	
2021	93	50	20	13		
2022	60	29	16			
2023	76	43				
Totals					16	8

$$P_5 = \text{Probability of surviving the 5th year} = \frac{8}{16 - 6} = 0.800$$

were enrolled late in the study. Therefore, we subtract 16 from 71 and calculate the probability of surviving the third year, given survival to the end of the second year (P_3), as:

$$P_3 = \frac{36}{71-16} = 0.655$$

We then ask, "If a person survives to the end of the third year, what is the probability that he or she will survive to the end of the fourth year?"

As seen in Table 6.6, a total of 36 people survived the third year, but we have no further information for 13 of them. Because 16 survived the fourth year, the probability of surviving the fourth year, if the person has survived the third year (P_4), is:

$$P_4 = \frac{16}{36-13} = 0.696$$

Finally, we do the same calculation for the fifth year (Table 6.7). We see that 16 people survived the fourth year, but that no further information is available for 6 of them.

Because 8 people were alive at the end of the fifth year, the probability of surviving the fifth year, if the person has survived the fourth year (P_5), is:

$$P_5 = \frac{8}{16-6} = 0.800$$

Using all of the data that we have calculated, we ask, "What is the probability of surviving for all 5 years?" Box 6.1 shows all of the probabilities of surviving for each individual year that we have calculated.

Now we can answer the question, "If a person is enrolled in the study, what is the probability that he or she will survive 5 years after beginning treatment?" The