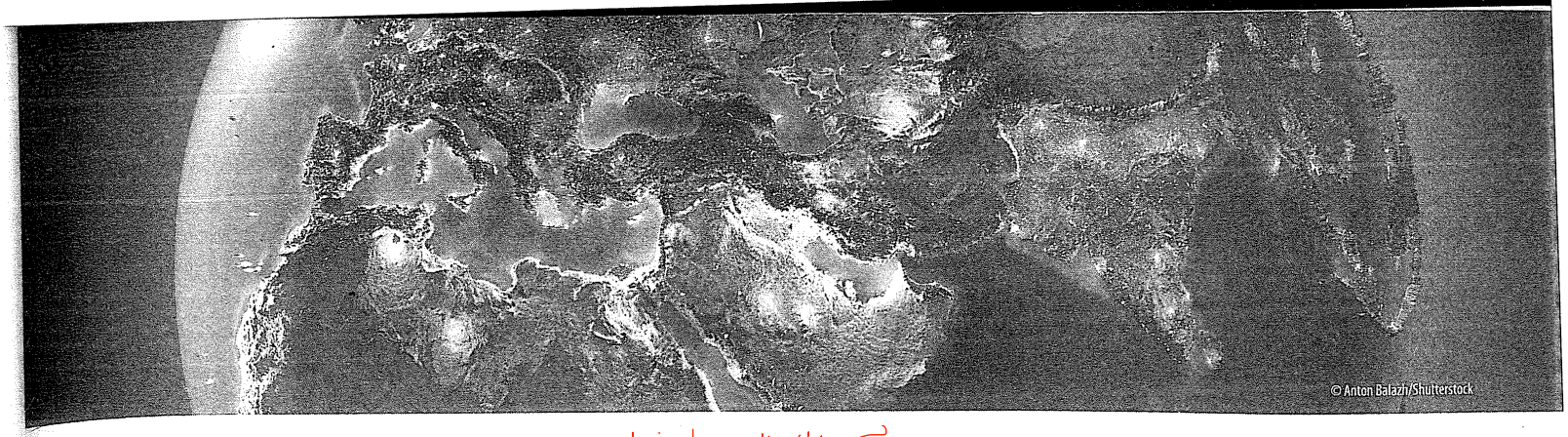


بسم الله الرحمن الرحيم

CHAPTER 1

Measures of Health and Disease in Populations

تعبر أساسيات القرارات في مجال الصحة والمرض

Abdulgafoor M. Bachani and Adnan A. Hyder

As Health & disease are extremely complex concepts ⟹ so we use prevention & treatment as indicators of health (bx. it's easier.)

In its 1948 charter, the World Health Organization (WHO) defined health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." Although this is an important ideological conceptualization, for most practical purposes, objectives of health programs are more readily defined in terms of prevention or treatment of disease. Disease has been defined in many ways and for a variety of reasons; distinctions may be made between disease, sickness, and illness. For purposes of defining and measuring disease burden, a general definition will be used in this text: *Disease* is anything that a person experiences that causes, literally, "dis-ease"—that is, anything that leads to discomfort, pain, distress, disability of any kind, or death constitutes disease. It may be due to any cause, including injuries or psychiatric conditions.

Very very general !! نحيل أكثر لتغطية لأكثر قريباً من الصحة بشكل عام.

It is also important to be able to diagnose and classify specific diseases to the extent that such classification aids in determining which health intervention programs would be most useful. Thus, defining disease, understanding the pathogenesis of the disease process, and knowing which underlying risk factors lead to this process are critical for understanding and classifying causes so as to determine the most effective prevention and treatment strategies for reducing the effects of a disease or risk factor. Just as the purpose of diagnosis of a disease in an individual patient is to provide the right treatment, so the major purpose of working through a burden of disease analysis in a population is to provide the basis for the most effective mix of health and social program interventions.

لأن التعريف واسع جداً.

Developments in the measurement of population health status and disease burden over the past two decades include the increasing use of summary, composite measures of health that combine the mortality and morbidity effects of diseases into a single indicator; the availability of results of Global Burden of Disease (GBD) studies, which make use of such summary indicators; and developments in the measurement of disability and risk factors. The more traditional approaches to measuring health are widely available in other public health textbooks and will be used for illustrative and comparative purposes here.

This chapter is divided into five sections. The first section explains the reasons for and approaches to measuring disease burden in populations, describes the need for using quantitative indicators, highlights the importance of using data for decision making in health, and lists a variety of major health indicators currently in widespread use. The second section critically reviews methods for developing and using composite measures that combine the mortality and morbidity from diseases in populations at national and regional levels. It explores the potential utility of these measures and discusses their limitations

هذا يكون

and implications. The third section demonstrates the application of these methods for measurement of health status and assessment of global health trends. It reviews current estimates and forecasts trends in selected countries and regions, as well as examines the global burden of disease. The fourth section reviews important underlying risk factors of disease and discusses recent efforts to measure the prevalence of major risk factors and to determine their contributions to regional and global disease burdens. The final section provides conclusions for the chapter.

Reasons for and Approaches to Measuring Health and Disease

Rationale

The many reasons for obtaining health-related information all hinge on the need for data to guide efforts toward reducing the consequences of disease and enhancing the benefits of good health. These include the need to identify which interventions will have the greatest beneficial effect, to identify emerging trends and anticipate future needs, to assist in determining priorities for expenditures, to provide information for education to the public, and to help in setting health research agendas. The primary information requirement is for understanding and assessing the health status of a population and its changes over time. In recent years, practitioners have emphasized the importance of making evidence-based decisions in health care. There is little reason to doubt that evidence is better than intuition, but realizing its full benefits depends upon recognizing and acting upon the evidence. This chapter examines evidence—the facts of health and disease—and demonstrates how to assemble this evidence so that it can assist in better decision making concerning health and welfare.

A well-documented example of the relationship between decision making and data can be seen in a classic health systems project in Tanzania (**EXHIBIT 1-1**). This

EXHIBIT 1-1 Using Evidence to Improve a Health System: An Example from Africa

The **Tanzania Essential Health Interventions Project** (TEHIP), a joint venture of the Tanzanian Health Ministry, the International Development Research Centre (IDRC), and the Canadian International Development Agency, starting in 1996 was conducted in two rural districts—Morogoro and Rufiji—with a combined population of approximately 700,000. The annual health spending in Tanzania was about $8 per capita. In Morogoro and Rufiji, TEHIP added resources on the condition that they be spent rationally; in other words, the amount of money spent on interventions should reflect the burden of disease. TEHIP conducted burden of disease analysis for the two districts and established a demographic surveillance system (DSS). The routine data from the DSS provided information for the district teams to support resource allocation based on disease burden. The organization found that the amount that the local health authorities spent on addressing each disease bore little relation to the actual burden of disease. Although childhood problems (e.g., pneumonia, diarrhea, malnutrition, measles) constituted 28% of the disease burden, only 13% of the budget was devoted to addressing them. Other conditions, meanwhile, attracted more than their fair share of resources. For example, 22% of the budget was targeted to tuberculosis, even though it accounted for less than 4% of years of life lost.

TEHIP promoted the use of burden of disease analysis, district accounts, and other mapping tools for more rational decision making in the districts. It also brought management tools and community voice techniques to the district teams. The result was better ability of district-level healthcare workers and managers to control and allocate resources and processes related to healthcare provision. The district teams decided to spend more on neglected diseases for which cost-effective treatments or preventive measures were available. The extra $1 per capita was enough to allow the district health authorities to align their spending to reflect the real disease burden. For example, sexually transmitted diseases received 3% of the budget prior to TEHIP's intervention; that percentage changed to 9.5% after the realignment. Malaria accounted for 30% of the years of life lost because of death and debilitating illness; the budget for malaria prevention and treatment programs increased from 5% of total spending in 1996 to 25% in 1998.

The results of TEHIP were documented as changes in health outcomes. In Rufiji, for example, infant mortality fell by 40% in 5 years. In fact, just between 1999 and 2000, infant mortality fell from 100 deaths per 1,000 live births to 72 deaths per 1,000 live births, while the proportion of children dying before their fifth birthdays dropped by 14%, from 140 per 1,000 to 120 per 1,000. The success of TEHIP and its approach led to replication and further innovation in not only Tanzania but also many other low- and middle-income countries (LMICs), including Burkina Faso, Ghana, and Nigeria.

For additional information on the TEHIP success story, visit the websites identified in the following sources.

Data from https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7814.pdf; https://www.idrc.ca/en/article/tanzanias-healthcare-breakthrough; https://www.researchgate.net/publication/268341357_Tanzania_Essential_Health_Interventions_Project_An_Overview_of_TEHIP_Interventions_Tanzania_Ministry_Of_Health_TANZANIA_ESSENTIAL_HEALTH_INTERVENTIONS_PROJECT_TEHIP_Interventions_-An_Overview

case illustrates how able people with good intentions had been making decisions routinely, only to find that using established methods to collect evidence on the burden of disease changed the nature and effectiveness of their own decisions. A major reason for the effective use of the evidence was that it was collected locally and put forward in a form helpful to decision makers.

Measuring Health and Disease

The relative importance (burden) of different diseases in a population depends on their frequency (incidence or prevalence), severity (the mortality and extent of serious morbidity), consequences (health, social, economic), and the specific people affected (gender, age, social and economic position).

A Counting Disease (frequency)

The first task in measuring disease in a population is to count its occurrence. Counting disease frequency can be done in several ways, and it is important to understand what these different methods of counting actually mean. The most useful way depends on the nature of the disease and the purpose for which it is being counted. There are three commonly used measures of disease occurrence: cumulative incidence, incidence density, and prevalence.

Cumulative incidence, or *incidence proportion*, is the number or proportion of new cases of disease that occur in a population at risk for developing the disease during a specified period of time. For this measure to have meaning, three components are necessary: a definition of the onset of the event, a defined population, and a particular period of time. The critical point is *new cases* of disease—the disease must develop in a person who did not have the disease previously. The numerator is the number of new cases of disease (the event), and the denominator is the number of people at risk for developing the disease. Everyone included in the denominator must have the potential to become part of the group that is counted in the numerator. For example, to calculate the incidence of prostate cancer, the denominator must include only men, because women are not at risk for prostate cancer. The third component is the period of time. Any time unit can be used as long as all those counted in the denominator are followed for a period comparable with those who are counted as new cases in the numerator. The most commonly used time denominator is one year.

Incidence density, which is often simply called *incidence rate*, is the occurrence of new cases of disease per unit of person-time. This metric directly incorporates time into the denominator and is generally the most useful measure of disease frequency; it is often expressed as new events per person-year or per 1,000 person-years. Incidence is a measure of events (in this case, the transition from a nondiseased state to a diseased state) and can be considered a measure of risk. This risk can be looked at in any population group, defined by age, sex, place, time, sociodemographic characteristics, occupation, or exposure to a toxin or any other suspected causal factor.

Prevalence is a measure of present status rather than of newly occurring disease. It measures the proportion of people who have the defined disease at a specific point of time. Thus, it is a composite measure made up of two factors—the incidence of the disease that has occurred in the past and its continuation to the present or to some specified point in time. That is, prevalence equals the incidence rate of the disease multiplied by the average duration of the disease. For most chronic diseases, prevalence rates are more commonly available than are incidence rates.

B Severity of Disease

To understand the burden of disease in a population, it is important to consider not only the frequency of the disease but also its severity, as indicated by the morbidity and premature mortality that it causes. *Premature mortality* is defined as death before the expected age of death had the disease not occurred. *Morbidity* is a statement of the extent of disability that a person suffers as a consequence of the disease over time and can be measured by a number of indicators, as discussed later in this chapter.

Mortality

Traditionally, mortality has been the most important indicator of the health status of a population. John Grant developed the first known systematic collection of data on mortality with the *Bills of Mortality* in the early 1600s in London. He described the age pattern of deaths, categorized them by cause as understood at the time, and demonstrated variations from place to place and from year to year. Mortality rates according to age, sex, place, and cause continue to be central information about a population's health status and a crucial input for understanding and measuring the burden of disease. Considerable literature exists on the use of mortality to indicate health status and its application to national and subnational levels (Murray & Chen, 1992).

The fact of death by age, sex, and place is required by law in most countries through death registration, and in many countries the cause of death through death

certification is required as well. Both provide essential information about the health status of a population. Nevertheless, in many low-income countries, the fact of death, let alone its cause, is still not reliably available.

In high-income countries, vital statistics (i.e., the registration of births and deaths by age, sex, and place) are routinely collected and highly reliable. In most middle-income countries, the reliability and completeness of these data have been steadily improving and often are fairly satisfactory. In contrast, the collection of vital statistics remains grossly incomplete in many low-income countries. An analysis of death registration in the course of the Global Burden of Disease study showed that vital registration data together with sample registration systems still do not cover 100% of global mortality. Survey data and indirect demographic techniques are needed to provide information on levels of child and adult mortality to paint a complete picture of global mortality (GBD 2015 Mortality and Causes of Death Collaborators, 2016). Nevertheless, even in low-income countries, increasing use of survey methods is delivering useful estimates of the mortality rates for the population younger than age 5 years and other populations.

Obtaining information about cause of death remains difficult even in many middle-income countries; a lot of information depends on special surveys or studies of select populations. Verbal autopsies (VAs) have been used increasingly for judging the likely cause of death, especially for children younger than age 5. This method comprises structured questions administered by trained interviewers with family members after a death; the information is then reviewed by physicians (or computers) to assign a cause of death using algorithms. VAs are useful for assessing some causes of death such as neonatal tetanus and severe diarrhea, but their sensitivity and specificity may be limited for diseases whose symptoms are variable and nonspecific, such as malaria (Anker et al., 1999; Thatte, Kalter, Baqui, Williams, & Darmstadt, 2009). Recently, automated systems for analyzing VA data have been developed and are being tested.

Age-specific mortality profiles are a prerequisite for a burden of disease analysis. Although extensive work has been done to document and analyze child mortality in low- and middle-income countries (LMICs), less has been done for adult mortality (Hill, 2003). LMICs have higher rates of age-specific adult mortality than do high-income nations (GBD 2015 Mortality and Causes of Death Collaborators, 2016; Lopez et al., 2002; Murray & Chen, 1992). Indeed, mortality rates are higher for both women and men in LMICs at every age when compared with the high-income world. In Africa, the enormous increase in deaths of young and middle-aged women and men from acquired immunoeficiency syndrome (AIDS) has had a profound impact on mortality and survival (**EXHIBIT 1-2**).

Traditional indicators of mortality have been the standard for assessing population health status. Neonatal mortality rates (NMR; deaths of live-born infants before 28 days of age per 1,000 live births), infant mortality rates (IMR; deaths of live-born infants before 12 months of age per 1,000 live births), and child mortality (deaths of children younger than 5 years of age) are considered sensitive indicators of the overall health of nations. The United Nations Children's Fund (UNICEF) publishes an annual global report that includes a ranking of nations based on these indicators (UNICEF, 2015). These indicators have the added advantage of having been studied for their relationships with other indicators of the social and economic development of nations. For example, a clear relation exists between the gross national product (GNP) per capita, which is an indicator of national wealth, and child mortality. In general, the higher the level of economic development, the lower the rate of child mortality. However, there are exceptions, and they need to be examined carefully. For example, Sri Lanka and the Indian state of Kerala are both low-income regions that have traditionally had low child mortality rates. These examples demonstrate that the relationship between mortality and poverty is complex and needs in-depth investigation.

There are major deficiencies in cause-specific mortality data in low- and most middle- income countries. In keeping with demographic and epidemiologic transitions, the pattern of cause-specific mortality changes at different levels of total mortality, with a general trend of decreasing infectious and parasitic disease cause-specific mortality with declining total mortality. Indeed, mortality from these communicable causes has been a major reason for the historical difference between high- and low-mortality populations (Murray & Chen, 1992).

The cause of death certification system based on WHO's *International Classification of Diseases* (ICD) has been used widely in many countries for many years (WHO, 2016). Despite the existence of this standardized process for categorizing deaths, variations in the reliability of these data occur because of variations in the training and expertise of the people who are coding causes of death, as well as the supervision and feedback provided. Nevertheless, there have been steady improvements in many countries, including automation of data collection and analysis, and these kinds of data provide some of the best information available on major causes of mortality.

EXHIBIT 1-2 Trends of the HIV/AIDS Epidemic

Globally, AIDS is the leading infectious cause of death among 15- to 49-year-olds (Global Burden of Disease Risk Factors Collaborators et al., 2015). Untreated disease caused by the human immunodeficiency virus (HIV) has a case fatality rate that approaches 100% (WHO, 2003). Unknown more than 30 years ago, the HIV/AIDS epidemic has killed more than 35 million people.

At the end of 2016, an estimated 36.7 million people were living with HIV/AIDS, with 69.5% of those individuals living in sub-Saharan Africa (WHO, 2017) . The prevalence of new HIV infections has increased by 60% in eastern Europe and Central Asia (**TABLE 1-1**) (UNAIDS, 2017; WHO, 2017).

HIV/AIDS is the tenth leading cause of disability-adjusted life years, accounting for 2.7% of this global burden. In terms of mortality, it is the eleventh leading cause of death among people of all ages, accounting for 2.1% of all deaths (Global Burden of Disease Risk Factors Collaborators et al., 2015). Nearly 42% of the 1 million global deaths from HIV/AIDS have occurred in East and Southern Africa (**FIGURE 1-1**) (UNAIDS, 2017).

TABLE 1-1 Global Summary of HIV and AIDS Epidemic

Number of people living with HIV	**Total**	**36.7 million**	**(30.8–42.9 million)**
	Adults	34.5 million	(28.8.4–40.2 million)
	Women	17.8 million	(15.4–20.3 million)
	Children	2.1 million	(1.7–2.6 million)
Number newly infected with HIV	**Total**	**1.8 million**	**(1.6–2.1 million)**
	Adults	1.7 million	(1.4–1.9 million)
	Children	160,000	(100,000–220,000)
AIDS deaths	**Total**	**1.0 million**	**(830,000–1.2 million)**
	Adults	890,000	(740,000– 1.1 million)
	Children	120,000	(79,000–160,000)

Reproduced from Joint United Nations programme on HIV/AIDS (UNAIDS). (2017). *UNAIDS Data 2017*. Geneva, Switzerland: Author. Retrieved from http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf

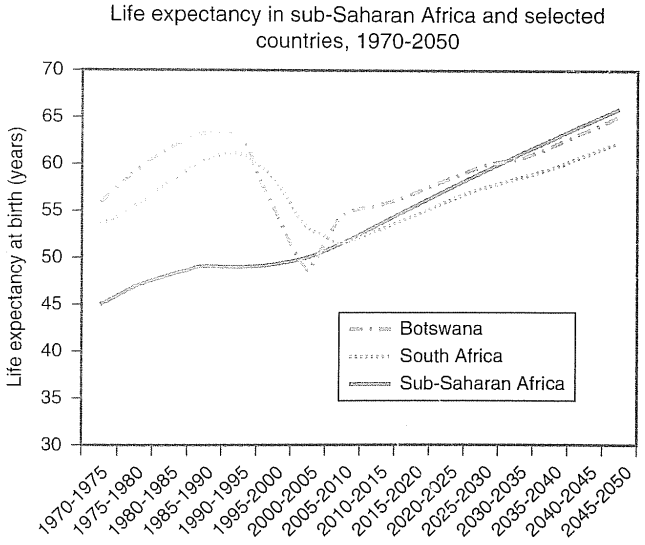


FIGURE 1-1 Trends in life expectancies and the HIV/AIDS epidemic.

Courtesy of Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. (2008). *World population prospects: The 2008 revision*. Retrieved from www.un.org/esa/population/publications/wpp2008/wpp2008_highlights.pdf

Reproduced from Joint United Nations programme on HIV/AIDS (UNAIDS). (2017). *UNAIDS Data 2017*. Geneva, Switzerland: Author. Retrieved from http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf; World Health Organization (WHO). (2017). *Global health observatory (GHO) data – HIV/AIDS*. Geneva, Switzerland: Author. Retrieved from http://www.who.int/gho/hiv/en/

Mortality can be expressed in two important quantitative measures: (1) *mortality rate* (MR) and (2) *case fatality ratio* (CFR). The MR, a form of incidence rate, is expressed as the number of deaths in a defined population in a defined time period. The numerator can be total deaths, age- or sex-specific deaths, or cause-specific deaths; the denominator is the number of persons at risk of dying in the stated category as defined earlier for incidence. Demographers use the notation XqY for the probability of dying in the Y years following age X at the then prevailing age-specific mortality rates for the population. Thus, 5q0 is the probability of death of newborns by age 5 years (see Table 1-2 later in this chapter), and 30q15 is the probability of death in young adults from age 15 to 45. The CFR is the proportion of those persons with a given disease who die of that disease (at any time, unless specified). The MR is equal to the CFR multiplied by the incidence rate of the disease in the population.

The distinction between the proportion of deaths attributable to a cause (number of deaths due to the cause divided by total number of deaths in a given population in a given time period) as compared to the probability of death from the cause (disease-specific MR) is important to understand. For example, the probability of death (and disability) from noncommunicable causes (indeed, from virtually all causes) is higher in low- and middle-income regions than in the high-income world. However, the proportion of deaths and disability attributable to these chronic causes is smaller in LMICs than in wealthier countries because of the much larger toll taken by infectious and nutritional causes. With increasing economic development, the risk of death and disability from chronic disease does not increase; rather, the proportion of deaths attributable to chronic disease increases as the proportion of deaths attributable to communicable and nutritional disease declines.

Demographic and Epidemiologic Transitions

The demographic transition describes the changes in birth and death rates that historically have accompanied the shift from a traditional society to a modern society; it is detailed in other chapters. With modernization, sharp declines in mortality have been followed by a reduction in fertility, albeit one that commonly lags behind the change in the death rate by years or decades. The term *transition* refers to the shift away from a stable population in which very high birth rates are balanced by very high death rates to a stable population in which low birth rates are balanced with low death rates. In between these extremes, as a society undergoes modernization, there is a lag between falling mortality, especially in the under-5 age group, and the drop in birth rates that leads to explosive population growth. Thereafter birth rates fall and a new stage is reached in which birth and death rates are low and balance resumes. The result is a striking change in the age structure of the population, with a decreased proportion of children and an aging population. These changes in the population age distributions are reflected in the shift from a wide-based pyramid, reflecting larger numbers in the younger age groups, to a structure with a narrow base, nearly rectangular configuration, and nearly equal percentages in each age group.

In 1971, Omran described the underlying reasons for the demographic transition and used the term *epidemiologic transition* to explain the changing causal factors of disease that accounted for the dramatic drop in under-5 mortality, which was largely due to reduction in malnutrition and communicable diseases. Although high rates of maternal mortality are characteristic of the low- and middle-income world, reductions in maternal mortality occur in a different time frame from those in under-5 mortality. Reductions in maternal mortality require a better-developed infrastructure, including ready availability of surgical and blood transfusion capacity plus improved communication and transportation systems. Thus, drops in maternal mortality occur much further along the road toward economic development, and changes occur only after shifts in the child mortality have been seen.

Major changes in the patterns and causes of injury are also likely to occur with modernization. For example, road traffic injuries tend to increase as countries go through the stage of development in which there is a great increase in vehicles and in the speeds at which they are operated before improved roads, appropriate laws and regulations, and law enforcement are in place (Crooper & Kopits, 2003; WHO, 2015). There may also be important shifts in the nature of violence and the people toward whom it is directed, related to crime patterns, civil unrest, ethnic conflicts, and intrafamily tensions (WHO, 2002b). The profound impact of the HIV/AIDS epidemic was discussed earlier in Exhibit 1-2.

Other Health-Related Metrics

In addition to basic measures of mortality, morbidity, and life expectation that are central for population health status assessment, a variety of important health-related indicators are useful for specific purposes. Many are discussed more fully in other chapters of this text; they are summarized in **TABLE 1-2**. Those related to the Sustainable Development Goals (SDGs) are discussed in **EXHIBIT 1-3**.

TABLE 1-2 Health-Related Metrics

Type	Indicator	Definition/Interpretation
Demographic indicators: reproductive health	Maternal death	Death of a woman while pregnant or up to 42 days post-delivery from any cause except accident
	Maternal mortality ratio	Maternal deaths per number of pregnancies (maternal deaths per 100,000 live births)
	Maternal mortality rate	Maternal deaths per number of women of reproductive age (maternal deaths per 100,000 women aged 15–49)
	Lifetime risk of maternal mortality	Cumulative loss of human life due to maternal death over the female life course
	Total fertility rate	Average number of children a woman would bear if she lived to the end of her reproductive period
	Life expectation at birth	Average number of years a newborn would live if his or her life were lived under the mortality conditions for the place and year in question
Anthropometric indicators: nutrition	Weight for age	Underweight
	Height for age	Stunting
	Weight for height	Wasting
	Mid-upper arm circumference	Wasting
Mortality (death) indicators	Mortality rate	Number of deaths in a specified time period/number of persons at risk of dying during that period
	Infant mortality rate	Number of deaths of live born infants before 12 months of age per 1,000 live births
	Under-5 mortality rate	Number of deaths of children younger than age 5 per 1,000 live births averaged over the last 5 years
	5q0	Probability of death of a newborn by age 5
	Neonatal mortality rate	Number of deaths of live-born infants before 28 days of age per 1,000 live births
	Stillbirth rate	Number of babies born with no signs of life at or after 28 weeks' gestation per 1,000 births
	Perinatal mortality rate	Number of fetal deaths (28 or more weeks of gestation) + postnatal deaths (first week) per 1,000 live births
Disease frequency	Endemic	Usual occurrence of a given disease in a defined population
	Epidemic	Occurrence of a given disease in a defined population clearly in excess relative to its usual occurrence
	Pandemic	A worldwide epidemic involving large numbers

EXHIBIT 1-3 Sustainable Development Goals

On September 25, 2015, UN member states adopted 17 *Sustainable Development Goals* (SDGs) as part of the post–Millennium Development Goals (MDGs) development agenda. The SDGs build on the goals identified under MDGs and broaden the scope to include new areas of focus—for example, climate change, economic inequality, innovation, sustainable consumption, and peace and justice—to improve overall well-being and life of current and future generations through sustainable means.

The 17 SDGs were divided into 169 quantifiable targets that are measured by 230 indicators. Of these, 21 targets and 39 indicators are directly related to health. The health-related indicators include a variety of indicator types: incidence rates, prevalence "rates," mortality rates, mortality ratios, birth rates, and proportion of target populations receiving an intervention.

Appendix 1 provides examples of some of the health-related SDG indicators.

Morbidity and Disability

Measures of mortality have been the principal indicators of population health status for generations. Their relative ease of observation, availability of data, and history of use make mortality information useful for assessing and monitoring the health status of populations. However, the key limitation with mortality-based indicators is that they "note the dead and ignore the living" (Kaplan & Anderson, 1996). Measurements of morbidity, by comparison, are more problematic because there is not a clearly defined endpoint such as death provides. In addition, several components of disability need to be assessed, and there may be a substantial subjective aspect to grading the extent or severity of a condition.

The *International Classification of Impairments, Disabilities, and Handicaps* (ICIDH) was developed in the 1970s to classify nonfatal health outcomes as an extension of WHO's ICD system (WHO, 1980). It was developed to more fully describe the impact of a given disease on an individual and on society, and to account for that disease's heterogeneity of clinical expression and evolution in different individuals and societies. ICIDH categories included *impairment* (loss or abnormality of psychological, physiological, or anatomic structure or function), *disability* (restriction or lack of ability to perform an activity considered normal), and *handicap* (disadvantage from a disability or impairment for a given individual based on the inability to fulfill a normal role as defined by age, sex, or sociocultural factors). These distinctions clarified more than just processes—they helped define the contribution of medical services, rehabilitation facilities, and social welfare to the reduction of disability.

In 2002, WHO built on the ICIDH to develop the *International Classification of Functioning, Disability, and Health*, commonly known as ICF (WHO, 2002c). In this system, health-related domains are classified from the perspectives of the body, of the individual, and of society by means of two lists: a list of body functions and structures, and a list of domains of activity and participation. Because an individual's functioning and disability occur within a context, the ICF also includes a list of environmental factors that provide a description of that context. The ICF has become WHO's framework for measuring health and disability at both individual and population levels. It was officially endorsed by all 191 WHO member states in the Fifty-Fourth World Health Assembly on May 22, 2001 (resolution WHA 54.21). Unlike its predecessor, which was endorsed for field trial purposes only, the ICF was endorsed for use in member states as the international standard to describe and measure health and disability.

Using such classifications, indicators of disability—such as *impairment-free*, *disability-free*, and *handicap-free* life expectancies—have been developed. These, in turn, have been used to estimate health-adjusted life expectancies using severity and preference weights for time spent in states of less than perfect health.

Hospital inpatient discharge records—when they are based on good clinical evidence and coded by staff well trained in coding procedures—can provide high-quality data on the major causes of morbidity serious enough to require hospitalization. They also can provide good cause-of-death data for hospitalized persons, and some sense of the outcome status of those with serious conditions. Hospital data are generally improving in quality, especially in middle-income countries and in selected sentinel (usually tertiary care) teaching hospitals in low-income countries. Such information is inevitably biased because of the highly skewed distribution of those using such hospitals, but in many situations it is possible to have a good understanding of those biases and make appropriate adjustments to draw useful conclusions.

Generally, outpatient records in most of the world are highly deficient in terms of diagnosis; indeed,

they often identify only the patient's chief complaint and the treatment dispensed. The main value of most such records is limited to establishing the fact of using a facility. There are usually strong biases in terms of those patients who use outpatient facilities because of access factors (distance and cost of use), nature and severity of the disease problem, and opportunity for using alternative services.

Visits to healthcare facilities, functional disability (a measure of activity that is less than the norm), and time spent away from work (absenteeism, work days lost) have all been used to assess the magnitude of morbidity from various conditions. A commonly used approach to evaluating morbidity in a population has been the assessment of the impact on social roles or functional performance, such as days missed from work or spent in bed (Kaplan & Anderson, 1996). A considerable body of literature focuses on the wide variety of instruments used to measure such functional capacity, especially in the clinical medical literature, that are not directly useful for population-based morbidity assessment.

Data about morbidity are often based on self-perceived assessments, and are frequently gleaned from survey-based interview information. The perception of morbidity and its reporting, the observation of morbidity and its impact, and other factors are responsible for the wide variations between reported and measured prevalence of conditions (Murray & Chen, 1992). This has resulted in an underestimation of the presence and impact of morbidity in both LMICs as compared with high-income nations. This situation also underscores the variations in morbidity data, which are often interpreted as indicating that wealthy individuals and low-mortality populations report higher rates of morbidity (*Global Burden of Disease and Risk Factors*, 2006; Woolf et al., 2015).

Measurement of health-related quality of life has also been discussed in the medical literature for decades. *Health-related quality of life* refers to how well an individual functions in daily life and his or her perception of well-being. Various domains of quality have been defined, such as health perception, functional status, and opportunity, and several instruments have been developed to evaluate them. Both disease-specific and general instruments exist, with such tools abounding in fields dealing with patients having chronic disabled states, such as psychiatry, neurology, and counseling. These scales are often dependent on self-reported information, although some incorporate observational data as well. However, concerns have been raised about their reliability and validity. These measures are not discussed further in this text, because they have been primarily used in clinical assessments of individuals, rather than larger populations.

Measuring Disability

If all the various forms of disability—physical, functional, mental, and social—are to be compared with mortality, they must be measured in an equivalent manner for use in health assessments. To do so, measurement of disability must quantify the duration and severity (extent) of this complex phenomenon. A defined process is needed that rates the severity of disability as compared with mortality, measures the duration of time spent in a disabled state, and converts various forms of disability into a common scale. General measures of disability without regard to cause (often carried out by special household surveys) are useful to determine the proportion of the population that is "disabled" and unable to carry out normal activities, but are not much help for quantifying the extent of disability.

In general, three components of disability need to be assessed. The first component is the *case disability ratio* (CDR)—the proportion of those diagnosed with the disease who have disability. For most diseases that are diagnosed clinically, the CDR will be 1.00 because, by the definition of disease given earlier, patients will have signs or symptoms. In contrast, when the diagnosis is based on, for example, infection rather than disease (such as tuberculosis) or on a genetic marker rather than the physical manifestation (such as sickle cell trait), the CDR is likely to be less than 1.00.

The second component of disability is its *extent or severity*—how incapacitated the person is as a result of the disease. The extent of disability is expressed on a scale, such as from 0 (indicating no disability) to 1.00 (equivalent to death). The assessment of severity can be quite subjective, particularly because so many different types and dimensions of disability exist. A number of methods have been introduced in an effort to achieve comparability and obtain consistency (Murray, Salomon, Mathers, & Lopez, 2002).

Measurement of individual preferences for different health states to determine relative severity of disability has been done by a variety of methods (Kaplan & Anderson, 1996; Murray et al., 2002; Torrance, 1986). Factors that influence the assessment of such preferences include the type of respondent, the type of instrument used to measure the response, and the time from entry into the disabled state. Individuals who are in a particular state, healthy individuals, healthcare providers, caretakers, and family members have all

go to Page 11 please

been interviewed in studies. Adaptation, conditioning, development of special skills, and vocational training can all change the response of individuals over time within a particular health state, thereby affecting the value of that state to the individual. As a consequence, the valuation is time dependent—for example, the value placed on a year of life by a paraplegic soon after entering that health state would be different from that obtained after several years of adjustment to that state (Murray & Lopez, 1994).

Instruments used to extract such preferences involve visual and interview techniques (*Global Burden of Disease and Risk Factors*, 2006; Torrance, 1986). Two alternative scenarios are often presented to the subject and the point of indifference sought (as in standard gamble techniques). Despite much work in this area, there is no consensus or accepted standard method for such elicitation. Severity of disability scales have been developed by group consensus using community surveys (Kaplan & Anderson, 1996), a mixture of community and expert groups (Ghana Health Assessment Team, 1981), experts only (World Bank, 1993), and population surveys (Global Burden of Disease Risk Factors Collaborators et al., 2015; Murray et al., 2002; Salomon et al., 2012). These scales usually compare perfect health states to death on a scale of 0 to 1 (**TABLE 1-3**).

In the first Global Burden of Disease 1990 study, the disability severity estimates were based on expert opinion. Twenty-two indicator conditions were selected and used to construct seven disability classes (Table 1-3). Outcomes from all other health conditions were categorized within these seven classes (with special categories

TABLE 1-3 Examples of Disability Classification Systems

Ghana Health Assessment Team, 1981		
Class	**Severity**	**Equivalent to (Maximum)**
1	0	Normal health
2	0.01–0.25	Loss of one limb's function
3	0.26–0.50	Loss of two limbs' function
4	0.51–0.75	Loss of three limbs' function
5	0.76–0.99	Loss of four limbs' function
6	1	Equivalent to death

Global Burden of Disease Study, 2013*	
Category	**Disability Weight for Unique Health States**
Infectious disease	Acute episode, mild: 0.006 (0.002–0.012) Acute episode, moderate: 0.051 (0.032–0.074) Diarrhea, mild: 0.074 (0.049–0.104) Ear pain: 0.013 (0.007–0.024)
Cancer	Diagnosis and primary treatment: 0.288 (0.193–0.399) Metastatic: 0.451 (0.307–0.600)
Cardiovascular and circulatory disease	Acute myocardial infarction (MI), days 1–2: 0.432 (0.288–0.579) Acute MI, days 3–28: 0.074 (0.049–0.105) Heart failure, mild: 0.041 (0.026–0.062)
Diabetes and digestive and genitourinary disease	Diabetic foot: 0.020 (0.010–0.034) Gastric bleeding: 0.325 (0.209–0.462) Infertility, primary: 0.008 (0.003–0.015)

Chronic respiratory disease	Asthma, controlled: 0.015 (0.007–0.026) Chronic obstructive pulmonary disease (COPD) and other chronic respiratory diseases, mild: 0.019 (0.011–0.033) COPD and other chronic respiratory diseases, severe: 0.408 (0.273–0.556)
Neurologic disorders	Dementia, mild: 0.069 (0.046–0.099) Multiple sclerosis, moderate: 0.463 (0.313–0.613) Parkinson's disease, severe: 0.575 (0.396–0.730)
Mental, behavioral, and substance use disorders	Alcohol use disorder, very mild: 0.123 (0.082–0.177) Anxiety disorders, moderate: 0.133 (0.091–0.186) Anorexia nervosa: 0.224 (0.150–0.312)
Hearing and vision loss	Hearing loss, mild: 0.010 (0.004–0.019) Hearing loss, profound, with ringing: 0.277 (0.182–0.387) Distance vision, severe impairment: 0.184 (0.125–0.258)
Musculoskeletal disorders	Low back pain, moderate: 0.054 (0.035–0.079) Neck pain, severe: 0.229 (0.153–0.317) Gout, acute: 0.295 (0.196–0.409)
Injury	Burns, lower airway, with or without treatment: 0.376 (0.240–0.524) Crush injury, short or long term, with or without treatment: 0.132 (0.089–0.189) Concussion: 0.110 (0.074–0.158)
Other	Abdominopelvic problem, mild: 0.011 (0.005–0.021) Anemia, moderate: 0.052 (0.034–0.076) Hypothyroidism: 0.019 (0.010–0.032)

* Health states included in this table are only examples. The full list of 235 unique health states and their disability weights are available in Salomon et al., 2015.
Data from Salomon, J. A., Haagsma, J. A., Davis, A., de Noordhout, C. M., Polinder, S., Havelaar, A. H., et al. (2015). Disability weights for the Global Burden of Disease 2013 study. *Lancet Global Health*, *3*(11), e712-723. doi: 10.1016/S2214-109X(15)00069-8

for treated and untreated groups). This approach was revised for the 2010 iteration of the study, for which the process included empirical studies comprising simple paired questions, and survey of the general public through household surveys in countries such as Bangladesh, Indonesia, Peru, Tanzania, and the United States, as well as a web-based survey in English, Spanish, and Mandarin. The result was a categorization that included 220 health states (Salomon et al., 2012). Generally, for most conditions a reasonable degree of consensus can be reached within broad categories (e.g., 25% disabled as compared with 50%), but efforts to reach much finer distinctions have proved equivocal. The need to seek out more refined scales for purposes of health program decision making ought to be a national or local decision.

The third component of disability is its *duration*. The duration is generally counted from onset of the disability until cure, recovery, or death. Sometimes there is continuing permanent disability after the acute phase is completed; in such a scenario, the duration would be the remaining life expectation from the time of onset of disease.

Data for Decisions

In the collection and assessment of information, the level of precision required should be guided by the purpose for collecting the information and depend on the decisions to be taken. Even rough estimates may be helpful; though disconcerting to some, the time and cost of efforts to realize further precision need to be justified by the increased precision's potential impact on decision making. LMICs, with their scarce resources, need timely and appropriate information to plan and implement health interventions that maximize the health of their populations. Methods, indicators, and assessments of disease must support and contribute to this primary purpose of health systems.

Decisions concerning deployment of interventions against diseases and underlying risk factors ideally should be taken such that maximum healthy life per resource expenditure is obtained in an equitable, fair, and just fashion. The ultimate reason for obtaining health data is to have the information to guide such decision making.

Summary Measures of Population Health

This section focuses on the major approaches used for developing composite measures of population health status that summarize mortality and morbidity occurring in a population through the use of a single number. It discusses the rationale for composite measures, reviews the origins of each major approach, examines methodological differences among these approaches, and outlines the advantages and limitations of each.

Rationale for Composite Measures

Rationing of healthcare resources is a fact of life everywhere; choices about the best use of funds for health must inevitably be made (Hyder, Rotllant, & Morrow, 1998; WHO, 2000). The global scarcity of resources for health care is a challenge for every country, rich and poor (Evans, Hall, & Warford, 1981; Figgis & Walters, 2015–2017; World Bank, 1993), but the realities in LMICs paint the choices in much starker terms. It is even more important for LMICs to choose carefully how to optimize health expenditures so as to obtain the most health in the most equitable fashion from these expenditures. Important tools under development to assist in making better choices for health spending are based on measures of the effectiveness of health interventions in improving health status in relation to their cost.

In most sectors, decisions on resource allocation are based on perceived value for money. The health sector, however, has had no coherent basis for determining the comparative value of different health outcomes (from different health programs). To make decisions about whether to put money into programs that reduce mortality in children, as compared with those programs that reduce disabling conditions in adults, a common denominator is needed. In recent decades, work has been carried out to develop composite indicators combining morbidity and mortality into a single measure that may serve as a common denominator for comparing different health outcomes. A common unit of measure for these different health outcomes is *time lost from healthy life*.

The most important reason for attempting to capture the complex mix of incommensurable consequences resulting from disease within a single number is the need to weigh the benefits of health interventions against their costs. Costs of health programs are expressed in a unidimensional measure, such as U.S. dollars; therefore, the benefits to be achieved from their expenditure should be expressed in the same manner. *Healthy lifetime* is a unidimensional measure that can be used to compress health benefits and losses into a single time dimension. An explicit, objective, quantitative approach should enable better budgetary decisions and permit resource allocation in the health sector to be undertaken in a more effective and equitable fashion.

Note that a composite indicator is simply a tool to be used to assist decision makers in resource allocation. Like any tool, it can be misused. Conclusions that are reached on the basis of these indicators must be carefully examined. Not only do problems arise in trying to put so many dimensions together, which inevitably may lead to distortions, but serious issues also emerge concerning the reliability and validity of the information on which these indicators are based. Thus, all the problems associated with determining causes of death, counting the number of cases of disease, and assessing the extent of disability from a condition will lead to uncertainties when these factors are added and multiplied together. The development of a single indicator consisting of a specific number implies deceptive stability about something that may actually be composed of fragile data. Continuing vigilance in how these data are obtained, compiled, and used is critical, and those responsible for using the tool must have a clear technical understanding of what lies behind the numbers and which underlying assumptions and limitations are associated with these approaches. Despite all of these caveats, alternative approaches to improved decision making leave even more to be desired.

Uses of Composite Indicators

Measures of health status that combine mortality and morbidity facilitate comparisons both within and across populations. They can be used to estimate the quantitative health benefits from interventions and serve as tools to assist in the allocation of resources. The development of such measures entails two major processes: the measurement of healthy life, including losses of time from premature mortality and disability; and the valuing of life, which incorporates issues of duration, age, extent of future life, productivity, dependency, and equity (Morrow & Bryant, 1995). The purpose of developing such measures and the need for refining them become clear if the following objectives are to be achieved:

- The use of such methods at the country level for evaluating the impact of diseases
- Their use in the allocation of resources within the health sector
- The generation of more relevant and useful data for policy makers

Understanding Summary Measures

Precursors of composite indicators have been discussed in the literature for decades and generally were developed to assist with prioritization of health issues. Usually these metrics were based on the measurement of losses of time, losses of productive time, income forgone, or other costs incurred as a result of diseases. The earlier indicators generally focused on economic losses and estimated time loss due to disease and converted these losses into a dollar value. Thus, these measures are more economic measures than disease burden measures.

Two types of composite summary measures have been developed: *health gap measures* (healthy life lost), such as healthy life years (HeaLYs) or disability-adjusted life years (DALYs), and *health expectancies*, such as disability-free life expectancy (DFLE) or health-adjusted life expectancy (HALE). Both types use healthy lifetime lost through disability and death as a common measure of the impact of mortality and nonfatal health outcomes. These two types of measures are complementary and can be studied using survivorship curves, as discussed by Murray and Lopez (1994) (**FIGURE 1-2**).

In Figure 1-2 (Murray et al., 2002), the darker line is the survivorship curve based on a standard hypothetical life table population that demonstrates the proportion (y-axis) of an initial birth cohort that remains alive at any age (x-axis). The area A + B is the total life expectancy at birth of this cohort. A part of this life is spent in full health (area A); the lighter line is the survivor curve of those persons in full health. Thus, area A represents time lived in full health, whereas area B is time lived in suboptimal health (with disability). Area C represents time lost due to mortality. The area of the complete rectangle (A + B + C) represents the ideal survivorship curve—the theoretical maximum of healthy life for a cohort who lived in full health until a maximum age when all died.

Health expectancies are summary measures that estimate expectancy of life in a defined state of health. Examples include DFLE, active life expectancy, and HALE. These indicators extend the concept of life expectancy to expectations of various states of health, not just of life per se. Health expectancies assign lower weights to life lived in less than full health on a scale of 0 to 1, in which full health is rated 1. In Figure 1-2, health expectancy is given by the following equation:

$$\text{Health expectancy} = A + f(B)$$

where f is some function that assigns weights to years lived in suboptimal health.

Health gaps are summary measures that estimate the difference between actual population health and some specified norm or goal. In Figure 1-2, that difference is indicated by area C (loss due to mortality) plus some function of area B—that is, survivorship with disability:

$$\text{Health gap (healthy life lost)} = C + g(B)$$

where g is some function that assigns weights to health states lived during time B. Weights range between 0, meaning no disability (full health), and 1, meaning complete disability (equivalent to death). Note that this measure is equivalent to healthy life lost based on the natural history of disease in a population as discussed in the section "Healthy Life Year" later in this chapter.

Although some believe that health expectancies such as the HALE indicator are more readily understood (because they are conceptual extensions of the

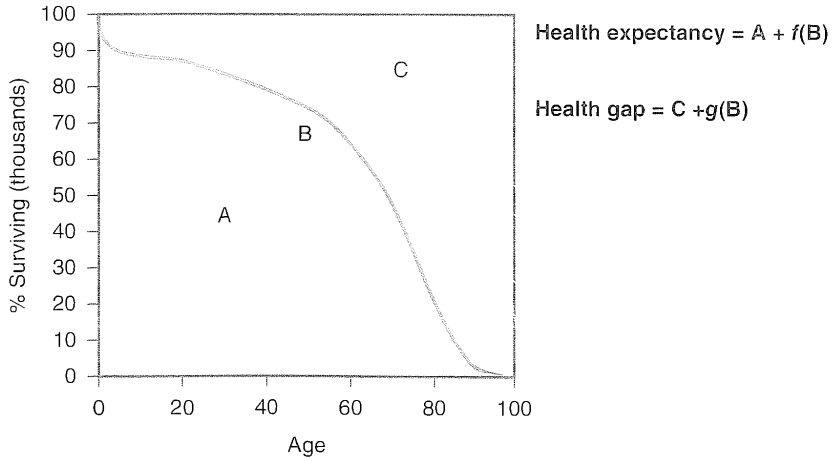


FIGURE 1-2 Survivorship curve of a hypothetical population showing health gaps and health expectancies.

Modified from Murray, C. J. L., et al. (2002). *Summary measures of population health*. Geneva, Switzerland: World Health Organization.

widely used life expectancy measure), health gap measures have important advantages for the purposes of health policy, planning, and resource allocation decisions. Both HeaLYs and DALYs are developed on the basis of disability and death attributable to a specific disease in an individual person. In their construction, great care is taken to ensure that there is *categorical attribution* using the ICD, so that each event (death or disability) is mutually exclusive and collectively exhaustive. With these measures, therefore, summing deaths and disabilities from each disease provides the total amount of death and disability for the population (a property termed *additive decomposition*). Health gap measures have this property, whereas health expectancies do not (Murray et al., 2002).

Composite Indicators

A number of composite summary indicators for burden of disease assessment have been developed. We will focus on four of these indicators: three of the health gap type (the healthy life year, the disability-adjusted life year, and the quality-adjusted life year) and one of the health expectancy type (HALE). In addition to measures of morbidity and mortality per se, these composite indicators may incorporate certain social value choices either explicitly or implicitly: the choice of life expectancy tables, valuing future life as compared with present life, valuing life lived at different ages, valuing social or economic productivity, and valuing equity in relation to cost-effectiveness. These social value choices are discussed later in this chapter (see the section "Valuing Life: Social Value Issues"), but because some social value choices are integral to the calculations of some composite indicators, they are briefly mentioned in this section.

Healthy Life Year

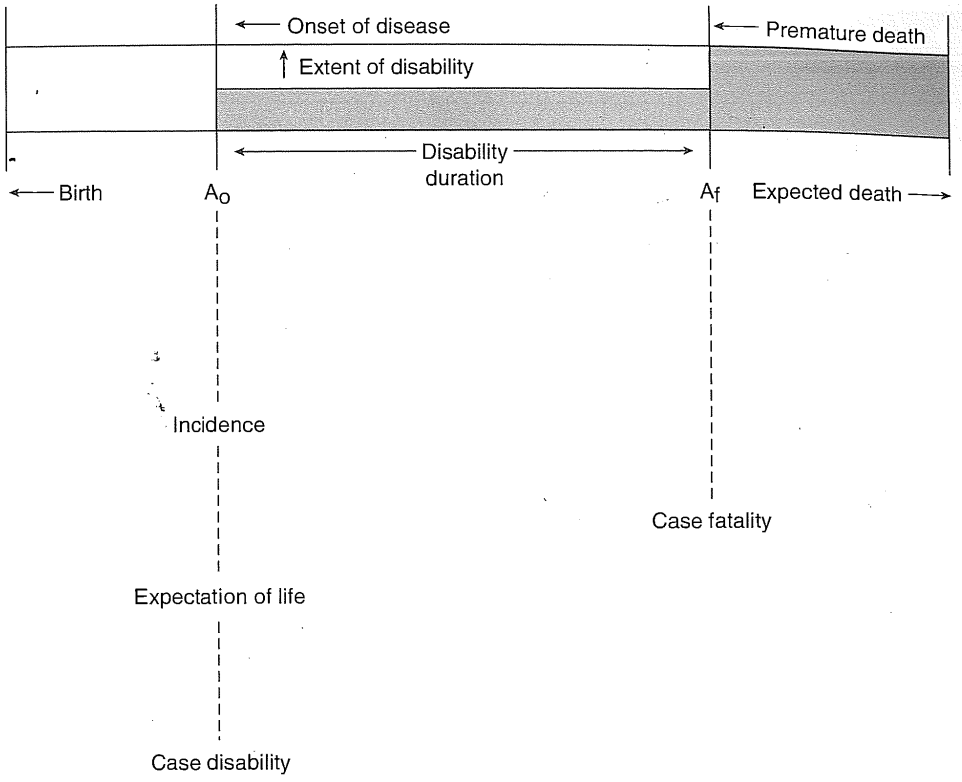
The healthy life year (HeaLY) is a composite measure that combines the amount of healthy life lost due to morbidity with that lost due to death—that is, loss of life expected had the disease not occurred (Hyder et al., 1998). We discuss the HeaLY first because it is conceptually straightforward, serves as a prototype for other health gap indicators, and was the first of the composite measures to be used as a tool in national health planning (Ghana Health Assessment Team, 1981).

Measuring the loss of healthy life from disability is more challenging than measuring the comparable loss from death, and many approaches have been used (Murray & Lopez, 1994). To incorporate loss from disability in a composite measure, such a loss must have comparable dimensions to that for life lost due to death. The HeaLY includes three components for disability: case disability ratio (comparable to the case fatality ratio), extent of disability, and duration of disability. The CDR and duration of disability can be determined objectively, but assessment of the extent of disability, which usually ranges from 0 (no disability) to 1 (equivalent to death), has a substantial subjective element (Morrow & Bryant, 1995).

The healthy life approach focuses on knowledge of the pathogenesis and natural history of disease (Last, Spasoff, & Harris, 2000) as the conceptual framework for assessing morbidity and mortality and for interpreting the effects of various interventions (**FIGURE 1-3**).

The onset of disease usually will be dated from the start of symptoms or signs, as determined by the individual afflicted, a family member, or a medical practitioner, or as the result of a lab test. Several different patterns of disease evolution are possible, of course. **FIGURE 1-4** illustrates healthy life lost from disability and premature death due to typical cases of cirrhosis, polio, and multiple sclerosis, respectively, in terms of onset, extent and duration of disability, and termination. The conclusion of the disease process depends on the natural history of the disease as modified by possible interventions. The possible outcomes include clinical recovery (the complete disappearance of clinical signs and symptoms), progression to another disease state (such as chronic hepatitis progressing to cirrhosis), and death. The last outcome includes death directly caused by the disease as well as death indirectly brought on by the disease as a result of disability.

The definitions of variables and formulas to calculate HeaLYs are provided later in this section and summarized in **TABLE 1-4**. Each disease will have a distribution of ages at which onset or death may occur, but for most diseases the average age will provide a satisfactory approximation for a population. In view of the limitations of data, this is the starting assumption for the application of the HeaLY method in LMICs. Nevertheless, as with other choices in this method, if sensitivity testing indicates that the average age is not satisfactory, then estimates may be based on age distributions. Similarly, if the natural history of a disease or response to interventions is different in different age groups, then the disease can be specifically classified by age (e.g., neonatal tetanus as compared with adult tetanus, and childhood pneumonia as compared with adult pneumonia). In recurrent diseases or diseases with multiple episodes (e.g., diarrhea), age at onset denotes the average age at first episode.



Note: A_o = average age at onset; A_f = average age at death; ▨ = healthy life lost.

FIGURE 1-3 The HeaLY model: Loss of healthy life from disability and death.

The expectation of life in HeaLYs were based on normative expectations of what should occur under usual circumstances at the time of that work. Women in Japan, who had the highest global expectation of life, approximated this norm with an expectation of life at birth of 82.5 years for females (Model Life Table West, level 26) (Coale, Demeny, & Vaughan, 1983; Coale & Guo, 1989).

The definition of disease ("dis-ease") makes the value of the case disability ratio 1 by default for most disease states, because all cases are disabled (to varying degrees and duration) if those persons have been labeled as diseased. For some conditions (e.g., sickle cell trait or HIV positivity) and risk factors, however, cases may not be considered diseased by definition, but the condition nonetheless needs to be assessed.

The duration of disability can be either temporary or permanent (lifelong). If the disability is temporary, then Dt is the duration of that disability until recovery (see Table 1-4). If the disability is permanent and the disease does not affect life expectation, then Dt is the expectation of life at age of onset of disease [Dt = E(Ao)]. If the disability is permanent and the disease reduces life expectation, then Dt is the expectation of life at age of onset reduced by the difference between ages of fatality and onset [Dt = E(Ao) − (Af − Ao)]. A disability severity scale needs to be used to estimate extent (severity) of the disability (see Table 1-4).

The HeaLYs lost from death and from disability are added and expressed as the total years of life lost per 1,000 population per year. The loss is attributed to the year in which disease onset occurs and includes the stream of life lost from disability and death at any time after onset, even if these events happen many years later. This method offers a prospective view of the event (disease onset) and its natural history (or as modified by interventions) over time.

An important benefit of the HeaLY formulation is that the effects of different kinds of interventions can be readily explored to determine their expected gains in terms of healthy life. The HeaLY spreadsheet (available upon request from the authors) incorporates these concerns; it also includes options for considering the proportion of the population that will be covered by an intervention and allows for different levels of coverage for different segments of the population for each intervention.

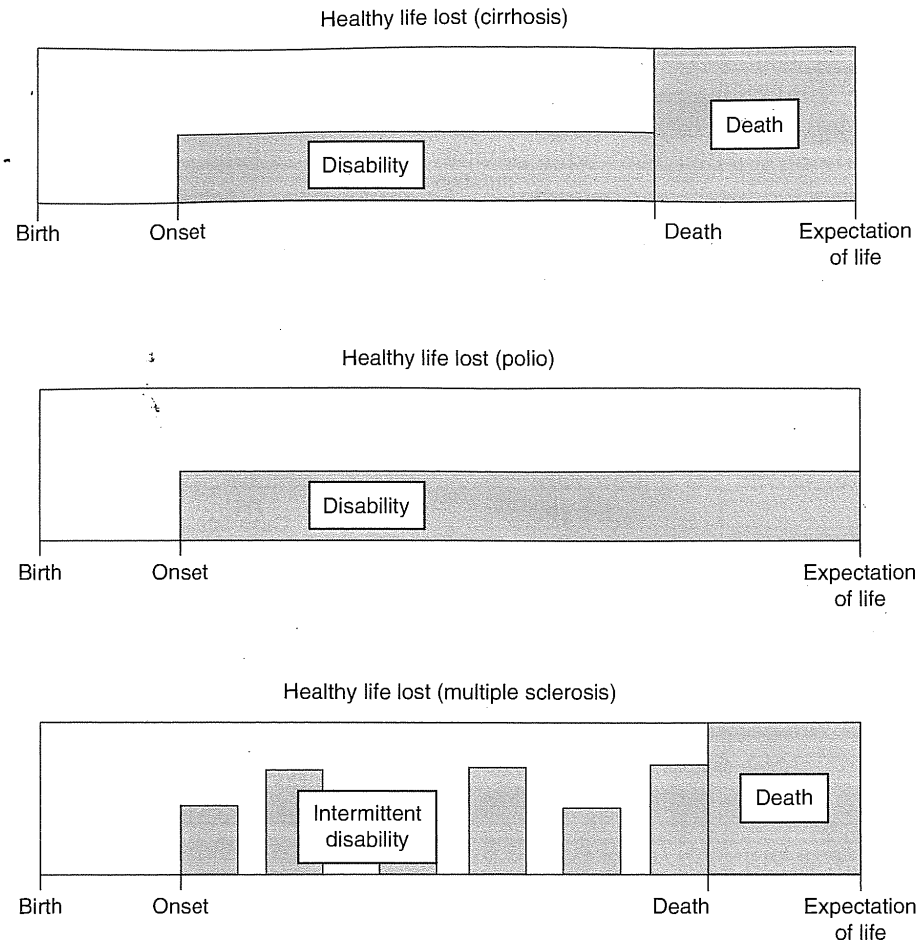


FIGURE 1-4 Different patterns of healthy life lost.

Reproduced from Hyder A., Rollant G., and Morrow R. H. (1998). Measuring the burden of disease: Healthy life-years. *American Journal of Public Health, 88*(2), 196–202. Figure 1, p. 197.

Disability-Adjusted Life Year

The disability-adjusted life year (DALY) is a health gap population summary measure that combines time lost due to disability with that lost due to death (life that would have been expected had the disease not occurred), in a manner similar to the HeaLY measure. It first appeared in the World Development Report of 1993 and has become the most widely used composite measure of population health (Global Burden of Disease Risk Factors Collaborators, 2016; Global Burden of Disease Risk Factors Collaborators et al., 2015; Jamison et al., 2006; Lim et al., 2012; Lopez et al., 2002; Murray & Lopez, 1994; Murray et al., 2002).

DALYs are calculated as two separate components for the measurement of life lost due to disease, and they may also directly include three social value choices. The two components are (1) years of life lost (YLL), referring to the loss of healthy life from death, and (2) years of life lived with disability (YLD), referring to the loss of healthy life from disability. Thus

$$\text{DALY} = \text{YLL} + \text{YLD}$$

The social value choices that may be included in DALYs are (1) life expectation values, (2) discount rates for future life, and (3) weighting for life lived at different ages, as discussed later.

Since the GBD 2010 study (Lozano et al., 2012), age weighting is no longer a default value choice for the DALY. Instead, users have the option to calculate these values with or without age weighting, as well as with or without discounting. The following three options could be applied: (1) use both age weights and discounting, (2) use either age weights or discounting, and (3) use neither age weights nor discounting.

The calculation for YLL in a population uses the age distribution of all deaths by cause in one year multiplied by life expectation at each age to estimate the loss of life for each disease that would have been expected if not for that disease. The expectation of life can be obtained either from a model life table (Coale & Guo, 1989) or based on the best achievable low levels of mortality such as those found in Japan; thus the DALY, as does the HeaLY, directly incorporates this social value choice.

TABLE 1-4 Variables for Estimating Healthy Life Years (HeaLYs)

Symbol	Explanation	Expression
I	Incidence rate per 1,000 population per year	/1,000/year
Ao	Average age at onset	years
Af	Average age at death	years
E(Ao)	Expectation of life at age of onset	years
E(Af)	Expectation of life at age of death	years
CFR	Case fatality ratio: proportion of those developing the disease who die from the disease	0.00–1.00
CDR	Case disability ratio: proportion of those developing the disease who have disability from the disease	0.00–1.00
De	Extent of disability (from none to complete disability equivalent to death)	0.00–1.00
Dt	Duration of disability in years	years
	Disability can be either permanent or temporary: ▪ If temporary, then Dt = duration of that disability (i.e., until recovery or death) ▪ If permanent and disease does not affect life expectation, then Dt = E(Ao) ▪ If permanent and the disease does reduce life expectation, then Dt = Af − Ao	
HeaLY	Healthy life years lost per 1,000 population per year: $I \times \{[CFR \times \{E(Ao) - [Af - Ao]\}] + [CDR \times De \times Dt]\}$	HeaLYs per 1,000 per year

For disability, the DALY uses estimates of incidence, duration, and severity to calculate the time lived with disability (YLD) for each disease. The YLD component equals the number of incident cases in the period multiplied by the average duration of disease multiplied by a weight factor for the degree of severity (extent) of the disease. A description of the severity scale used in one version of DALY was given earlier in this chapter, in the section on measurement of disability (see Table 1-4).

The second social value choice, which was directly incorporated in the original version of DALY, is the discount rate of 3% per annum. This social time preference has been used for most estimates; recently, DALY results discounted at 0% have also become available.

The third social value choice concerns weighting life lived at different ages. Earlier DALYs were age weighted according to an arbitrary exponential curve designed to give the most value to life lived as a young adult (Hyder et al., 1998; World Bank, 1993). Weighting by age was the most controversial component of the DALYs when they appeared and caused great dissent from other health professionals (see the section "Valuing Life Lived at Different Ages" later in this chapter). Recent DALY listings from GBD studies also include results with no age weighting (all years equally valued). It has been argued that age weighting

of DALYs does not affect final results, but this depends on the purpose for making the estimates and has been challenged (Anand & Hanson, 1997; Barendregt, Bonneux, & Van der Maas, 1996; Barker & Green 1996; Hyder et al., 1998).

An important difference between the HeaLY and the DALY is the fact that the starting point for the HeaLY is the onset of disease; that is, the loss of healthy life is based on the natural history of the disease (as modified by interventions), illustrated in Figures 1-2 and 1-3. This is true for the YLD component of the DALY, but the YLL is based on mortality in the current year. In a steady state, there is no difference in these perspectives. However, when incidence is changing—such as with HIV in many parts of the globe—the DALY approach can potentially understate the true situation (Hyder & Morrow, 1999).

The calculation for DALYs can be expressed in the form of an integral that was first published in the World Bank literature (Murray & Lopez, 1994). This single equation incorporating all technical and value choices had the advantage of standardization to ensure comparability of the multiple calculations undertaken in the GBD studies, and greatly facilitated the actual computations. Nevertheless, for national and local priority setting, it may be preferable to use an indicator constructed such that the social value choices can be adjusted to suit national and local preferences (Bobadilla, 1998; Hyder et al., 1998; Morrow & Bryant, 1995). Recent DALY formulations allow for this possibility; indeed, it is useful to think of DALYs as a family of related measures using terminology specifying the following formulation: DALY (r, K) uses a discount rate of r and age weighting indexed to K. Other parameters can be added in a similar fashion (Jamison et al., 2006).

HeaLYs and DALYs are both "health gap" measures and can be considered part of the same family of measures. In fact, DALYS exactly equal HeaLYs when the following conditions are met: (1) the condition in question is in steady state or equilibrium (that is, the incidence, CFR, and disability variables remain constant during the time intervals under consideration); (2) age weighting is not applied ($K = 0$); and (3) the same measures of disability (weights) are used.

The health status & the life extend are extremly relative. so it's important in cost utility analysis ⟹ to guide the resources to best alternative for patients

Quality-Adjusted Life Year

The quality-adjusted life year (QALY) was introduced in 1976 to provide a guiding principle for selecting among alternative tertiary healthcare interventions (Zeckhauser & Shepard, 1976). The idea was to develop a single measure of quality of life that would enable investigators to compare expected outcomes from different interventions—a measure that valued possible health states both for their quality of life and for their duration.

The central notion behind the QALY is that a year of life spent in one health state may be preferred to a year spent in another health state. This generic measure sums time spent in different health states using weights on a scale of 0.00 (dead) to 1.00 (perfectly healthy) for each health state; it is the arithmetic product of duration of life and a measure of quality of life (health state weight). For example, 5 years of perfect health = 5 QALYs; 2 years in a state measured as 0.5 of perfect health followed by 5 years of perfect health = 6 QALYs.

The QALY was originally developed as a differentiating indicator for individual choices among tertiary healthcare procedures, not as a measure of disease burden in a population. It was used to assess individual preferences for different health outcomes from alternative interventions (Morrow & Bryant, 1995). The QALY, too, comprises a large family of measures. Since its introduction, a wide variety of QALY measures have been developed, along with a voluminous literature on alternative methods incorporating a range of disability domains and a diversity of methods to assign weights to generate QALYs (Kaplan & Anderson, 1996; Nord, 1993). The most widely used measure is the EQ-5D (European Quality of Life with Five Domains and three levels of quality for each domain; www.euroqol.org).

Perhaps the most important use of QALYs has been as a common denominator to measure utility in cost-utility analysis (and effectiveness in cost-effectiveness analysis) to assist in resource allocation among alternative health interventions by ranking interventions in terms of cost per QALY (Kaplan & Anderson, 1996; Nord, 1992; Torrance, 1986). An early and widely publicized attempt to make the best use of healthcare resources by maximizing QALYs per dollar expended was the well-intentioned but rather unfortunate effort undertaken in Oregon in the early 1990s (**EXHIBIT 1-4**).

In the United Kingdom, as part of its 1997 National Health Service (NHS) reforms, the National Institute for Clinical Excellence (NICE; www.nice.org.uk) was created to advise public health officials about the effectiveness and cost-effectiveness of various health interventions. In an explicit attempt to introduce economic considerations in addition to medical judgments for the allocation of resources, NICE has produced a large collection of studies on the cost per QALY produced by the interventions it appraises. Some of these appraisals have been the source of considerable controversy. If a treatment is considered cost-effective for a group of patients,

EXHIBIT 1-4 Oregon: Historical Application of the QALY for Allocation of Resources

An early and well-known attempt to apply the QALY approach for allocation of health resources occurred in the state of Oregon (Blumstein, 1997). In 1988, Oregon faced a budgetary shortfall for its Medicaid program, and coverage for organ transplants was denied. In an effort to prioritize its health services, Oregon undertook a bold attempt to explicitly ration health services. A coalition including consumers, healthcare providers, insurers, business, and labor representatives launched a broad and courageous healthcare reform. It began with a series of "experiments" in which the decision-making process was based on a cost-effectiveness approach using quality of well-being (QWB—essentially a QALY) for comparing the outcomes of treatment options among people.

The initial list, published in 1990, consisted of 1600 condition/treatment pairs drawn up as follows:

Cost-effectiveness ratio = cost of services / (health gain × duration)
Cost of services = charges for treatment including all services and drugs
Quality of well-being (QWB) = sum of QWB weight (W) × each QWB state × probability that symptoms of that QWB state would occur
Health gain = QWB with treatment − QWB without treatment

From the beginning, there was great opposition to the very notion of rationing; consequent denial of services to those who had conditions that did not make the list contributed to the rancor. There were also unfortunate technical blunders in the generation of the first list. For example, treatment for thumb sucking was ranked higher than hospitalization for starvation, and treatment for crooked teeth higher than early treatment for Hodgkin's disease. Such inconsistencies, together with objections raised by groups advocating for the disabled, gave rise to alternative approaches for establishing rankings.

Although enormous public effort went into the reform and much was accomplished, the explicit cost-effectiveness approach with QALYs as the outcome measure was eventually dropped (Blumstein, 1997; Eddy, 1991; Morrow & Bryant, 1995; Nord, 1993).

NICE will recommend its use throughout the NHS; if not, it will recommend against its use in the NHS. The use of these cost-effectiveness studies as an aid to decision making is intended to increase the total healthcare benefits gained from the money spent by the NHS.

The QALY as originally used is essentially equivalent to the YLD of the DALY; in fact, it would be exactly the same as the YLD when the following conditions are met: (1) there is no discounting ($r = 0$); (2) there is no age weighting ($K = 0$); and (3) the same disability weights are used. More recently (as used in some cost-effectiveness studies), QALYs have incorporated life expectation as well.

Health-Adjusted Life Expectancies

Several types of health expectancies exist in the literature. During the 1990s, *disability-free life expectancy* (DFLE) and related measures were calculated for many countries (Robine, 1994; Mathers et al., 2001). However, these measures incorporate a dichotomous weighting scheme in which time spent in any health state categorized as disabled is assigned, arbitrarily, a weight of zero (equivalent to death). Thus, DFLE is not sensitive to differences in the severity distribution of disability in populations. In contrast, *disability-adjusted life expectancy* (DALE) adds up expectation of life for different health states with adjustment for severity weights. In 2001, WHO replaced the DALE terminology with *health-adjusted life expectancy* (HALE); the latter term will be used throughout the remainder of this text.

The HALE is a composite summary measure of population health status that belongs to the family of health expectancies; it summarizes the expected number of years to be lived in what might be termed the equivalent of "full health." WHO has used it as the measure of the average level of health of the populations of member states for annual reporting on population health for a few years (WHO, 2000).

Health expectancy indices combine the mortality experience of a population with the disability experience. The HALE is calculated using the prevalence of disability at each age so as to divide the years of life expected at each age (according to a life table cohort) into years with and without disability. Mortality is captured by using a life table method, while the disability component is expressed by additions of prevalence of various disabilities within the life table. This indicator allows an assessment of the proportion of life spent in disabled states. When compared with the total expectation of life, it translates into a measure of the total disability burden in a population.

Comparison of the various methods and specific indicators is available in the literature (Robine, 1994). Alternative methods are given in WHO's *National Burden of Disease Studies* manual (Mathers et al., 2001a). As originally designed, the HALE does not relate to