CHAPTER

1

Measures of Health and Disease in Populations

ADNAN A. HYDER, PRASANTHI PUVANACHANDRA, AND RICHARD H. MORROW

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 book: *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.

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.

Recent developments in the measurement of population health status and disease burden 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 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 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 (Figure 1-1). 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 health systems project in Tanzania (Exhibit 1-1). This 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).

Counting Disease

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 common 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, 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 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 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.

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 and the International Development Research Centre (IDRC), 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 must 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. The organization found that the amount 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 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.

Although this is a dramatic example of how data can be used to recognize and correct misplaced health resource expenditures, it should be emphasized that health system expenditures should be equitably distributed based on which intervention programs maximize healthy life gains, not according to disease problems per se.

For additional information on the TEHIP success story, visit the following websites: http://www.odi.org.uk/Rapid/Tools/Case_studies/TEHIP.html and http://www.idrc.ca/tehip.

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 Graunt 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.

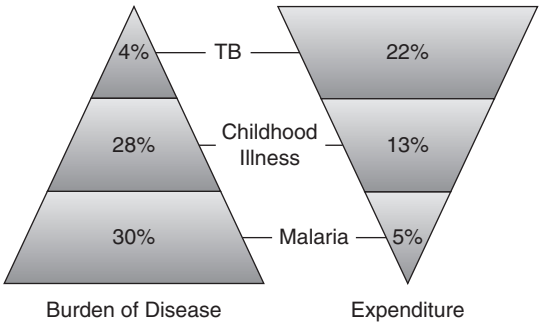


Figure 1-1 Using evidence to improve a health system: an example from Africa.

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, their reliability and completeness have been steadily improving and often are fairly satisfactory. In many low-income countries, however, the collection of vital statistics remains grossly incomplete. 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 provide a complete picture of global mortality (Murray et al., 2001). 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 five years and other populations.

Obtaining information about cause of death remains difficult even in many middle-income countries; most 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 five. 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 quite 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 et al., 2009).

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, less has been done for adult mortality (Hill, 2003). Developing countries have higher rates of age-specific adult mortality than do high-income nations (Lopez et al., 2002; Murray & Chen, 1992). Indeed, mortality rates are higher for both women and men 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 immune deficiency 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. 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 (United Nations Children's Fund, 2009). 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, 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 low child mortality rates. These examples demonstrate that the relationship between mortality and poverty is complex and needs in-depth investigation.

There continue to be major deficiencies in cause-specific mortality data in low- and most middle-income countries. In keeping with demographic and epidemiologic transitions (see Exhibit 1-3, later in this chapter), 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 is a major reason for the 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, 1992). 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

Exhibit 1-2 **Trends of the HIV/AIDS Epidemic**

Acquired immune deficiency syndrome (AIDS) is the leading infectious cause of adult death in the world. Untreated disease caused by the human immunodeficiency virus (HIV) has a case fatality rate that approaches 100% (WHO, 2003). Unknown 30 years ago, this disease has already killed more than 25 million people, and an estimated 31 to 36 million others are living with HIV/AIDS (UNAIDS, 2008, 2009). The most heavily burdened continent is Africa, home to two-thirds of the world's people living with HIV/AIDS. The prevalence of this disease is rising most rapidly in eastern Europe and Central Asia (e.g., Estonia, Latvia, Ukraine, and the Russian Federation) and in other parts of Asia (Indonesia, Pakistan, and Vietnam) (See Table 1-1) (UNAIDS, 2008, 2009).

Of the leading causes of disease burden among men and women of all ages, HIV/AIDS is the fifth cause, accounting for 4% of the global burden of disease. In terms of mortality, it is the sixth leading cause of death among people of all ages, accounting for 3.5% of all deaths (WHO, 2004). Nearly 72% of the two million global deaths from HIV/AIDS have occurred in sub-Saharan Africa (See Figure 1-2) (UNAIDS, 2009).

Table 1-1 **Global Summary of HIV and AIDS Epidemic**

Number of people living with HIV	**Total**	**33.4 million**	**(31.1–35.8 million)**
	Adults	31.3 million	(29.2–33.7 million)
	Women	15.7 million	(14.2–17.2 million)
	Children	2.1 million	(1.2–2.9 million)
Number newly infected with HIV	**Total**	**2.7 million**	**(2.4–3.0 million)**
	Adults	2.3 million	(2.0–2.5 million)
	Children	430,000	(240,000–610,000)
AIDS deaths in 2008	**Total**	**2.0 million**	**(1.7–2.4 million)**
	Adults	1.7 million	(1.4–2.1 million)
	Children	280,000	(150,000–410,000)

Source: UNAIDS, 2009.

Figure 1-2 Trends of the HIV/AIDS epidemic. *Source:* Based on data from Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, *World Population Prospects: The 2008 Revision*, http://esa.un.org/unpp.

and feedback provided. There have been steady improvements in many countries, however, and these kinds of data provide some of the best information available on major causes of mortality.

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 five (see Table 1-1), 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 poor countries 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 Chapter 3. 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-five 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 (see Exhibit 1-3).

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-five 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 of maternal mortality occur in a different time frame from those of under-five 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 (see Chapter 3).

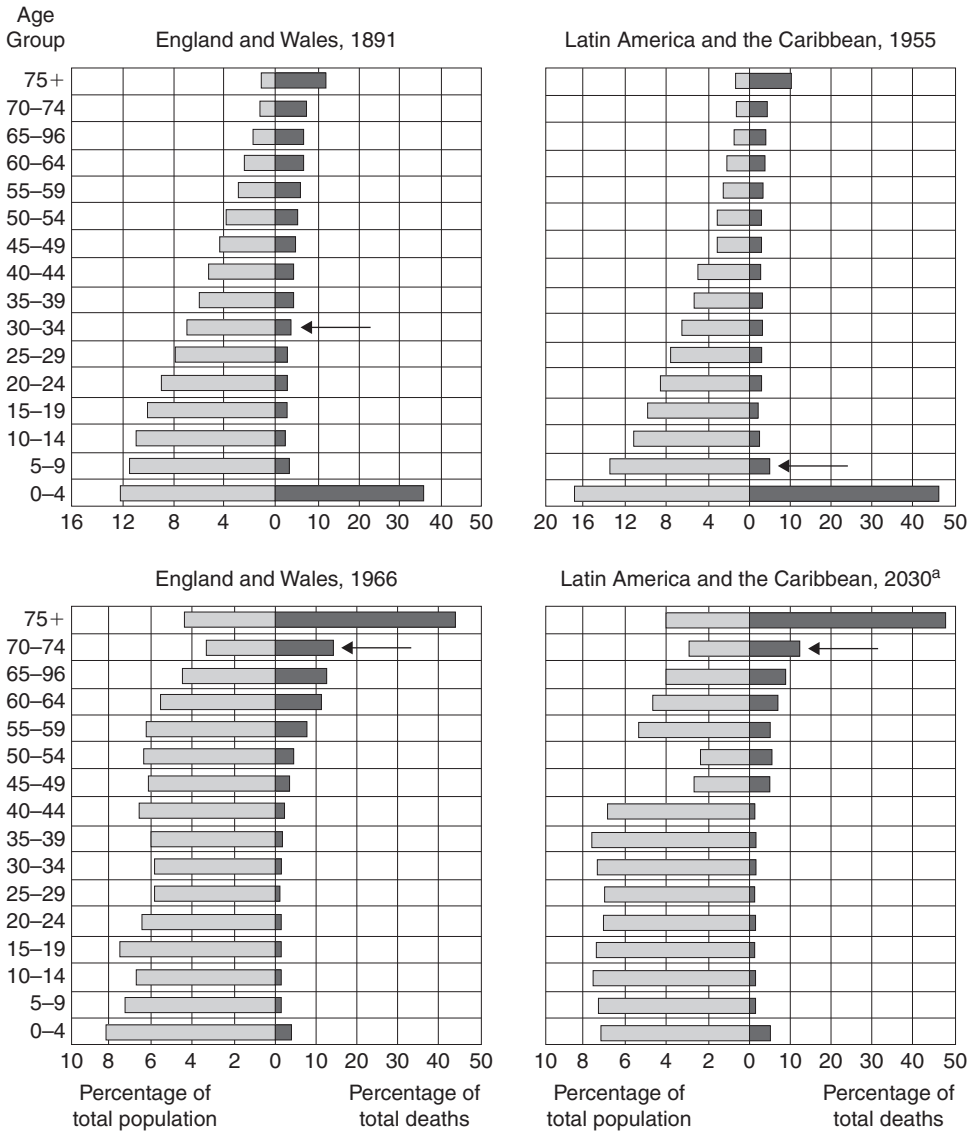
Major changes in the patterns and causes of injury are also likely to occur with modernization. For example, road traffic injuries tend to increase as countries go through the stage of development in which there is a great increase in vehicles and in the speeds at which they are operated before improved roads and law enforcement are in place (Crooper & Kopits, 2003). 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

Exhibit 1-3 The Demographic Transition

The graphs that follow show the shape of the age structure of the population on the left and the percentage of total deaths attributable to each age group on the right. The age structure for both England and Wales and for Latin America and the Caribbean shifted from a broad base and narrow top to a fairly uniform rectangular shape. At the same time there was a marked shift in the percentage of deaths by age, from children younger than age five to the elderly.



[a]Projected
[b]The age below which half of all deaths in a year occur.

Table 1-2 Health-Related Metrics

Type	Indicator	Definition/Interpretation
Demographic indicators: reproductive health (see Chapter 4)	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 age 15–49)
	Lifetime risk of maternal mortality	
	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 (see Chapter 6)	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-five 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
	Perinatal mortality rate	Number of fetal deaths (28 or more weeks of gestation) + post natal 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

health-related indicators are useful for specific purposes. Many are discussed more fully in other chapters of this book; they are summarized in Table 1-2. Those related to the Millennium Development Goals (MDGs) are discussed in Exhibit 1-4.

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, 1990). 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.

Exhibit 1-4 Millennium Development Goals

In 2001, UN member states adopted eight Millennium Development Goals (MDGs) to spur social and economic development in the world's poorest countries:

- Goal 1: Eradicate extreme poverty and hunger
- Goal 2: Achieve universal primary education
- Goal 3: Promote gender equality and empower women
- Goal 4: Reduce child mortality
- Goal 5: Improve maternal health
- Goal 6: Combat HIV/AIDS, malaria, and other diseases
- Goal 7: Ensure environmental sustainability
- Goal 8: Develop a global partnership for development

The 8 MDGs were divided into 21 quantifiable targets that are measured by 60 indicators. Of the 21 targets, eight are directly related to health. Of the 60 indicators, 22 are directly 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. For examples, the following is a list of MDG indicators:

Target 1c: Reduce by half the proportion of people who suffer from hunger

- 1.8 Prevalence of underweight children younger than five years of age
- 1.9 Proportion of population below minimum level of dietary energy consumption

Target 4a: Reduce by two-thirds the mortality rate among children younger than five

- 4.1 Under-five mortality rate
- 4.2 Infant mortality rate
- 4.3 Proportion of one-year-old children immunized against measles

Target 5a: Reduce by three-fourths the maternal mortality ratio

- 5.1 Maternal mortality ratio
- 5.2 Proportion of births attended by skilled health personnel

Target 5b: Achieve, by 2015, universal access to reproductive health

- 5.3 Contraceptive prevalence rate
- 5.4 Adolescent birth rate
- 5.5 Antenatal care coverage (at least one visit and at least four visits)
- 5.6 Unmet need for family planning

Target 6a: Halt and begin to reverse the spread of HIV/AIDS

- 6.1 HIV prevalence among population aged 15–24 years
- 6.2 Condom use at last high-risk sex
- 6.3 Proportion of population aged 15–24 years with correct knowledge of HIV/AIDS
- 6.4 Ratio of school attendance of orphans to school attendance of non-orphans aged 10–14

Target 6b: Achieve, by 2010, universal access to treatment for HIV/AIDS for all who need it

- 6.5 Proportion of population with advanced HIV with access to antiretroviral drugs

Target 6c: Halt and begin to reverse the incidence of malaria and other major diseases

- 6.6 Incidence and death rates associated with malaria
- 6.7 Proportion of children younger than five sleeping under insecticide-treated bed nets
- 6.8 Proportion of children younger than five with fever who are treated with appropriate antimalarial drugs
- 6.9 Incidence, prevalence, and death rates associated with tuberculosis
- 6.10 Proportion of tuberculosis cases detected and cured under directly observed treatment short course

Target 7c: Reduce by half the proportion of people without sustainable access to safe drinking water and basic sanitation

- 7.8 Proportion of population using an improved drinking water source
- 7.9 Proportion of population using an improved sanitation facility

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 society, and to account for that disease's heterogeneity of its clinical expression and evolution in different individuals and societies. ICIDH categories included *impairment* (loss or abnormality of psychological, physiological, or anatomical 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 and 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, 2002a), in which 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 occurs 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. The ICF 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 for 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 alternate 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 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, 1990). 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 is 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 low- and middle-income as compared with high-income nations. This situation also underscores the variation in morbidity data, which are often interpreted as indicating that wealthy individuals and low-mortality populations report higher rates of morbidity (Lopez et al., 2006).

Measurement of individual preferences for different health states to determine relative severity of disability has been done by a variety of methods (Kaplan, 1990; Murray et al., 2002; Torrence, 1986). Factors that influence the assessment of such preferences in-

clude 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 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 (Lopez et al., 2006; Torrence, 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.

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 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, there have been concerns about their reliability and validity. These measures are not discussed further in this book, because they have been primarily used in clinical assessments of individuals and do not directly relate to measures of population health.

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 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, usually 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 et al., 2002). For example, severity of disability scales have been developed by group consensus using community surveys (Kaplan, 1990), a mixture of community and expert groups (Ghana Health Assessment Team, 1981), experts only (World Bank, 1993), and population surveys (Murray et al., 2002). These scales usually compare perfect health states to death on a scale of 0 to 1 (Table 1-3).

In the 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 (see Table 1-3). Outcomes from all other health conditions were categorized within these seven classes (with special categories for treated and untreated groups). 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 decisions ought to be a national or local decision.

The third component of disability is its *duration*. The duration is generally counted from onset until

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 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, 1990		
Disability Class	**Severity Weight**	**Indicator Conditions**
1	0.00–0.02	Vitiligo, height, weight
2	0.02–0.12	Acute watery diarrhea, sore throat, severe anemia
3	0.12–0.24	Radius fracture, infertility, erectile dysfunction, rheumatoid arthritis, angina
4	0.24–0.36	Below-knee amputation, deafness
5	0.04–0.50	Rectovaginal fistula, major mental retardation, Down syndrome
6	0.50–0.70	Major depression, blindness, paraplegia
7	0.70–1.00	Psychosis, dementia, migraine, quadriplegia

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 of 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 further precision need to be justified by its potential impact on decision making. Low- and middle-income countries, 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, makes explicit the value choices that each entails, 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 be made (Hyder et al., 1998; World Bank, 1993). The global scarcity of resources for health care is a challenge for every country, rich and poor (Evans et al., 1981; World Bank, 1993), but the realities in low-

and middle-income countries make the issue of choice that much starker. It is even more important for poor countries 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 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 substantiality 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 is 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 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 (HeaLY) and disability-adjusted life years (DALY), and *health expectancies,* such as disability-free life expectancy (DFLE) and

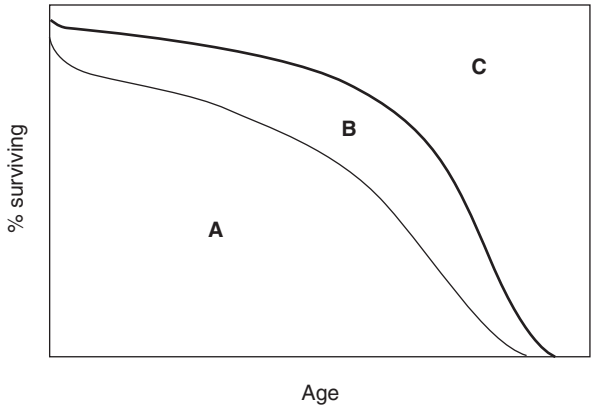


Figure 1-4 Survivorship curve of a hypothetical population showing the areas of health expectancies.
Source: C.J.L. Murray et al. Summary Measures of Population Health (Geneva, Switzerland, WHO, 1999).

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 (Figure 1-4).

In Figure 1-4, the bold 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 thin 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 1-4, 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 1-4, 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 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 classification of disease 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.

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 healthy life year 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). The HeaLY approach is a direct derivative of the work done in Ghana that incorporates several additional features.

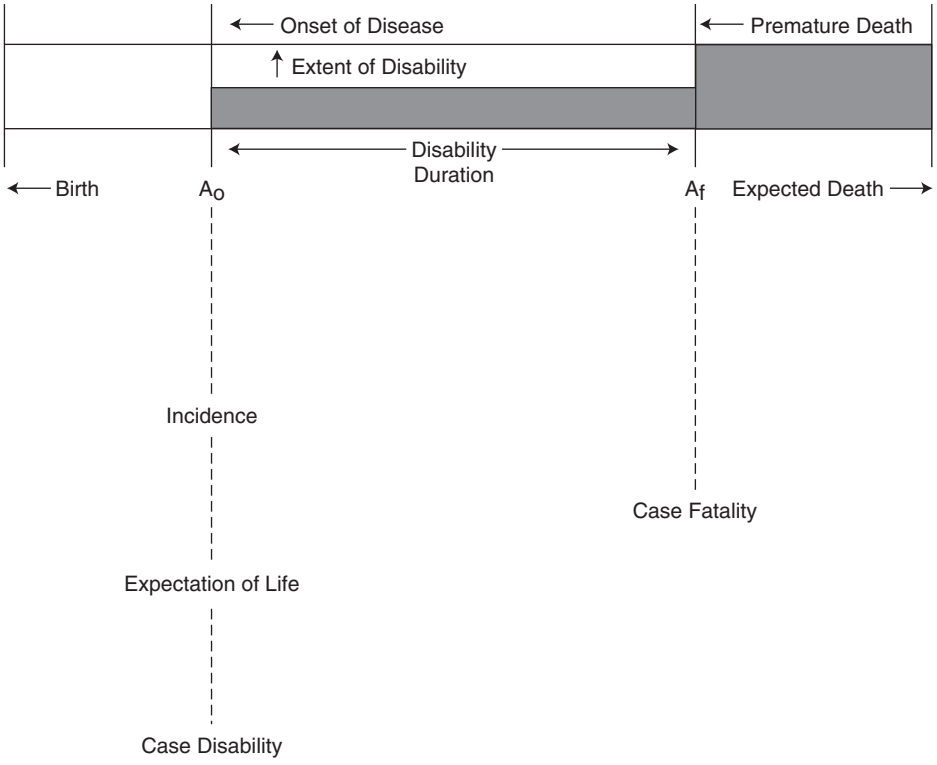
The measure of loss from death is based on the years of life that would have been lived had the disease not occurred. The information needed in addition to the incidence rate and case fatality ratio is the age of disease onset, the age of death, and the expectation of life at these ages. All of this information is objective in nature and potentially available in every country. The main issue centers on the choice for life expectation (see also the section "Expectation of Life" later in this chapter). The original Ghana work was based on expectation-of-life tables specific to Ghana. In later work, considerations of global equity and comparability across countries made it preferable to use the best possible life expectation—that of the female population in Japan.

Measuring the loss of healthy life from disability is more challenging than measuring that 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 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, 2000) as the conceptual framework for assessing morbidity and mortality and for interpreting the effects of various interventions (Figure 1-5). For the purpose of estimating healthy life lost or gained, disease is defined as stated earlier in this chapter: anything that an individual (or population) experiences that causes, literally, "dis-ease"—anything that leads to discomfort, pain, distress, disability of any kind, or death, including injuries and psychiatric disabilities. With some exceptions, those persons with infection or some biological characteristic (such as sickle-cell trait) are considered healthy unless they have specific identifiable symptoms or signs. Preclinical or subclinical disease is not generally counted. However, the diagnostic criteria for some conditions such as hypertension, HIV infection, or onchocerciasis (diagnosed by skin snip) include individuals without signs or symptoms. Such criteria (e.g., indicators of infection, high blood pressure, or genetic markers) are appropriate when they serve as the basis for intervention programs. Interventions may also be directed at reducing identifiable risk factors, such as tobacco smoking or risky sexual behavior. To the extent that risk reduction can be translated into disease reduction, the approach to measuring the benefits and costs of a risk reduction intervention program remains the same as that for disease reduction.

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 1-6 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 developing countries. Nevertheless, as with other



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

Figure 1-5 The HeaLY Model: Loss of Healthy Life from Disability and Death.

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. For some diseases, such as malaria, which is characterized by recurrent episodes, and schistosomiasis, in which reinfection occurs at frequent intervals, it may be useful to view them as single lifetime diseases. For example, malaria in Africa may be considered for each individual as a single, lifelong disease with chronic, usually asymptomatic, parasitemia but with intermittent severe clinical attacks (which result in high mortality in late infancy and early childhood while immunity is being acquired), followed by recurring, nonfatal clinical episodes after age 10.

The expectation of life in HeaLYs is based on normative expectations of what should occur under usual circumstances. Women in Japan, who have the highest global expectation of life, approximate this norm with an expectation of life at birth of 82.5 years for females (model life table west, level 26) (Coale & Demeney, 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 D_t 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 D_t is the expectation of life at age of onset of disease [$D_t = E(A_o)$]. If the disability is permanent and the disease

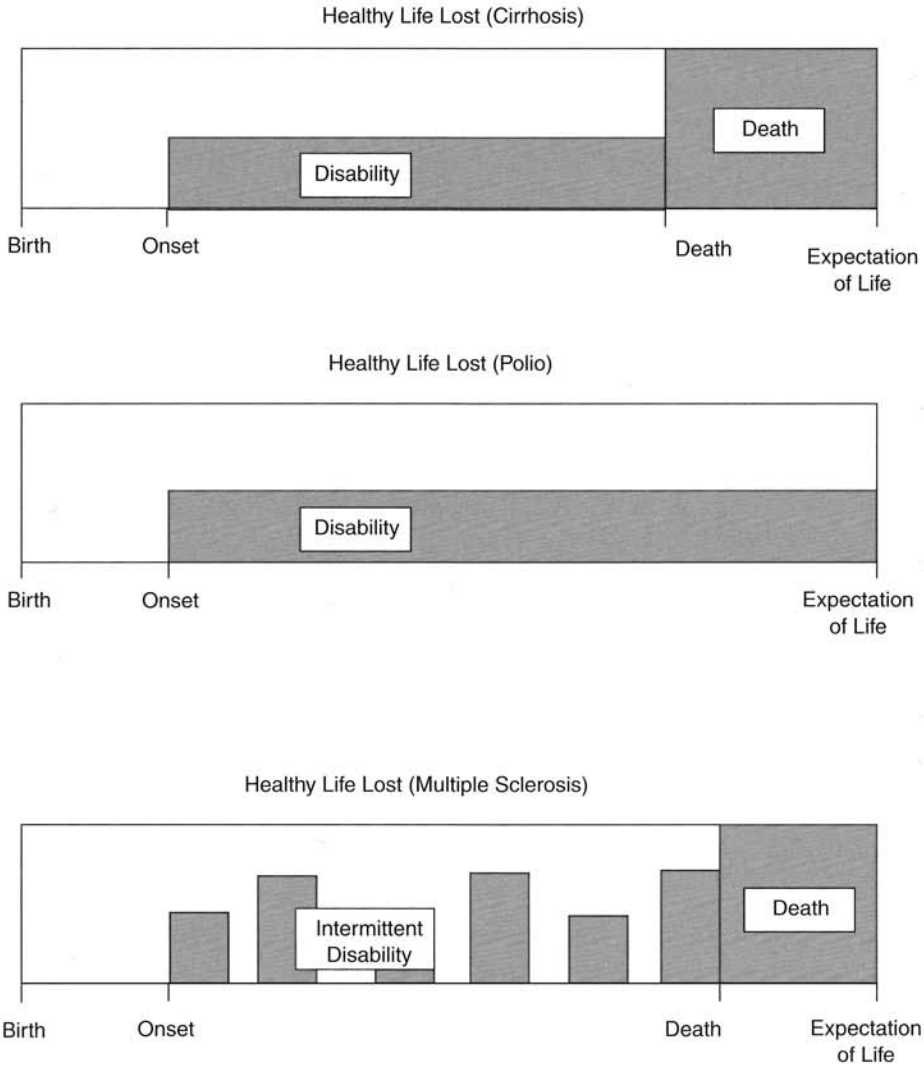


Figure 1-6 Different Patterns of Healthy Life Lost.
Source: Hyder A, Rollant G, Morrow RH. Measuring the Burden of Disease: healthy life-years. AM J Public Health, Feb. 1998; 88: 196–202. Figure 1 p. 197.

does reduce life expectation, then D_t is the expectation of life at age of onset reduced by the difference between ages of fatality and onset [$D_t = E(A_o) - (A_f - A_o)$].

A disability severity scale needs to be used to estimate extent (severity) of the disability (see Table 1-4).

The healthy life years 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.

The health status of a population can be considered as the amount of healthy life it achieves as a proportion of the total amount that the people could achieve under optimal conditions. A cohort of 1,000 newborns with an expectation of life of 82.5 years has the potential of 82,500 years of healthy life, for example. In a steady state, a random sample of 1,000 people from a population made up of successive such cohorts has the potential of 41,250 years of healthy life (Hyder et al., 1998; Morrow & Bryant, 1995). Each year this population would experience events

Table 1-4 Variables for Estimating Healthy Life Years (HeaLY)

Symbol	Explanation	Expression
I	Incidence rate per 1,000 population per year	/1,000/year
A_o	Average age at onset.	years
A_f	Average age at death.	years
$E(A_o)$	Expectation of life at age of onset.	years
$E(A_f)$	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
D_e	Extent of disability (from none to complete disability equivalent to death)	0.00–1.00
D_t	Duration of disability in years.	years
	Disability can be either permanent or temporary. • If temporary, then D_t = duration of that disability(i.e., until recovery or death) • If permanent and disease does not affect life expectation, then $D_t = E(A_o)$ • If permanent and the disease does reduce life expectation, then $D_t = A_f - A_o$	
HeaLY	Healthy life years lost per 1,000 population per year: $I \times \{[CFR \times \{E(A_o) - [A_f - A_o]\}] + [CDR \times D_e \times D_t]\}$	HeaLYs per 1,000 per year

leading to 1,000 years of healthy life lost attributable to mortality, with a distribution of age at death equivalent to that which leads to a life expectation of 82.5. Any disease that leads to disability or to death earlier than that set by this age-at-death distribution would increase the amount of healthy life lost beyond this minimum. This formulation is equivalent to the health gap, as indicated in Figure 1-4. Discounting future life or adding productivity, dependency, or age weighting would modify these denominator numbers.

HeaLYs measure the gap or loss between the current situation in a country as compared to that of an ideal or standard population. In recent work, researchers have used a standard based on the life expectation approximated in Japan. Thus, if exactly the same method were used to estimate the HeaLY losses for females in Japan, they would amount to 0 per 1,000 people for loss due to mortality; only those losses due to disability would be counted. Because the population under study is the ideal (standard), assuming stability of the population with constancy of mortality rates and no disability, there would be no gap to measure. This does not mean that the population is not having a loss of healthy life, but simply that such loss is the minimum as defined by the structure of the population and the expectation of life, as described previously. Any country that is experiencing losses greater than this minimum, either as a result of excess mortality or disability, will have a gap that can be measured; that gap is what the HeaLYs register.

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. Interventions may usefully be divided into two broad categories: those that prevent the initiation of the disease process and those that treat a disease process already under way. Some interventions fall into both categories. The primary effect of preventive strategies is to reduce the incidence of new cases of disease. The main effect of treatment strategies is to interfere with the natural history of the disease process, thereby reducing the case fatality and/or case disability ratios or extending life by providing a later age at death for conditions such as diabetes and AIDS. The HeaLY spreadsheet (available upon request from the authors at ahyder@jhsph.edu) 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.

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 healthy life year measure. It first appeared in the World Development Report of 1993 (World Bank, 1993) and has become the most widely used composite measure of population health (Jamison et al., 2006; Lopez et al., 2002; Murray & Lopez, 1994,1999; 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.

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 is obtained from a model life table based on best achievable low levels of mortality, such as those found in Japan (Coale & Guo, 1989); thus the DALY, as does the HeaLY, directly incorporates this social value choice.

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 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. DALYs are 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 & Ranaan-Eliya, 1996; Barendregt et al., 1996; Barker & Green, 1996; Hyder et al., 1998).

An important difference between the HeaLY and DALY is the fact that the starting point for the HeaLY is the onset of disease; the loss of healthy life is based on the natural history of the disease (as modified by interventions), illustrated in Figures 1-4 and 1-5. 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. When incidence is changing, however—such as with HIV in many parts of the globe—the DALY approach can greatly 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 has the advantage of standardization to ensure comparability of the multiple calculations undertaken in the GBD studies, and it has certainly 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 the 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 formulation as follows: DALYs (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 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.

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 & Shephard, 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, five years of perfect health = 5 QALYs; 2 years in a state measured as 0.5 of perfect health followed by five years of perfect health = 4 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, 1990; 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, 1990; Nord, 1992; Torrence, 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 in Oregon in the early 1990s (Exhibit 1-5).

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, NICE will recommend its use throughout the NHS; if not, it will recommend against its use in the NHS. The hope is that use of these cost-effectiveness studies as an aid to decisions will 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.

The 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 (Mathers et al., 2001; Robine, 1994). 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, the *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 book.

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." Some consider the HALE measure to provide the best available summary measure for measuring the overall level of health for populations (Mathers et al., 2001). 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 (WHO, 2000).

Health expectancy indices combine the mortality experience of a population with the disability expe-

Exhibit 1-5 Oregon: 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 1,600 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 dropped (Blumstein, 1997; Eddy, 1991; Morrow & Bryant, 1995; Nord, 1993).

rience. 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., 2001).

As originally designed, the HALE does not relate to specific diseases but rather to the average extent of disability among that proportion of each age group that is disabled. The lack of correlation between a condition or disease entity and the measure makes it less valuable for resource allocation and cost-effectiveness calculations. It is possible to convert health gap measures for specific diseases or interventions and risk factors into HALEs, but it is not clear what would be gained from this exercise.

Although the HALE is conceptually interesting and is now being calculated and included regularly in the WHO annual reports, it is not clear what additional information the HALE provides beyond the standard life expectancy data. At a national level, the amount of healthy life lost due to disability very closely parallels, and is closely proportional to, that lost due to death. As a result, the relative ranking of countries by HALEs is virtually identical to the ranking based on life expectation at birth.

Summary

Table 1-5 summarizes these four summary measures in terms of origins, purposes, level of use, sources of data, and disciplinary background of originators.

Valuing Life: Social Value Issues

The very idea of valuing some lives more than others is jarring, yet these notions are regularly reflected in our actions. The value of life is often implicit in the way resource allocation decisions are made; therefore, as much as possible such decisions should be explicit, open, and transparent. Many thoughtful people have serious reservations about assigning a single number to such a complex multidimensional phenomenon as

Table 1-5 Comparisons of Composite Summary Measures of Population Health

	Healthy Life Years	Disability-Adjusted Life Years	Quality-Adjusted Life Years	Health-Adjusted Life Expectancy
Origin	Ghana Ministry of Health, 1981	World Bank Development Report, 1993	North America, 1976	World Health Organization Report, 2000
Purpose	Assist in resource allocation decisions	Compare disease burdens in many different populations	Assess individual preferences for various outcomes from complex interventions	Compare national disease burdens
Level of use	National and district level decisions	Broad policy decisions	Personal decisions	Global comparisons
Data	National and local data from multiple sources; expert review	Global data and expert opinion	Tertiary hospital data and personal interviews	Global data and expert opinion
Original discipline base	Epidemiologists, clinicians, national planners	Economists, statisticians	Economists, clinicians	Demographers, economists, statisticians
Social values that may be incorporated	Future life discounted	Age weighting, future life discounted	Generally not included	Not relevant

health. But what is the alternative for use as a measure of utility or effectiveness in economic analyses? Outcome measures must be expressed as a unidimensional measure to be comparable to unidimensional monetary expenditure units for costs. (However, decisions about allocation should not be made on a mechanical basis; other factors, including the effect on equity, may need to be considered in decisions in addition to the goal of maximizing healthy life per unit expenditure.)

To construct composite measures of population health, important social value choices must be made. Choices about which expectation for life should be used and about valuing life lived at different ages, valuing future life as compared with the present, valuing life in terms of economic and social productivity, and valuing equity in relation to efficiency—all raise major ethical concerns.

Expectation of Life

Years of life lost due to death and to chronic disability are based on life expected had the disease not occurred. To estimate the expectation of life in a population, a choice must be made between using a local, national, or model life table. This choice should be determined by the purpose of the study.

For assisting in national and local decision making, it may be more suitable to use national life tables based on the mortality and fertility of the population in question than to use model life tables. Conversely, a model life table might be selected to reflect the best health state possible in the world, such as the west model. This selection allows a fair comparison with other countries. For example, from a global perspective it would be unfair to use national life tables to compare gains that could be achieved from a particular intervention in Ghana with those in the United Kingdom, even if both costs and lives saved were the same in each country. The reason is that those lives saved in Ghana would have a lower life expectancy than those in the United Kingdom, resulting in less healthy life saved for the same expenditure. From the global viewpoint in this example, the priority would be to fund the intervention in the United Kingdom because it would produce more healthy life per expenditure than for Ghana.

Model life tables in common use are the United Nations model life tables and the Coale and Demeney (Coale & Guo, 1989) life tables, which were used in the HeaLY and GBD studies (Hyder et al.,1998; Lopez et al., 2006). The West model life table does not refer to any geographical entity but is considered to represent a mortality pattern typical of the most technologically advanced countries. Level 26 has a female life expectancy at birth of 82.5 years, as actually experienced by women in Japan; therefore, it represents a level that could be achievable elsewhere.

Valuing Life Lived at Different Ages

Age weighting refers to the valuing of a year of life according to the age at which it is lived. This practice immediately raises questions about the basis for valuing human life. Does a day of one person's life have the same value as a day of anyone else's life? Does the value vary with age, economic productivity, or social status? Should life itself be valued separately from what is done with that life?

The Ghana Health Assessment Team (1981) judged that all human life was intrinsically valuable and that a given duration of any life was equal in value to that of any other life. The valuing of a year of life equally, irrespective of age, has been considered egalitarian (Busschbach et al., 1993; Morrow & Bryant, 1995). This choice was incorporated into the development of the HeaLY approach: A year of life lived at any age is equally valued.

The original DALY formulation assigned an exponential function to provide a value chosen so that life lived as a dependent (e.g., infants, children, the elderly) is given less value than life lived during the productive years. With this approach, the intrinsic value of life increases from zero at birth to a maximum at age 25 and declines thereafter, so that a day of life of a 50-year-old is worth about 25% less than that of a 25-year-old. Paradoxically, the age weighting used in the original DALY formulation leads to higher valuation of life lived before age 15 than does the HeaLY formulation, in which life lived at all ages has equal value (Barendregt et al., 1996; Hyder et al., 1998). Current formulations of the DALY leave age weighting as an option, and such weighting is not used with the HALE.

Age-related valuing has been justified by studies showing that individuals value their own life lived at different ages differently. Such values have been reported in the literature, and studies have reported that they are consistent across respondents of different ages (Busschbach et al., 1993). Murray and Lopez (1994) report studies from many countries that reveal a preference for saving younger lives as compared with older ones. Nevertheless, it is not clear how much of the differential valuing of life at different ages is related to an underlying appreciation that economic and social productivity varies at different ages. If it is decided that healthy life should be valued according to economic and social productivity, then an alternative to age weighting might be to explicitly add a productivity factor or subtract for the societal costs of dependents, such as education (see the section "Valuing Life for Its Economic and Social Productivity").

Valuing Future Life Compared with Present Life: Discounting

Discounting is the process for determining the present value of future events. Social time preference takes into account the phenomenon that people value events at present more highly than those in the future (independent of inflation and of uncertainty). For investments in other sectors, time preference is normally taken into account by discounting future returns and costs by some appropriate discount rate. Thus the discount rate can be considered the inverse of an interest rate. The main issue concerning discounting in relation to summary measures is whether discounting life itself is appropriate. There seems little problem about the usefulness of discounting the future value of what is produced by healthy life, but should the life itself be discounted (Morrow & Bryant, 1995)?

Discounting has been applied in the health sector because both the losses from a disease and the benefits from a health intervention often occur in the future. An intervention today may not produce immediate benefits (such as in immunization), or it may result in benefits being sustained over a long time (such as in supplementary nutrition). The costs for these activities must be borne now, but the benefits are realized in the future and are less valuable than if they could occur now. This is equivalent to investing money now so as to obtain more in the future. A healthy life year now has greater intrinsic value to an individual or community than one in the future (Gold et al., 1996; Weinstein et al., 1996).

The rate at which society is supposed to discount has been termed the social discount rate (SDR), a numeric reflection of societal values regarding intertemporal allocation of current resources. There is no consensus about the most appropriate choice of a discount rate in health, but most agree that it should be lower than that used in the private commercial sector. The WDR in 1993 and the GBD studies used a discount rate of 3% per year; in lieu of other information, this rate has come to be used in most international health cost-effectiveness studies. Nevertheless, the impact of using a range of discount rates, including zero, should be explored with each study.

Valuing Life for Its Economic and Social Productivity

Whether and how to value economic and social productivity for purposes of healthcare decision making is highly contentious; to a large extent, the age weighting incorporated in the original DALY formulation was considered by many to be a proxy for productivity. The consensus now seems to be that any such

valuations should be considered separately, made explicit, and very much dependent on the purpose of the valuations.

In general, productivity may be attributed to adults aged 15 to 64, and those in these age groups could be given a higher value. Persons younger than age 15 and older than age 65 may be considered as dependents and given a lower value. Many variations for differential valuing are possible, including type of employment. People at different socioeconomic levels in a society are expected to have different capacities for productivity—yet, to value life according to income levels or social class would not seem fair and generally would not be acceptable. In poor countries, the value of marginal wages for subsistence agriculture is negligible, but the value of the workers' lives certainly is not.

A fundamental question is whether to consider adding a productivity component to the summary measure. Health issues do not readily conform to the requirements of market economics; information is inadequate, and misinformation is rife on the part of the providers as well as the public. Externalities from good health are generally large. Demand for costly services is largely determined by the healthcare providers rather than by the consumers. Competitive market forces have not worked well for those in greatest need. In the private sector, demand for services is clearly related to productivity and willingness (and ability) to pay. If left to market forces alone, inequitable distribution would be inevitable.

Economic arguments have been put forward for valuing life according to productivity, but counterclaims have been made that human life cannot and should not be expressed in economic terms for decision-making purposes. Nevertheless, efforts to avoid such expression result in implicit valuation of life. Barnum (1987) has argued for adding productivity to the valuing of human life, stating that it has been ignored in health policy, is readily quantifiable, and does not ignore the welfare of children because the whole population is dependent on adult productivity for quality and sustenance. Such economic appraisal of human life is often based on the net transfer of resources from the "producers" to the "consumers" and the consequent interdependence of people.

In relation to this issue, in the *Report of the Commission on Macroeconomics and Health* (WHO, 2001), a DALY was stated to be worth at least an average annual income per head. Although the basis for such a valuation was not adequately justified, the basic notion seems right. More work on explicit valuations of human life and what it produces are needed, and will certainly affect health-related cost-effectiveness decisions.

Valuing Equity in Relation to Efficiency

A child born in Malawi or Uganda will likely live only half as long as one born in Sweden or Singapore; one in three babies born in Niger or Sierra Leone will not live to see his or her fifth birthday. These inequalities are unfair and harmful and, therefore, qualify as inequities. In terms of social justice, equity has to do with a fair distribution of benefits from social and economic development. However, the term *equity* is used in different conceptual senses: equal access to health services for all (opportunity equality), equal resources expended for each individual (supply equality), equal resources expended on each case of a particular condition (equality of resource use to meet biological need), equal healthy life gained per dollar expended (cost-effectiveness), care according to willingness to pay (economic-demand equality), care according to biological or socioeconomic need, and equal health states for all.

Decisions based on cost-effectiveness (e.g., cost per healthy life year), therefore, may not accord well with concerns about equity. These calculations are generally indifferent to equity; they are designed to steer interventions to what is efficient, whatever the differential need may be. To meet the requirements of equity, health system planners need to go beyond ensuring equality of access to health care and require a balance so that health system responses are in accord with equity as well as efficiency.

Provided that health information is available according to socioeconomic and vulnerable groups, use of these summary indicators as tools for equity by calculating healthy life per dollar to be gained by all socioeconomic and vulnerable groups could readily be undertaken. It would be straightforward to assess the impact of specific health decisions to ensure that they enhance equity. Summary measures such as HeaLYs and DALYs can be used to guide allocation of resources to ensure equitable distribution of those resources so as to reach those most in need. Cost-effectiveness by itself does not provide adequate guidance; equity should be an associated criterion to govern the distribution of societal benefits.

Data for Composite Measures

Types of Data

The data needs for estimating the burden of disease in a region or country are extensive, and obtaining even reasonable estimates in low- and middle-income coun-

tries has been a source of concern (Anand & Ranaan-Eliya, 1996; Barker & Green, 1996; Bobadilla, 1998; Murray et al., 2002). Brief descriptions of the types of data required follow; available data need to be carefully reviewed and optimally utilized.

Demographic Data. Population data are integral to burden of disease estimations and are needed both as denominators and for consistency checks. In a national setting, a recent census is useful for providing population counts by age, sex, and geographic location. Particularly helpful, when there is inadequate death registration, is to have a one-year post-census follow-up on a sample of enumeration areas so as to obtain robust age, sex, and place mortality. The age and sex distribution of the population is critical, and often is a major factor that determines the nature of the disease burden. A good vital registration system is a key asset that will provide both birth and death numbers. Underreporting, age misreporting, and other biases in data may have to be addressed (using standard demographic methods) prior to use of these data in burden of disease estimation.

Mortality. Mortality data are required for any burden of disease analysis. Specifically, age, sex, and place mortality rates greatly assist the analysis by defining the contribution of mortality to the pattern of disease burden. They also serve as an essential framework that constrains estimates obtained from a variety of special studies that fill important information gaps but may be incomplete or biased in the populations covered. Reporting errors, such as underreporting of deaths and reporting of age at death, need to be carefully examined. In particular, information has to be evaluated for deficiencies in the under-five group and older age groups. For the youngest ages, the probabilities of deaths in the first year (1q0) and in the next four years (4q1) provide better estimates of the risk of death than do overall mortality rates. Methods such as the Brass method for indirect estimates of mortality provide useful ways to assess age-specific mortality data for potential errors (Hill, 2001).

For burden of disease studies, cause of death data are required for all ages, but reliable cause of death records are rarely available in low- and middle-income countries, especially for deaths that do not occur in healthcare facilities. Even if available, the classification system used may be outdated rather than ICD based, and the reliability of coding may vary by the type and location of the hospital. Young-adult deaths may be better recorded than deaths of infants and the elderly. Especially in low-income countries, it can be helpful to cross-check death records with other information, using postmortem interviews and hospital registers to assist in defining causes of death or to extrapolate from other data or other regions to assist in the estimates.

Morbidity. Meaningful data on disability are even more difficult to find and interpret than mortality data. Often morbidity information is institution based or restricted to one or two sources, such as hospital inpatient and clinic outpatient records. The representativeness of small studies and the range and types of morbidity covered in any survey need careful evaluation. National disability surveys or regional studies conducted for the evaluation of disabled people may be available; such research is useful in providing some estimate of the prevalence of serious disabilities and their age and sex distribution. However, linkage between disability and disease is often not available, and attributing one type of disability to specific causes is difficult. For example, because many conditions can lead to blindness—for example, diabetes, hypertension, injuries, trachoma, and cataracts—the attribution of proportions of blindness in a population to its cause can be problematic. Information on the duration of disability may be found in specialized studies and the experience of institutions. The severity of disability will have to be rated on a scale; the various methods used in the literature were described earlier in this chapter.

Variables

The types of data just described need to be processed in the form of specific disease-based estimates. The key variables are defined in Table 1-4.

The incidence rate (usually expressed per 1,000 general population per year) is central to the natural history of disease concept. Although incidence is a basic epidemiologic indicator, it is usually not found in routine data collection systems. Special studies, prospective surveys, or calculations based on the prevalence (which is more commonly available than the incidence) and knowledge of the average duration of the disease can be helpful in developing this measure.

The case fatality rate is the proportion of those developing the disease who die from it at any time. It is expressed as a decimal value between 0 (for nonfatal conditions) and 1 (for universally lethal conditions such as AIDS). The case disability ratio (analogous to the CFR) is the proportion of those diagnosed with a disease who have signs or symptoms, and is usually 1 (as discussed earlier).

Age is required in various formats. Age at onset is when disease onset occurs in a population; age at

fatality denotes the age at death as a result of the disease.

The expectation of life at age of onset is the years of life expected at that age had the disease not occurred. Similarly, expectation of life at fatality is the years of life expected at that age had the death not occurred.

Checking Data

Data used for generation of indicators need to be evaluated for validity, reliability, and consistency, using defined qualitative and quantitative criteria. Large population-based studies may be given preference over smaller sample-based work if both are available and the quality of their data is comparable. Better conclusions may be possible by cross-checking different sources of data. Community-based studies, which may be representative of the population but have limited diagnostic validity, may be compared with hospital-based work, in which diagnosis may be valid but would come from a biased population sample. The following subsections profile simple types of checks for data quality.

Comparison of Total Numbers. Cross-checks should be done to compare total numbers. It is essential to check that the number of deaths in a year in a region is the same as the sum of all deaths from all causes in the same region. Similarly, program-based data can be compared with data from other sources to ensure better estimates of causes of death. The comparison of totals allows one to work within a frame of mortality and avoids double counting of one death. However, it does not assist in the distribution of deaths within that frame.

Relationship Between Variables. Checks based on the epidemiologic relationship between parameters refer to the application of simple, yet vital, relationships such as the following:

- Prevalence (point) = incidence × average duration of disease
- Cause-specific mortality rate = incidence × case fatality rate

These checks allow estimates from different sources to be compared for internal consistency. Such relationships can also be used to derive one of the estimates in the equations when the others are known.

Sensitivity Analysis. Sensitivity analysis is a useful tool to determine whether data that are more precise are required for the purposes of a particular decision. A one-way sensitivity analysis (Petiti, 1994) evaluates the effect of manipulating one variable at a time on the dependent variable. If the outcome is sensitive to one or more variables, their precision is more important in the estimation.

Disease Groups: Classification

Murray and Chen (1992) introduced a disease group system based on the WHO ICD classification system. Group I includes conditions characteristic of low-income countries: communicable diseases, maternal and prenatal conditions, and nutritional deficiencies. These conditions decline at rates faster than overall mortality rates as socioeconomic conditions improve; thus group I contributes to a relatively small percentage of deaths in the high-income world. Group II, which consists of noncommunicable and chronic diseases, accounts for most loss of healthy life in the high-income countries and proportionately increases with the epidemiologic transition in low- and middle-income countries (see Exhibit 1-3). Group III consists of injuries, both intentional and unintentional (including violence).

The distribution of the disease burden among these three groups is one indicator of the type of disease burden and the level of epidemiologic transition in a country. It is important to distinguish between the proportions of deaths attributed to these groups, as opposed to the risk of dying from the conditions in these groups. For example, the proportion of deaths attributable to group II causes increases from high- to low-mortality countries (or to an older age structure of the population); however, the risk of death from group II conditions is higher in high-mortality countries.

Implementing a Burden of Disease Study

Knowing how to conduct a burden of disease analysis is important for all countries. Generic steps for a national burden of disease study include the following:

- Assess demographic information, including a census with age, sex, geographic (urban/rural), and selected socioeconomic status information, and vital statistics with births and deaths.
- Collect cause-of-death information for all deaths in a year by age, sex, geographic location, and socioeconomic status as possible, according to the WHO ICD system.
- Define disability by cause/disease, and develop a severity scale using expert and community input.
- Collate information by disease from all sources and assess reliability/validity, using expert

opinion when needed to define variables for a spreadsheet.

- Decide whether social value preferences such as age weighting, discounting, economic and social productivity, and expectation of life will be used and what their values will be.
- Estimate healthy life lost for each disease condition and by disease groups.
- Perform sensitivity analyses to check the robustness of results relative to critical variables and assumptions.
- Consider other variations, including assessment of losses by risk factors; regional, age, and sex breakdowns; and future projections.
- Review the policy implications on overall mortality and morbidity in the country and by cause; feed data into cost-effectiveness analysis and further research.
- Include other modifications as appropriate to the country setting.

To use summary measures to assist in health planning and resource allocation decisions, additional steps include the following:

- Estimate the effectiveness (gains of healthy life) of each intervention under consideration in terms of expected coverage and reductions in incidence and/or case fatality or case disability ratios.
- Work out the costs of the proposed interventions.
- Develop cost-effectiveness ratios to plan which combination of interventions targeted to which groups will provide a maximum return of healthy life per expenditure for the funds allocated to health.
- Review expected gains of healthy life according to age, sex, geographic area, and socioeconomic and vulnerable groups to ensure that all are better off (or at least none are worse off) and adjust as necessary.

Another very important consideration in this process is time. The conduct and analysis of such studies must be timely to assure its appropriate use by policy makers and useful for resource allocation decisions. The precision and comprehensive nature of the study must be balanced by the need for timely results.

The steps described previously may be carried out simultaneously or in some sequence, depending on the specific national situation. Modifications will likely be needed depending on the availability of data (Exhibit 1-6). An actual study requires careful planning on the part of those responsible for its conduct and may include many additional steps that are beyond the scope of this chapter. Even so, these generic steps summarize the essentials of applying the burden of disease methods to a country. Increasingly countries are obtaining, refining, and using these data on an ongoing fashion.

Comparisons and Trends in Disease Burden

This section reviews a number of country-based burden of disease studies so as to compare and assess trends in disease burden from place to place and over time.

Comparative Disease Burden Assessments

Comparing the burden of disease across populations, time, and place is an important aspect of national burden of disease studies. This subsection uses examples from burden of disease studies to illustrate how disaggregated data can help in understanding the distribution of ill health in a country.

The Andhra Pradesh Burden of Disease Study, 2001

The regional distribution (urban/rural, state, district) of the disease burden is important to explore in a national burden of disease study. Andhra Pradesh, a state in India, was the focus of one of the most meticulous burden of disease studies conducted between 1994 and 2001. It had a population of 76 million in 2001, 27% of whom lived in urban areas (20.8 million people); a 1:3 ratio of urban-to-rural disease burden in terms of DALYs lost was identified (Mahapatra, 2001). The burden of disease rates were 19% higher in rural areas than in urban areas, as measured by DALYs lost per 1,000 population (Figure 1-8).

The Burden of Disease and Injury in New Zealand, 1996

Age and ethnicity are key characteristics of a population that require a disaggregation of the burden of disease. The national burden of disease study of New Zealand (1996 population = 3.6 million) provides a clear example of how the DALYs lost in 1996 were predominant among the older age group (65 years and older), even though this group represented only

Exhibit 1-6 **The Burden of Disease in Pakistan, 1990**

Pakistan is a developing country in South Asia, whose population numbered 112 million in 1990. A study was undertaken to estimate the burden of disease in Pakistan in 1990 and to calculate the loss of healthy life from a spectrum of common conditions. Nearly 200 data sources were evaluated, including national surveys, population-based studies, sentinel survey systems, and disease-specific studies.

Overall, 456 discounted HeaLYs per 1,000 people were lost due to new cases of diseases in 1990. Diarrhea and pneumonia in children caused the greatest loss of healthy life. Sixty-three percent of healthy life was lost from mortality, while 37 percent was lost due to disability. Hypertension and injuries were the leading causes of healthy life lost from disability. Nearly half the healthy life was lost in the under-five age group, demonstrating the great burden on Pakistani infants and children.

Although communicable diseases dominated the burden of disease in Pakistan in 1990, noncommunicable diseases also took a heavy toll. Figure 1-7 and Table 1-6 review the top conditions responsible for loss of healthy life, and the proportion of loss from noncommunicable conditions can only be expected to increase. Injuries also need to be recognized as a major public health problem in the country. According to these estimates, Pakistan had a lower overall burden of disease than most countries in sub-Saharan Africa in 1990, but a burden higher than most countries in Latin America.

Table 1-6 **Loss of Healthy Life in Pakistan: Top 10 Conditions for 1990**

	Premature Mortality Only Rank	Disability Only	Healthy Life Years Lost
Rank	**Disease**	**Disease**	**Disease**
1	Diarrhea	Hypertension	Diarrhea
2	Childhood pneumonia	Injuries	Childhood pneumonia
3	Tuberculosis	Eye diseases	Tuberculosis
4	Rheumatic heart disease	Malnutrition	Birth diseases
5	Chronic liver disease	Birth diseases	Injuries
6	Congenital malformations	Congenital malformations	Hypertension
7	Birth diseases	Dental diseases	Congenital malformations
8	Ischemic heart disease	Ischemic heart disease	Chronic liver disease
9	Child septicemia	Adult female anemia	Ischemic heart disease
10	Other respiratory diseases	Mental retardation	Rheumatic heart disease

Source: Hyder & Morrow, 2000.

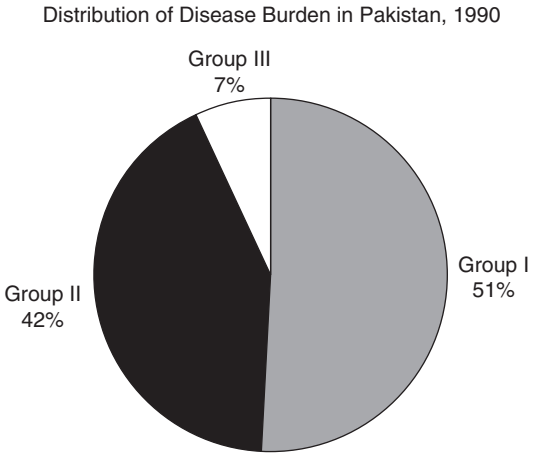
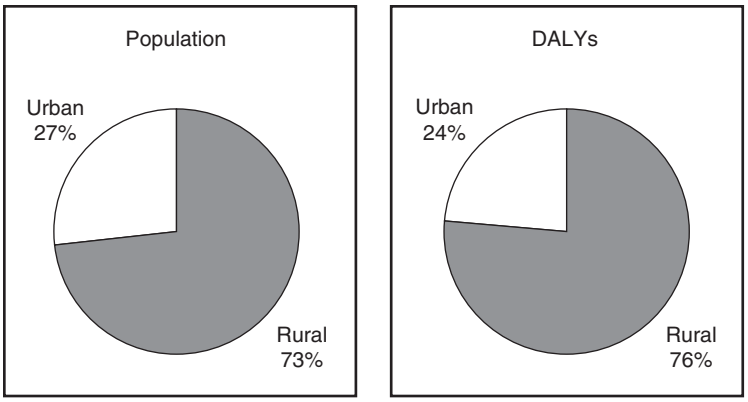


Figure 1-7 The Burden of Disease in Pakistan, 1990.
Source: (Hyder et al, 2000).



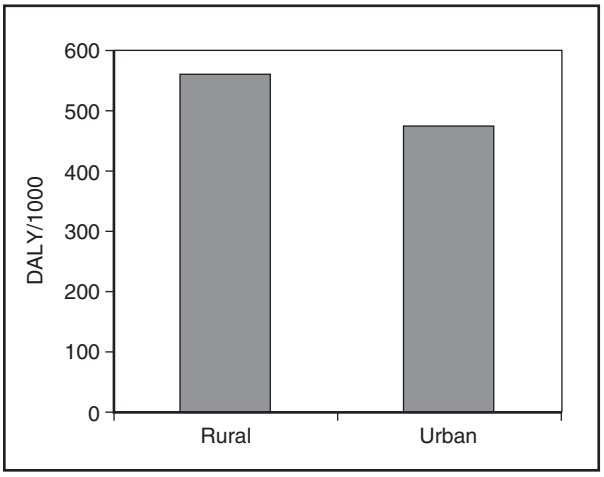


Figure 1-8 Burden of disease in Andhra Pradesh, 2001, by region.
Source: Based on Mahapatra, P. (2001). "Estimating National Burden of Disease: The Burden of Disease in Andhra Pradesh 1990s." Hyberdad: Institute of Health Systems.

Note: Total DALYs lost in Andhra Pradesh = 5 million.

12% of the population (New Zealand Ministry of Health, 2001). The identification of 15% of the burden of disease in the indigenous Maori population, who represent only 9.7% of the total New Zealand population, was an important equity finding (Figure 1-9).

Burden of Disease in Chile, 1993

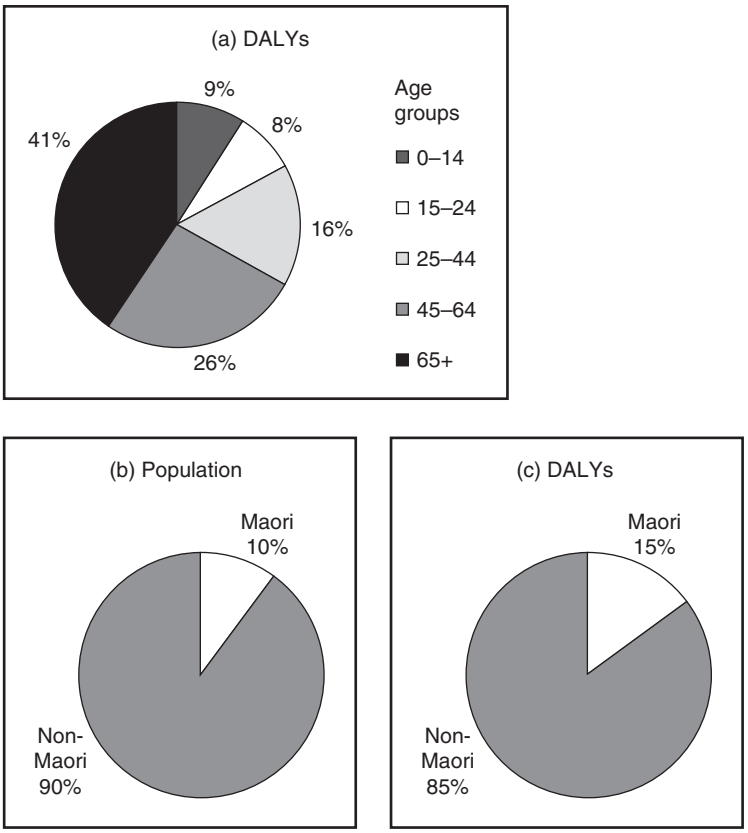
A disaggregated burden analysis by gender can also be seen in the work done in Chile in 1993, where at that time 49.6% of the population was male. The study found that 56% of the DALYs lost were attributable to males (Figure 1-10). The distribution of the burden by major disease groups—I (communicable, infectious), II (chronic, non-communicable), and III (injuries, violence)—showed the dominance of chronic conditions in the burden (Concha, 1996).

Burden of Disease Estimates for South Africa, 2000

HIV/AIDS is ravaging Africa; thus the impact of HIV/AIDS on the burden of disease in African countries can be significant. In South Africa, 30% of the 15 million DALYs lost in 2000 were attributed to HIV/AIDS (Figure 1-11) (Burden of Disease Research Unit, 2003); for a population of 45 million, this means 0.33 DALY is lost per capita. Such data are important for national decision making.

The Burden of Disease and Injury in Australia, 1996

The distribution of disease burden by socioeconomic variables is important for poverty and equity analysis. The national burden of disease analysis in Australia for 1996 presented results based on socioeconomic status (defined by the social and economic characteristics of the living area), disaggregated by gender, for both mortality (YLL) and disability (YDL) (Figure 1-12) (Mathers et al., 1999). These results show the high disability losses for women and for the poor. Such explorations of intranational distributions of disease burden are useful in studying the disproportionate effects of ill health on the poor and women.



Note: Total DALYs lost in New Zealand for 1996 = 500,000.

Figure 1-9 Burden of disease in New Zealand, 1996, by age (a) and ethnicity (b).
Source: Based on data from New Zealand Ministry of Health. (2001). "Burden of Disease and Injury in New Zealand."

Burden of Disease and National Income

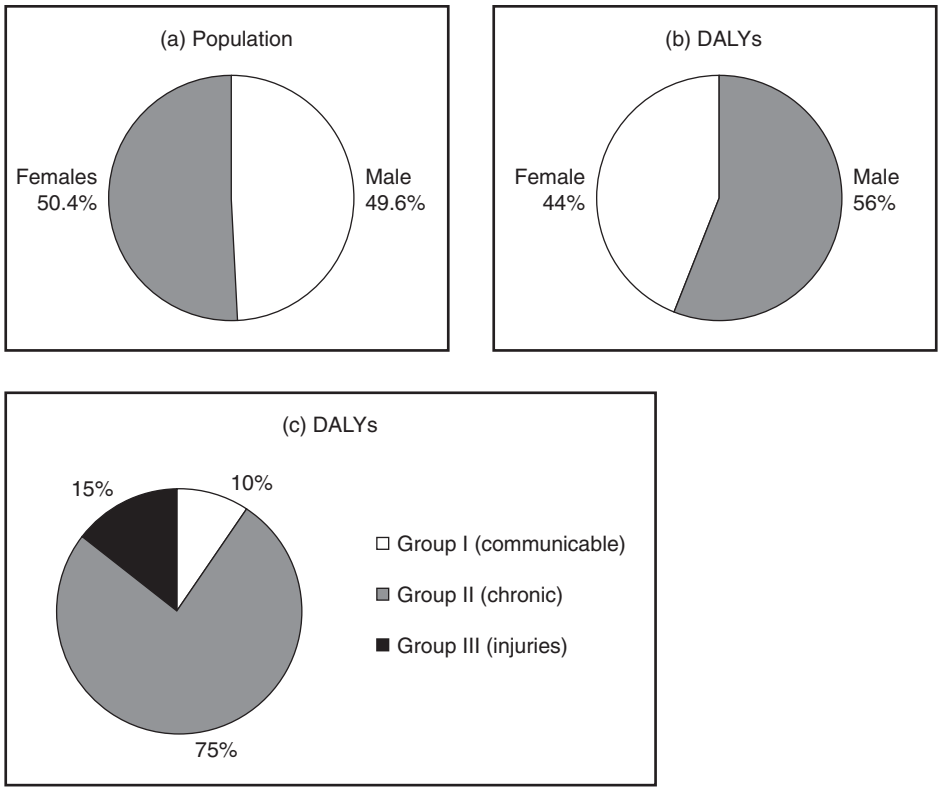
WHO has categorized its member states by income levels into high-, middle-, and low-income nations. The population of the world in 2004 totaled slightly more than 6.4 billion people, with 85% residing in low- and middle-income nations (Figure 1-13). As may be expected, more than 90% of the global burden is found in low- and middle-income nations, reflecting the double challenge faced by the majority of people in the world: They are poor and they are unhealthy. This relationship between ill health and poverty has long been recognized as complex and has been the object of much research and inquiry.

Burden of Disease by Disease Groups

Another way to disaggregate data is to explore the disease burden based on three disease groups: group I (communicable, infectious, maternal, and perinatal), group II (noncommunicable, chronic), and group III (injuries and violence). There is great variation in the portions allocated to these groups; for example, group I conditions may be responsible for anywhere from 12% to-70% of the burden of disease. When the countries are stratified by GNP per capita as a measure of development, an important trend can be seen from historical data (Table 1-7): As income rises, the proportion of the burden attributable to group I conditions decreases, while the share attributable to group II conditions increases. This effect is progressive, although countries such Turkmenistan (a middle-income country) still retain a high group I burden. This finding is consistent with the theory of epidemiological transition, which predicts a change in a country's disease profile with economic development.

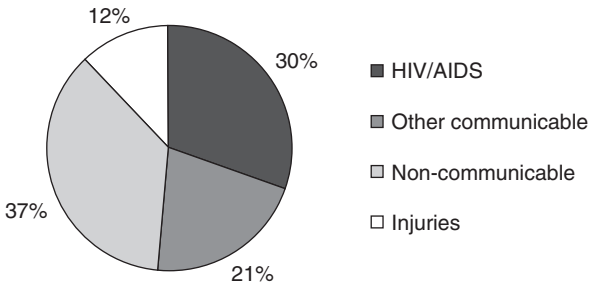
Global Assessments of Disease Burden

Information regarding health and disease for all countries of the world can be collated to provide a picture of global health status. In addition, global health assessments may be completed as a separate



Note: Total DALYs lost in Chile for 1993 = 2 million.

Figure 1-10 Burden of disease in Chile, 1993, by gender (a) (b) and disease groups (c).
Source: Based on data from Concha, M. (1993). "Burden of Disease in Chile."



Note: Total DALYs lost in South Africa for 2000 = 15 million.

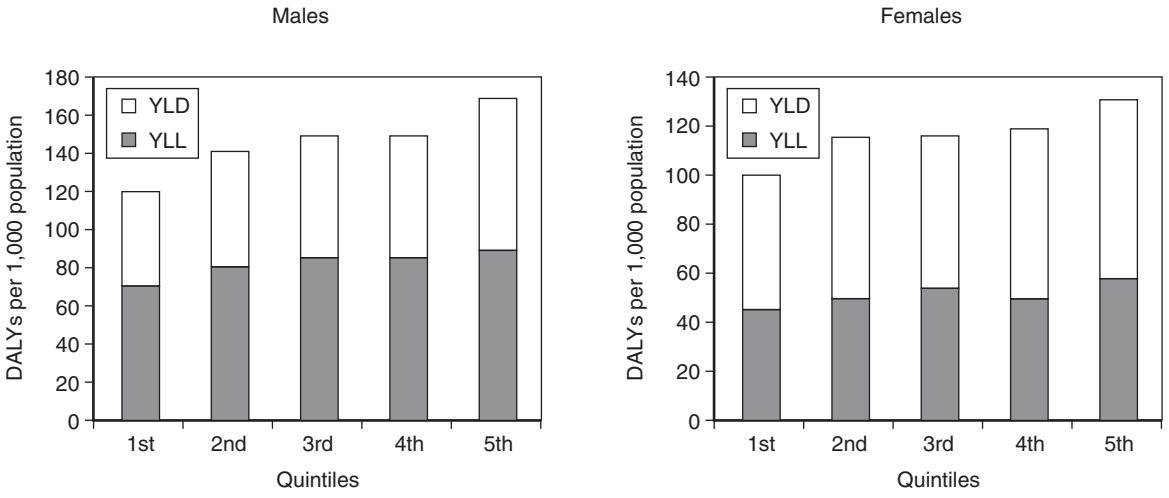
Figure 1-11 Burden of disease in South Africa, 2000, by disease groups.
Source: Based on data from Burden of Disease Research Unit. (2003). "Initial Burden of Disease Estimates for South Africa, 2000." South Africa: South African Medical Research Council.

activity, and such data can then be disaggregated into regional information. Global assessments serve to highlight major challenges facing the world community, and trends in such assessments indicate progress, if any, in improving the health of people worldwide. Such information is critical to the work of organizations such as WHO and UNICEF in their efforts to combat ill health and disease worldwide. This section highlights results of global exercises for assessment of the disease burden, recent evaluations, and projections for the future.

The Global Burden of Disease

The Global Burden of Disease (GBD) 2000 study constructed estimates of mortality, disability, and DALYs by cause for regions of the world. Demographic estimates of deaths in 2000 by age and sex form the basis of this work. Subsequently, WHO undertook an update of the GBD study to produce reliable estimates of mortality and morbidity for all regions of the world for 2004. The results were based on a variety of sources, including vital registrations systems, special studies, surveys, and expert opinion. This section reviews the 2004 GDB data.

Mortality. Globally, in 2004, an estimated 58.8 million deaths occurred, 53% of whom were males. Ischemic heart disease, cerebrovascular disease, and lower respiratory infections were the top three causes



Note: The first quintile corresponds to the highest socioeconomic group, and the fifth quintile to the lowest. Each quintile contains approximately 20% of the total Australian population. Total DALYs lost in Australia for 1996 = 2.5 million.

Figure 1-12 Burden of disease in Australia, 1996, by socioeconomic status and gender.
Source: Based on data from Mathers, C., Vos, T. & Stevenson, C. (1999). "The Burden of Disease and Injury in Australia." Australian Institute of Health and Welfare.

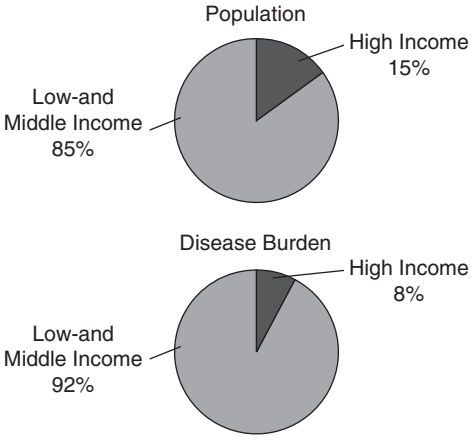


Figure 1-13 Global Burden of Disease, 2004, by Income Level of Countries.
Source: Based on data from World Health Organization. (2004).

of death, while 10 causes accounted for 50% of deaths worldwide. One death in 10 was from injuries, with road traffic accidents included in the top 10 causes of deaths. Approximately 10.4 million deaths occur in children younger than 5 years of age globally, with more than 50% of these fatalities being caused by just four communicable diseases. Of those under-five deaths, 75% occurred in the African and Southeast Asia regions. While the effect is less pronounced in the adult population, the proportion of deaths in the 15- to 59-year age range remains skewed toward low- and middle-income countries. Mortality in the African region is 40% higher than the next-highest-mortality region and is four times higher than in high-income countries. Thus an inordinate share of the mortality burden at the beginning of this decade was found in low- and middle-income countries, even among adults.

Table 1-8 shows the differences in the 10 leading causes of deaths for 2004 for the high-income countries and the low-income countries. The presence of perinatal conditions, tuberculosis, HIV/AIDS, and malaria in the low-income world is indicative of the high impact of these conditions on premature mortality. These conditions are absent from the top 10 causes in the high-income countries, reflecting the success in combating these infectious conditions in the modern era. It is important to note that noncommunicable diseases such as ischemic heart disease were already prominent causes of premature deaths in the low- and middle-income world in 2000 and remain among the top five causes of death in 2004.

Disability. The GBD study 2004 update also provides an evaluation of the contribution of conditions to disability in the world. Leading causes of disability in 2004 worldwide are shown in Table 1-9. Neuropsychiatric and behavioral conditions dominate the causes of disability, accounting for four of the top 10 conditions. However, a diverse spectrum of

Table 1-7 Historical Distribution of Disease Burden Within Countries

	Disease Burden in Disease Categories (of 100%)		
Country	**Group I**	**Group II**	**Group III**
Low-Income Nations (GNP per capita: $635 or less)*			
Andhra Pradesh	54	30	16
Guinea	70	23	7
Lower Medium-Income Nations (GNP per capita: $636–$2,555)			
Colombia	22	39	39
Jamaica	16	60	24
Turkmenistan	51	45	4
Uzbekistan	46	40	14
Upper Medium-Income Nations ($2,556–$7,911)			
Mauritius	16	74	10
Mexico	32	48	20
Uruguay	12	73	15

Note: Disease classification system: Group I: communicable, infectious, maternal, and perinatal; Group II: noncommunicable and chronic; Group III: injuries and accidents.

Source: World Bank, 1993.

*GNP per capita from World Bank (1993).

conditions, such as hearing loss, cataracts, and osteoarthritis, also appear on the list. A unique contribution of the GBD work has been its placement of nonfatal health outcomes in the center of international health policy in recent years. The important, and yet often ignored, effects of these conditions are obvious once disability is counted in these estimates of disease burden.

Disease Burden. Based on the estimation of deaths and disability presented in the preceding subsection, the global disease burden for 2004 was estimated using DALYs. Leading causes of the global burden in 2004 (Table 1-10) indicate how those conditions affect the low-middle income world. The top 10 list is a mixture of the unfinished agenda of communicable and perinatal conditions, noncommunicable diseases, and road traffic injuries. This situation highlights the challenge facing the global health community as it simultaneously continues to fight infectious diseases, seeks to improve the response to chronic conditions, and prepares to meet the increasing impact of injuries.

Age and Disease Distributions. Figure 1-14 illustrates the distribution of the global burden in 2004 by disease groups and demonstrates the growing relative impact of chronic diseases (group II) over infectious conditions (group I). Figure 1-15 provides comparable figures for loss of healthy life in sub-Saharan Africa, the Middle Eastern Crescent, Latin America, and the Caribbean. Note that communicable diseases still represent a considerable portion of the disease burden in 2004, especially in sub-Saharan Africa due to HIV/AIDS.

As the figures demonstrate, various subregions within middle- and low-income countries are at different stages of epidemiological transition. The influx of chronic diseases has added another layer of problems in some areas, while the burden of communicable

Table 1-8 Leading Causes of Deaths in High-Income and Low-Income Countries, 2004

High-Income Countries		Low-Income Countries	
Rank	**Cause**	**Rank**	**Cause**
1	Ischemic heart disease	1	Lower respiratory infections
2	Cerebrovascular disease	2	Ischemic heart disease
3	Trachea, bronchus, and lung cancers	3	Diarrheal diseases
4	Lower respiratory infections	4	HIV/AIDS
5	Chronic obstructive pulmonary disease	5	Cerebrovascular disease
6	Alzheimer's and other dementias	6	Chronic obstructive pulmonary disease
7	Colon and rectum cancers	7	Tuberculosis
8	Diabetes mellitus	8	Neonatal conditions
9	Breast cancer	9	Malaria
10	Stomach cancer	10	Prematurity and low birth weight

Source: Based on data from WHO, 2008.

Table 1-9 Leading Causes of Disability Losses Globally, 2004

Rank	Cause
1	Unipolar major depression
2	Refractive errors
3	Hearing loss, adult onset
4	Other unintentional injuries
5	Alcohol use disorders
6	Cataracts
7	Schizophrenia
8	Osteoarthritis
9	Bipolar disorder
10	Iron-deficiency anemia

Note: Disability losses are defined by years of life lived with disability (YLDs).

Source: Based on data from WHO, 2008.

Table 1-10 Leading Causes of Global Burden of Disease, 2004

Rank	Cause
1	Lower respiratory infections
2	Diarrheal diseases
3	Unipolar depressive disorders
4	Ischemic heart disease
5	HIV/AIDS
6	Cerebrovascular disease
7	Prematurity and low birth weight
8	Birth asphyxia and birth trauma
9	Road traffic accidents
10	Neonatal infections and other

Source: Based on data from WHO, 2008.

diseases has not yet been eradicated. This "double burden" poses a major challenge for the health systems in these nations. In addition, the scarcity of resources in many of these countries makes the situation even more critical, and it becomes imperative to define interventions that are cost-effective and able to reduce the burden of disease.

Other Ways Burden Can Be Measured

Mortality and morbidity alone have been used for decades for international comparisons of disease burden. Mortality among children younger than five years is considered a sensitive indicator of overall health of nations, but especially for the health of women and children. UNICEF publishes an annual *State of the World's Children* report that includes a ranking of nations based on this indicator (Table 1-11).

Gross national income (GNI) per capita is an indicator of national wealth, and the relationship between these variables usually follows an expected sequence, such that the country with the lowest GNI per capita has the worst indicators of health. However, as Table 1-11 indicates, even countries that have relatively higher per capita income can have poor indicators of health service accessibility (e.g., coverage of tetanus toxoid vaccination for pregnant women) and health impact (e.g., prevalence of anemia in pregnant women). For example, the per capita GNI for Bhutan is higher than that for Mongolia, yet Bhutan ranks lower than Mongolia in child mortality and life expectancy. Such examples demonstrate that the relationship between health and poverty is complex and needs in-depth investigation. When seeking to improve the health of nations, both absolute poverty and the disparities within societies serve as impediments to empowerment of the poor and needy, especially women and children.

Projections

Forecasts of disease burden have been attempted with the intent of providing some basis for health planning. Making such projections is a challenging task that requires further data manipulations and the use of assumptions. These assumptions must predict changes

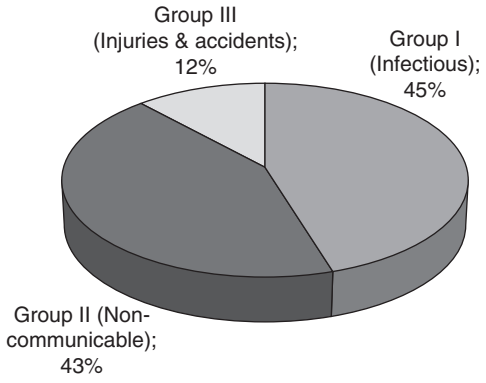


Figure 1-14 Global Burden of Disease 2004 by Disease Groups.
Source: Based on data from WHO. The global burden of disease: 2004 update. Geneva: World Health Organization, 2008.

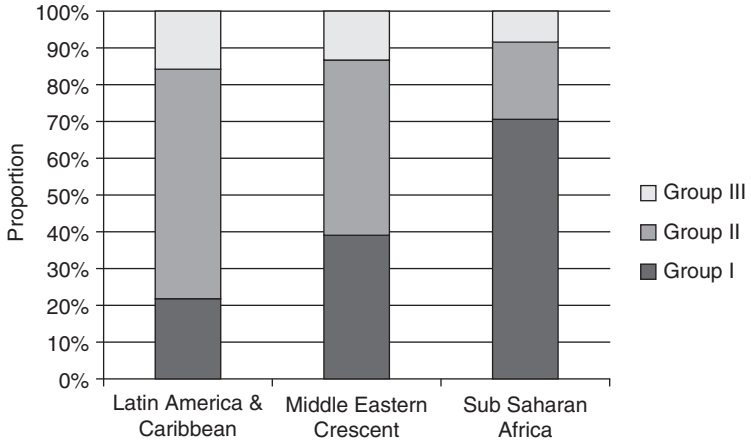


Figure 1-15 Proportion of Disease Burden by Disease Groups in Selected Regions, 2004. *Source*: Based on data from WHO. The global burden of disease: 2004 update. Geneva: World Health Organization, 2008.

in disease prevalence and incidence over time, the effects of interventions, and other factors. As a result, all projections are estimates with substantial variations that are highly dependent on the data used to derive them.

The GBD study updated for 2004 projects the global burden to the year 2030. These estimates are based on projected changes in life expectancy, age structure of the global community, disease profiles based on current states, and other relevant parameters (WHO, 2008). In addition, the projections are guided by forecasts for income per capita, human capital, and smoking intensity. The results of this exercise reveal the leading causes of projected global burden of disease for 2030, as summarized in Table 1-12. The domination of chronic diseases on this list is obvious, although respiratory conditions still appear to be important. Injuries from road traffic crashes are predicted to become the third leading cause of the global disease burden in the future. In addition, the lower ranking of HIV on the list reflects the assumption that interventions for this condition will succeed in reducing the burden in the intervening decades. This may or may not hold true, and other assumptions may be used to create a different scenario for the future.

The growing importance of noncommunicable diseases is a global phenomenon, and these conditions' impact on low- and middle-income countries and regions needs to be assessed. However, unlike

Table 1-11 Health Status Indicators and National Income for Selected Low- and Middle-Income Countries

Country	Ranking by Child Mortality (<5 years)	Life Expectation (years)	Stunted Children <5 years (%)	Coverage of Tetanus Vaccination Among Pregnant Women (%)	GNI per Capita (U.S. dollars)
Niger	2	46	40	36	170
Sierra Leone	1	34	34	60	140
Angola	3	40	45	62	660
Afghanistan	4	43	52	34	250
Mongolia	64	64	25	—	440
Pakistan	44	61	37	56	410
Bhutan	50	63	40	70	590
Nicaragua	82	69	20	95	370
Peru	86	70	25	57	2,050
Guatemala	74	67	46	38	1,750

Source: UNICEF, 2004.

Table 1-12 **Leading Causes of Disease Burden, 2004 and Projected, 2030**

2004		2030	
Rank	**Cause**	**Rank**	**Cause**
1	Lower respiratory infections	1	Unipolar depressive disorders
2	Diarrheal diseases	2	Ischemic heart disease
3	Unipolar depressive disorders	3	Road traffic accidents
4	Ischemic heart disease	4	Cerebrovascular disease
5	HIV/AIDS	5	Chronic obstructive pulmonary disease
6	Cerebrovascular disease	6	Lower respiratory infections
7	Prematurity and low birth weight	7	Hearing loss, adult onset
8	Birth asphyxia and birth trauma	8	Refractive errors
9	Road traffic accidents	9	HIV/AIDS
10	Neonatal infections and other	10	Diabetes mellitus

Source: Based on data from WHO, 2008.

the projected disease burden for the world, there is a persistent burden of respiratory infections and diarrheal diseases in these regions. The situation in the low- and middle-income world is one where the "triple burden" of persistent communicable diseases, prevalent non-communicable conditions, and increasing injuries will call for an appropriate response from public health officials.

Burden of Disease Attributed to Risk Factors

An analysis of risk factors that underlie many important disease conditions can be useful for assisting policy decisions concerning interventions directed toward health promotion and disease reduction. Smoking, alcohol, hypertension, and malnutrition are risk factors for a variety of diseases, for example, and specific interventions have been developed that may reduce their prevalence. Risk factors include an array of human behaviors, nutritional deficiencies and excesses, substance abuse, and certain characteristics such as hypertension. Some factors are both an outcome and a risk factor (hypertension), some are challenging to measure (violence), and yet others (smoking and alcohol) lead to many disease outcomes. The linkage between an identified risk factor and the set of associated disease outcomes may be difficult to directly quantify, and the portion of specific disease prevalence attributable to any one factor may be problematic.

Relationships such as those shown in Figure 1-16 require careful assessment to determine the proportion of heart disease to be attributed to hypertension in relation to other interacting causal factors. The best way to determine the portion of disease that may be ascribed to hypertension is through randomized trials with careful assessment of disease outcomes over time: Results from studies that control hypertension have shown a reduction of death and disability from not only cardiac disease, but also from cerebrovascular and renal diseases.

Because the most important purpose of risk factor analysis is to assist in decision making about the allocation of resources, the link between the risk factor and the potential intervention to reduce the risk should be clear. The effectiveness of interventions against risk factors ultimately should be judged by their ability to reduce the amount of healthy life lost attributed to the diseases that the risk factor affects. For the evaluation of an intervention that reduces

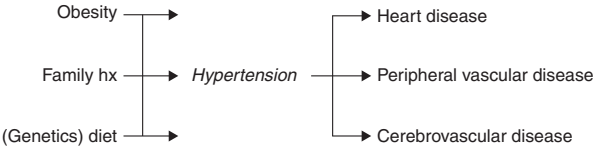


Figure 1-16 Flowchart.

hypertension, for example, the healthy life losses from the entire range of diseases that hypertension influences are therefore required.

Although understanding the underlying factors that lead to disease and the complex interrelations in the web of causation has long been a major focus of epidemiology (http://www.springerlink.com/content/n3mcyxyce7vqn66b/), most analyses of the relationships of risk factors to specific diseases have been done in the context of individual risk factors in limited settings and with wide variations in the criteria for risk assessment. As a consequence, comparisons of risk factors as determinants of disease on a population health level are problematic. The Comparative Risk Assessment (CRA) project of the GBD 2000 study carried out a systematic evaluation of 22 selected risk factors relative to global and regional burdens of disease using a specific model for analysis (Murray et al., 2001); it was updated in 2009 (WHO Global Health Risks) with data for 2004 (Mathers, 2009).

The Burden of Selected Major Risk Factors

The model used in CRA for causal attribution of health outcomes was based on counterfactual analysis that would result in the lowest population risk (Ezzati et al., 2002). Within this analysis, the contribution of one or a group of risk factors to disease or mortality was estimated by comparing the current or future disease burden with the levels that would be expected under an alternative hypothetical scenario. This involved an evaluation of the effect a risk factor has on the disease and its consequences, by setting the risk factor to its minimum while keeping all other factors constant. This method has the advantage of showing the potential gains by risk reduction from all levels of suboptimal exposure in a consistent way across risk factors (Ezzati et al., 2002).

The WHO Global Health Risks (2009) described 24 risk factors that are responsible for 44% of global deaths and 34% of DALYs. As shown in Figure 1-17(a), the five leading risks for mortality globally are high blood pressure (responsible for 13% of deaths globally), tobacco use (9%), high blood glucose (6%), physical inactivity (6%), and overweight and obesity (5%). These five factors especially increase risks for heart disease and cancer and have major consequences for countries across all income groups. In contrast, as Figure 1-17(b) shows, the main risks for burden of disease (DALYs) globally are underweight (6% of global DALYs), unsafe sex (5%), alcohol use (5%), and unsafe water, sanitation, and hygiene (4%); underweight, unsafe sex, and unsafe water, sanitation, and hygiene all contribute to infectious disease and overwhelmingly affect low-income countries. Alcohol abuse is largely a problem for men, but those mainly affected vary greatly by geographic region: This factor has its greatest impact on men in Africa, in middle-income countries in Latin America, and in a few high-income countries (e.g., Russia).

Eight risk factors—alcohol use, tobacco use, high blood pressure, high body mass index, high cholesterol, high blood glucose, low fruit and vegetable intake, and physical inactivity—account for more than 75% of ischemic heart disease (the leading cause of death worldwide) and 61% of total cardiovascular deaths. Although these major risk factors are associated with high-income countries, in fact more than 84% of the total global burden of disease that they cause occurs in low- and middle-income countries. Reducing exposure to these eight risk factors would increase global life expectancy by almost five years.

Globally, micronutrient deficiencies (including deficiencies of vitamin A, iron, and zinc), suboptimal breastfeeding, and preventable environmental risks account for more than four million deaths (nearly 40% of the 10.4 million under-five children who died in 1994); these deaths are readily preventable. In 1994, 82% of these deaths occurred in Africa and Southeast Asia.

Depression is the leading cause of years lost due to disability; rates of this disease are 50% higher in women than in men. Conversely, men aged 15 to 60 have a much higher risk of dying than do women of the same ages. The main causes of death that result in this differential are injury, particularly from violence and armed conflict, and heart disease.

Some risk factors may have few effects on the total global burden of disease, yet be very important locally within certain populations and regions. For example, iodine deficiency affects certain low- and middle-income countries and results in substantial disability in those populations.

The reasons for the demographic and epidemiologic transitions discussed earlier in this chapter and in Chapter 3 are largely related to shifts in these major risk factors as a result of changing social, economic and political trends and their complex interactions. Low-income countries continue to struggle against the high burdens of infectious diseases, malnutrition (including undernutrition and micronutrient deficiencies), and maternal and child health problems; at the same time, they must deal with the additional burdens of high levels of noncommunicable disease and injuries. The 2009 WHO Global Health Risks report estimates that had the risks analyzed in the report not existed, global life expectancy

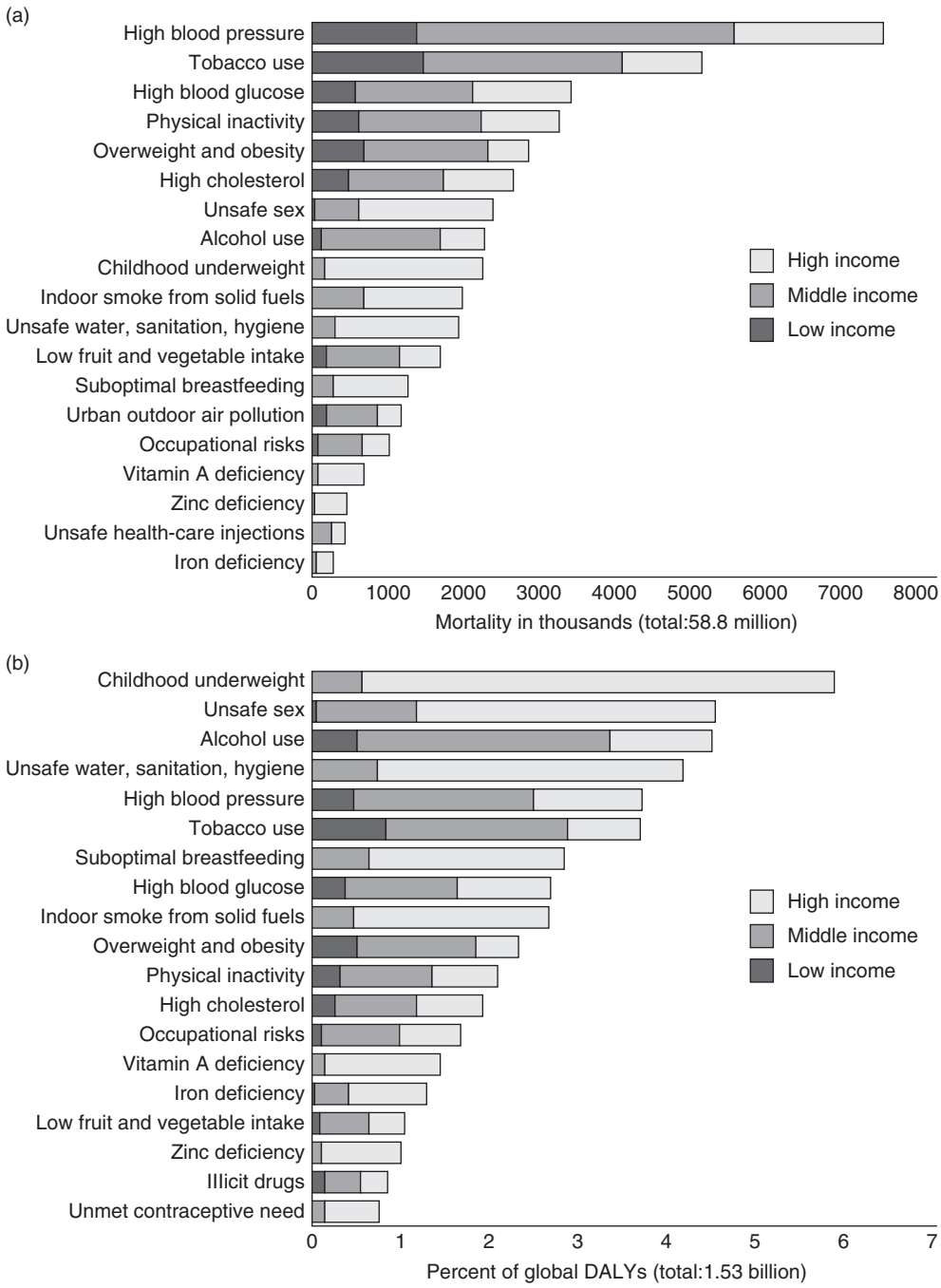


Figure 1-17 Mortality (a) and DALYs (b) due to 19 leading risk factors by country income level, 2004. *Source*: WHO Global Health Risks, 2009, Figures 6 and 7.

would be 10 years longer. This finding largely confirms the hopes first expressed in the 2000 Report to WHO by the Commission on Health and Macroeconomics (Commission on Macroeconomics and Health, 2001) and detailed with the MDGs (United Nations Millennium Declaration, 2000)—namely, that major improvements in the health status of those in low- and middle-income countries can be achieved:

> For example, reducing the burden of disease is possible since many cost-effective interventions are known, and prevention strategies can be transferred between similar countries.

Conclusion

The health of populations is the fundamental concern of global public health. The first step in the pursuit of population health improvement is the measurement of health and disease. Measurement is required to establish the magnitude of disease problems, define causal factors, explore potential solutions, and determine the impact of interventions. Measuring the impact of diseases on populations in terms of mortality and morbidity and their consequences is essential for planning effective ways to reduce the burden of illness and for setting priorities.

The burden of disease in populations has been gauged in many ways. Examples include measures of mortality, such as infant mortality rates; demographic measures, such as expectation of life at birth; and measures of morbidity, such as days away from work. However, for purposes of comparison among populations and for assisting in health planning and resource allocation, a common denominator combining these factors is needed. Summary measures of population health based on the amount of healthy life lost from disability and from death have been developed to serve that purpose.

Composite indicators (such as HeaLYs and DALYs) use duration of time (years, weeks, days) to measure the loss of healthy life from disease and the gain from interventions. These metrics are evolving into important tools for assisting health-related decision making. Nevertheless, to avoid misuse of such indicators, it is critical for those using them to understand the underlying assumptions and limitations and to meet the rather formidable data requirements. These summary measures also could be used to examine the burden of disease among subpopulations defined by socio-cultural-economic attributes and especially on vulnerable groups. Thus they could be used for ensuring that health-related decisions consider equity as well as cost-effective criteria.

Trends in disease burden provide important clues about the success of ongoing health programs and the need for development of new interventions. At the same time, they reflect non-health-related factors that are important to the production or maintenance of health in populations. Intercountry and interregional comparisons allow for measuring progress among nations; they can highlight inequalities in health status and examine these disparities in relation to social, economic, educational, and other factors.

Health systems across the world are greatly affected by ongoing changes in disease profiles and population dynamics. These systems must develop the capacity to respond to such changes effectively within the resources of each nation. Decisions must be based on evidence about the patterns of diseases, their risk factors, and the effectiveness of alternative interventions. Timely collection and analysis of appropriate, high-quality data to support such evidence are a prerequisite for improving equitable global health development.

Discussion Questions

1. How can data help achieve the main purpose of a health system in any country? Give examples.
2. What are the essential elements of a burden of disease assessment, and which types of data are the most challenging to obtain in a low-income country?
3. What are the relative strengths and weaknesses of summary measures such as HeaLYs and DALYs compared to more traditional indicators of disease burden such as infant or maternal mortality?
4. In your country or city, what would be the most appropriate set of indicators to assess the impact of chronic diseases on the population? Why?

References

Anand, S., & Ranaan-Eliya. (1996). *Disability adjusted life years: A critical review*. Working Paper No. 95-06. Harvard Center for Population and Development Studies.

Anker, M., Black, R. E., Coldham, C., Kalter, C. D., Quigley, M. A., Ross, D., et al. (1999). *A standard verbal autopsy method for investigating causes of death in infants and children*. WHO/CDS/CSR/ISR/99.4.

Barendregt, J. J., Bonneux, L., & Van Der Maas, P. J. (1996). DALYs: The age weights on balance. *Bulletin of the World Health Organization, 74*, 439–443.

Barker, C., & Green, A. (1996). Opening the debate on DALYs. *Health Policy and Planning, 11*(2), 179–183.

Barnum, H. (1987). Evaluating healthy days of life gained from health projects. *Social Science and Medicine, 24*, 833–841.

Blumstein, J. (1997). The Oregon experiment: The role of cost–benefit analysis in the allocation of Medicaid funds. *Social Science and Medicine, 45*, 545–554.

Bobadilla, J. L. (1998). *Searching for essential health services in low and middle income countries*. Washington, DC: Inter American Development Bank.

Burden of Disease Research Unit. (2003). *Initial burden of disease estimates for South Africa, 2000*. South Africa: South African Medical Research Council. Retrieved from http://www.mrc.ac.za/bod/bod.htm

Busschbach, J. J. V., Hesing, D. J., & de Charro, F. T. (1993). The utility of health at different stages of life: A qualitative approach. *Social Science and Medicine, 37*(2), 153–158.

Coale, A. J., & Demeney, P. (1983). *Regional model life tables and stable populations*. New York: Academic Press.

Coale, A. J., & Guo, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index, 55,* 613–643.

Commission on Macroeconomics and Health. (2001). *Macroeconomics and health: Investing in health for economic development*. Geneva, Switzerland: World Health Organization.

Concha M, Aguilera X (1996). *Burden of disease in Chile*. Santiago (Chile): Government of Chile Ministry of Health (MINSAL); 1996.

Crooper, M., & Kopits, E. (2003). *Traffic fatalities and economic growth*. World Bank Policy Research Working Paper 3035. Washington, DC: World Bank.

Eddy, D. (1991). Oregon's methods: Did cost-effectiveness analysis fail? *Journal of the American Medical Association, 266*, 2135–2141.

Evans, J.R., Hall, K.L., & Warford, J. (1981). Health care in the developing world: Problems of scarcity and choice. *New England Journal of Medicine, 305, 1117–1127*.

Ezzati, M., Lopez, A., Vander Hoorn, S., Rodgers, A., Murray, C. J. L., & Comparative Risk Assessment Collaborative Group. (2002). Selected major risk factors and global and regional burden of disease. *Lancet, 360*(9343), 1347–1360.

Ghana Health Assessment Team. (1981). A quantitative method for assessing the health impact of different diseases in less developed countries. *International Journal of Epidemiology, 10*, 73–80.

Gold, M. R., Siegel, J. E., Russel, L. B., & Weinstein, M. C. (Eds.). (1996). *Cost- effectiveness in health and medicine*. New York: Oxford University Press.

Hill, K. (2001). Demographic techniques: Indirect estimation. In *International Encyclopedia of Social and Behavioral Sciences* (p. 3461). Oxford, UK: Elsevier Science.

Hill, K. (2003). Adult mortality in the developing world: What we know and how we know it. *Training Workshop on HIV/AIDS and Adult Mortality in Developing Countries*. New York: United Nations Secretariat.

Hyder, A., & Morrow, R. (1999). Steady state assumptions in DALYs: Effect on estimates of HIV

impact. *Journal of Epidemiology and Community Health, 53*, 43–45.

Hyder, A. A., & Morrow, R. H. (2000). Applying burden of disease methods in developing countries: A case study from Pakistan. *American Journal of Public Health, 90*(8), 1235–1247.

Hyder, A. A., Rotllant, G., & Morrow, R. H. (1998). Measuring the burden of disease: Healthy life years. *American Journal of Public Health, 88*, 196–202.

Jamison, D.T., Brenan, J.G., Measham, A.R., Alleyne, G., Claeson, M., Evans. D.B., et al. (2006). Disease Control Priorities in Developing Countries. New York: Oxford University Press.

Kaplan, R.M. (1990). The general health policy model: An integrated approach. In B. Spilker (Ed.), *Quality of life assessment in clinical trials* (p. 156). New York: Raven Press.

Last, J. M. (Ed.). (2000). *A dictionary of epidemiology* (4th ed.). New York: Oxford University Press.

Lopez, A., Ahmad, O., Guillot, M., Inoue, M., Fergusson, B., Salomon, J., Murray, C. J. L., & Hill, K. (2002). *World mortality in 2000: Life tables for 191 countries*. Geneva, Switzerland: World Health Organization.

Lopez, A, Mathers, C, Ezzati, M, Jamison, D., & Murray, C. (Eds.). (2006). *Global Burden of Disease and Risk Factors*. The World Bank and Oxford University Press, New York.

Mahapatra, P. (2001). *Estimating national burden of disease: The burden of disease in Andhra Pradesh 1990s*. Hyderabad: Institute of Health Systems. Retrieved from http://www.ihsnet.org.in/BurdenOfDisease/APBurdenofDiseaseStudy.htm.

Mathers, C., Stevens, G., & Mascarenhas, M. (2009). *Global health risks: Mortality and burden of disease attributable to selected major risks*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/healthinfo/global_burden_disease/global_health_risks/en/index.html.

Mathers, C., Vos., T., Lopez, A., Salomon, J., Lozano, R., & Ezzati, M. (Eds.). (2001). National burden of disease studies: A practical guide (Edition 2.0). Geneva, Switzerland: World Health Organization.

Mathers, C., Vos, T., & Stevenson, C. (1999). *The burden of disease and injury in Australia*. Australian Institute of Health and Welfare. Retrieved from http://www.aihw.gov.au/

Morrow, R. H., & Bryant, J. H. (1995). Health policy approaches to measuring and valuing human life: Conceptual and ethical issues. *American Journal of Public Health, 85*, 1356–1360.

Murray, C. J. L., & Chen, L. C. (1992). Understanding morbidity change. *Population and Development Review, 18*(3), 481–503.

Murray, C. J. L., & Lopez, A. D. (1994). *Global comparative assessments in the health sector*. Geneva, Switzerland: World Health Organization.

Murray, C. J. L., & Lopez, A. D. (Eds.). (1996). *Global health statistics 1990*. Geneva, Switzerland: World Health Organization.

Murray, C. J. L., & Lopez, A. D. (1999). On the comparable quantification of health risks: Lessons from the Global Burden of Disease study. *Epidemiology, 10*, 594–605.

Murray, C. J. L., Lopez, A. D., Mathers, C. D., et al. (2001). *The Global Burden of Disease 2000 project: Aims, methods and data sources*. Geneva, Switzerland: Wohrld Health Organization.

Murray, C. J. L., Salomon, J., Mathers, C., & Lopez, A. (2002). *Summary measures of population health: Concepts, ethics, measurement and applications*. Geneva, Switzerland: World Health Organization.

New Zealand Ministry of Health. (2001). Burden of disease and injury in New Zealand. Retrieved from http://www.moh.govt.nz/moh.nsf/

Nord, E. (1992). Methods for quality adjustment of life years. *Social Science and Medicine, 34*, 559–569.

Nord, E. (1993). Unjustified use of the quality of well being scale in priority setting in Oregon. *Health Policy, 24*, 45–53.

Omran, A. (1971). The epidemiologic transition: A theory of the epidemiology of population change. *Milbank Memorial Fund Quarterly, 49*, 509–538.

Petiti, D. B. (1994). *Meta-analysis, decision analysis and cost-effectiveness analysis: Methods for quantitative synthesis in medicine*. New York: Oxford University Press.

Robine, J. M. (1994). *Disability free life expectancy trends in France: 1981–1991: International comparisons*. Chapter 2 In C. Mathers et al. (Eds.), *Advances in health expectancies*. Canberra, Australia: Australian Institute of Health and Welfare.

Thatte, N., Kalter, H. D., Baqui, A. H., Williams, E. M., & Darmstadt, G. L. (2009). Ascertaining causes of neonatal deaths using verbal autopsy: Current methods and challenges. *Journal of Perinatology: Official Journal of the California Perinatal Association, 29*(3), 187–194.

Torrence, G. W. (1986). Measurement of health state utilities for economic appraisal: A review. *Journal of Health Economics, 5*, 1–30.

UNAIDS. (2008). *Report on the global AIDS epidemic 2008*. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS.

UNAIDS. (2009). *AIDS epidemic update December 2009*. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS.

United Nations Children's Fund. (2009). *The state of the world's children 2008*. New York: UNICEF.

United Nations Millennium Declaration. (2000, September 18). Resolution adopted by the General Assembly. 55/2.

Weinstein, M. C., Siegel, J. E., Gold, M. R., Kamlet, M. S., & Russell, L. B. (1996). Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *Journal of the American Medical Association, 276,* 1253–1258.

World Bank. (1993). *World development report 1993: Investing in health*. New York: Oxford University Press.

World Health Organization (WHO). (1980). *International classification of impairments, disabilities and handicaps: A manual of classification relating to the consequences of disease*. Geneva, Switzerland: Author.

World Health Organization (WHO). (1992). *International statistical classification of diseases and related health problem (ICD-10): Tenth revision*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2000). *The world health report 2000*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002a). *World report on violence and health*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2002b). The international classification of functioning, disability and health: Introduction. Retrieved from http://www.who.int/classifications/icf/en/

World Health Organization (WHO). (2003). *The world health report 2003 – Shaping the Future*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2004). *The world health report 2004 – Changing History*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2008). *The global burden of disease: 2004 update*. Geneva, Switzerland: Author.

World Health Organization (WHO). (2009). *Global health risks: Mortality and burden of disease attributable to selected major risks*. Geneva, Switzerland: Author.

Zeckhauser, R., & Shephard, D. (1976). Where now for saving lives? *Law and Contemporary Problems, 40*(b), 5–45.