CHAPTER

1

Measures of Health and Disease in Populations

ADNAN ALI HYDER 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. For example, distinctions may be made between sickness and illness, but for purposes of defining and measuring disease burden a general definition will be used here. Disease is anything that a population (or an individual) experiences that causes, literally, "disease"; that is, anything that leads to discomfort, pain, distress of all sorts, disability of any kind, or death, constitutes disease from whatever cause, including injuries or psychiatric disabilities.

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, understanding the pathogenesis of the disease process and defining disease are critical for understanding and classifying causes in order 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, 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 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 the Global Burden of Disease studies making use of 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 has four sections: The first discusses the reasons for and approaches to measuring disease burden, the reasons for using quantitative indicators, and the importance of using data for decision making in health. The second section is a critical review of methods for developing and using composite measures combining 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 future trends of selected countries and regions, as well as 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.

Acknowledgments: The authors acknowledge Omar H. Amach for his assistance in the preparation of this chapter and for updating the necessary data and examples.

Reasons and Approaches for 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 would have the greatest 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 with time. In recent years, much has been made of the importance of evidence-based decisions in health. There is little reason to doubt that evidence is better than intuition, but that depends on how the evidence is used (Exhibit 1-1). This chapter examines evidence—the facts of health and disease—and demonstrates how to assemble this evidence so that it can best be used in assisting better decisions concerning health and disease.

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 type of people affected (gender, age).

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 of new cases of a disease that occur in a population at risk for developing the disease during a specified period of time. It is the proportion of people who develop new disease in a specific period of time. For this 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 1 year.

Incidence density, or often simply *incidence rate*, is the occurrence of new cases of disease per unit of person-time. This 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 by exposure to a toxin or any 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 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 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.

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 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.

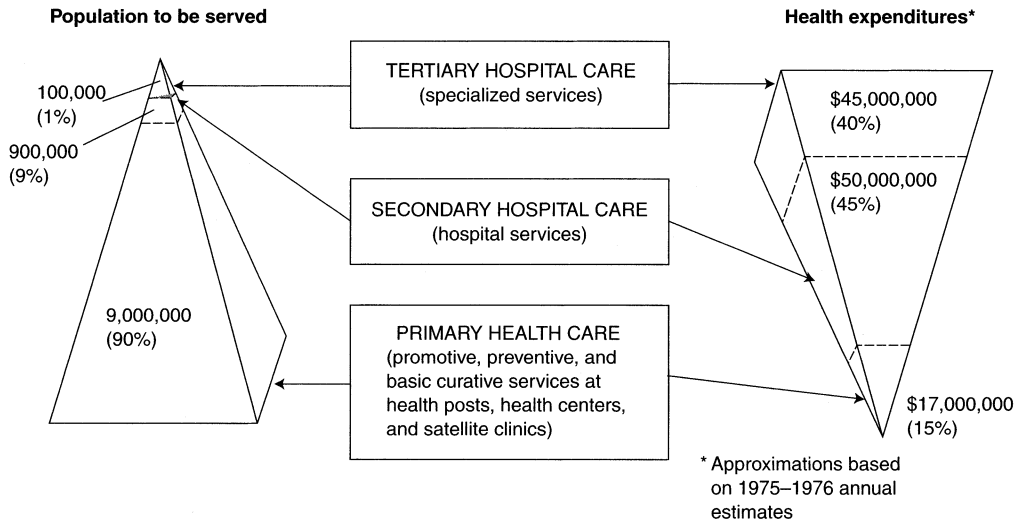
Mortality

Traditionally, mortality has been the most important indicator of the health status of a population. John

Exhibit 1-1 **Assessing a Health System with Data: Historical Example**

A well-documented historical example of the relationship between decision making and data can be seen in an assessment of the health status in Ghana in the mid-1970s. This case illustrates how able people with good intentions made decisions using established health system approaches, yet ignored good evidence gathered at the same time. A major reason for the failure to use available evidence was that it was not put forward in a form helpful to decision makers.

Assessment of the health status in Ghana in the late 1970s indicated that despite the remarkable increase in resources going into the health sector, the general health status of the population was still low. In the previous 10 years there had been little improvement in infant mortality rate, maternal mortality rate, and rate of communicable disease. That situation is strikingly illustrated in the accompanying figure. It shows how financial the resources of the nation were being allocated in reverse proportion to the numbers of people in need. The health system of Ghana could be likened to a pyramid, with the teaching hospital in Accra at the top and a network of health posts and dressing stations at the bottom. This was a system based on service delivery points and focused on buildings rather than on health services for the people. Such an emphasis on facilities creates false needs among the people for more facilities. Good health becomes synonymous with the availability of a doctor and a hospital rather than the enjoyment of a disease-free environment.



The health care dilemma in Ghana. The distribution of funds and personnel for primary health care compared with costly hospital-based care is in inverse proportion to the numbers of people that need to be reached. The health care pyramid for Ghana is upside down!

In response to this dilemma, the Health Planning Unit of the Ghana Ministry of Health developed a quantitative method for assessing the health impact of different diseases and for assisting in determining priorities for allocation of resources to alternative health programs. They used available data that were put in terms that had meaning for the decision makers—the gain in healthy life per dollar expended. This method was the first use of a common composite indicator that combined morbidity and mortality and could examine the gains in total healthy life per dollar expended on alternate health programs. At that time, the Ghana Ministry of Health had just developed the 1977–1981 health plan, which called for expanding and extending existing hospital and health center services. However, when the evidence was examined, an alternative community-based primary health care strategy that required equivalent expenditures was found to provide 20 times as much healthy life per dollar. This finding greatly strengthened the rationale for introducing a community-based primary health system in place of the hospital-centered system.

Source: Ministry of Health, Republic of Ghana, *A Primary Health Care Strategy for Ghana* (Accra, Republic of Ghana: National Health Planning Unit, 1978). Reprinted with permission.

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 variation 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), and paradigms such as the demographic transition are based largely on the decline of mortality in the under-five age group (Omran, 1971; Mosley et al., 1993).

The *fact* of death by age, sex, and place, which is required by law in most countries through *death registration*, and *cause of death*, as required by law in many countries through *death certification*, both provide essential information. Although death is a cardinal event and generally the most widely available kind of health information, in many low-income countries the fact of death, let alone cause of death, frequently is still not reliably available.

In high-income countries, vital statistics (i.e., the registration of births and deaths, usually 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 low-income countries, however, the collection of vital statistics remains highly incomplete, although improving. An analysis of death registration in the Global Burden of Disease 2000 study showed that complete or incomplete vital registration data together with sample registration systems cover 74% of global mortality. Survey data and indirect demographic techniques provide information on levels of child and adult mortality for the remaining 26% of estimated global mortality (Murray et al., 2001). Even in these countries, increasing use of survey methods provides estimates of the under-five and other mortality rates.

However, 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 under specific circumstances. Verbal autopsies have been used increasingly for judging likely cause of death. These can be quite useful for causes of death such as neonatal tetanus and severe diarrhea, but it has been found that sensitivity and specificity are limited for many diseases whose symptoms are variable and nonspecific (such as malaria).

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 (Feachem & Kjellstrom, 1992). Developing countries have higher rates of adult mortality than the high-income nations (Murray & Chen, 1992; Lopez et al., 2002), and 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 AIDS deaths in young and middle-aged women and men has had a profound impact on mortality and survival (see 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 thousand live births) and child mortality (under 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, 2004). These indicators have the added advantage of having been studied for their relationships with other indicators of the social and economic development of nations. There is a clear relation 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 these need to be examined carefully. For example, Sri Lanka and the Indian state of Kerala are both low-income regions that have low 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. Preston (1976) analyzed life tables for 43 national populations, including 9 developing countries, to develop cause-specific mortality profiles. In keeping with the 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 the WHO International Classification of Diseases

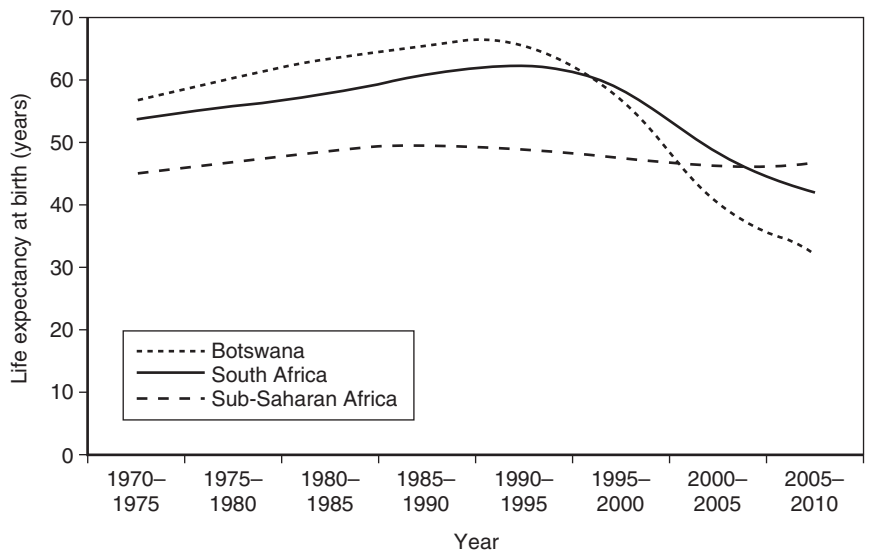
Exhibit 1-2 **Trends of the HIV/AIDS Epidemic**

Acquired immunodeficiency 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 a quarter of a century ago, today, an estimated 34 to 46 million people are living with HIV/AIDS (WHO, 2004). The most heavily burdened continent is Africa, which in 2003 was home to two-thirds of the world's people living with HIV/AIDS, but only 11% of the world's total population (WHO, 2004). However, developed countries are also afflicted. The Russian Federation and Ukraine, along with other countries in eastern Europe and countries in central Asia, have the most rapidly expanding HIV epidemics (WHO, 2003).

In 2002 (latest data available), of the leading causes of disease burden among men and women aged 15 years and over, HIV/AIDS was the number one cause for males and the second leading cause for females, accounting for around 6% of the global burden of disease (see accompanying table). In terms of mortality, despite declining global trends of communicable disease burden in adults, HIV/AIDS has become the leading cause of mortality among adults aged 15 to 59 years. Nearly 80% of the almost 3 million global deaths from HIV/AIDS in 2002 occurred in sub-Saharan Africa (see accompanying figure) (WHO, 2003).

Global Summary of HIV and AIDS Epidemic			
Number of People Living with HIV	**Total**	**37.8 million**	**[34.6–42.3 million]**
	Adults	35.7 million	[32.7–39.8 million]
	Women	17 million	[15.8–18.8 million]
	Children	2.1 million	[1.9–2.5 million]
Number Newly Infected with HIV	**Total**	**4.8 million**	**[3.6–5.6 million]**
	Adults	4.1 million	[4.2–6.3 million]
	Children	630 000	[570 000–740 000]
AIDS Deaths in 2003	**Total**	**2.9 million**	**[2.6–3.3 million]**
	Adults	2.4 million	[2.2–2.7 million]
	Children	490 000	[440,000–580,000]

Source: From Joint United Nations Programme on HIV/AIDS, *Global Summary of the HIV and AIDS Epidemic in 2003* (Geneva, Switzerland: UNAIDS, 2003). Reprinted with permission.



Trends in Life Expectancy in Sub-Saharan Africa and Selected Countries, 1970–2010.

Source: United Nations Population Division, *World Population Prospects: The 2002 Revision Population Database*, http://esa.un.org/unpp. Reprinted with permission.

(ICD) has been used widely in many countries (WHO, 1992). Despite a standardized process for categorizing deaths, variations in the reliability of these data occur because of variations in the training and expertise of people coding cause of death, as well as the supervision 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. The *mortality rate* is a form of incidence and 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. The *case fatality ratio* is the proportion of those with a given disease who die of that disease (at any time, unless specified). The mortality rate is equal to the case fatality ratio multiplied by the incidence rate of the disease in the population.

The distinction between the *proportion* of deaths attributable to a set of causes versus the *probability* of death from these causes is important to understand. For example, the probability of death 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 attributable to these chronic causes is less than those attributed to infectious causes. The risk of death and the rates of death by these causes do not increase; rather, the proportion of attributable deaths increases as the communicable proportion declines with development. For example, in 1990 the risk of dying from cancer was 50% greater in low- and middle-income countries than in high-income countries, even though cancer accounted for a much smaller proportion of total deaths in those countries.

Demographic and Epidemiologic Transitions

The term *demographic transition* was first used by F. W. Notestein in 1945 to describe the changes in birth and death rates that historically have accompanied the shift from a traditional to a modern society (Exhibit 1-3) (see also Chapter 3). With modernization (a complex term indicating social and economic development), sharp declines in mortality have been followed by a reduction in fertility, although usually lagging by years or decades. The term *transition* refers to the shift away from a stable, high-stationary stage of population in which very high birth rates are balanced by very high death rates and there is little or no population growth. As a society undergoes modernization, there is a transition with falling mortality, especially in the under-five age group, but with continuing high birth rates leading to explosive population growth. Birth rates then tend to drop, and a new, low-stationary stage is reached in which birth and death rates are low and balance resumes. The end results are 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).

Historically, all countries that have undergone modernization with a marked drop in under-five mortality have had rapid population growth. In the past, this population growth was always followed by falling fertility rates, but the reasons for the drop are not entirely clear. Maurice King has pointed to a potential major problem that may arise, termed the demographic trap, in which fertility rates do not drop (King, 1990). This situation would lead to the classic Malthusian scenario in which massive starvation and epidemic diseases overtake the population. King points out that there is no guarantee that there will be a drop in birth rate in all countries undergoing modernization and that changes in fertility depend very much upon social and cultural characteristics.

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. It is important to note that 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 much 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 under-five 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 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 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.

Morbidity and Disability

Measures of mortality have been the principal indicators of population health status for a long time. Their relative ease of observation, presence of data, and history of use make them suitable for assessing health status and consequent changes. However, the problem with mortality-based indicators is that they "note the dead and ignore the living" (Kaplan, 1990). Measurements of morbidity, on the other hand, are much more problematic because there is no clearly defined endpoint, such as death provides. In addition, several components of morbidity and disability need to be assessed: duration, severity, and consequences.

Concepts that distinguish among disease, illness, and sickness have been present in the literature for half a century, from the description of the sick role in 1929 to the development of a disability framework in the 1960s. This framework considered *disease* as an organic-level disorder confined to the individual, *illness* as a subjective state of dysfunction from the disorder at the individual level, and *sickness* as a social dysfunction within a society that goes beyond the individual. The International Classification of Impairments, Disabilities, and Handicaps (ICIDH) was developed by WHO to classify nonfatal health outcomes (WHO, 1980). This assessment was based on a progression from disease to handicap and was analogous to the ICD series. 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 sickness. Recently, WHO has built on the ICIDH and developed the International Classification of Functioning, Disability and Health (ICF), which is a classification of health that describes body functions and structures, activities, and participation (WHO, 2002a).

Using such classifications, indicators for disability, such as impairment-, disability-, 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 those hospitalized, 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

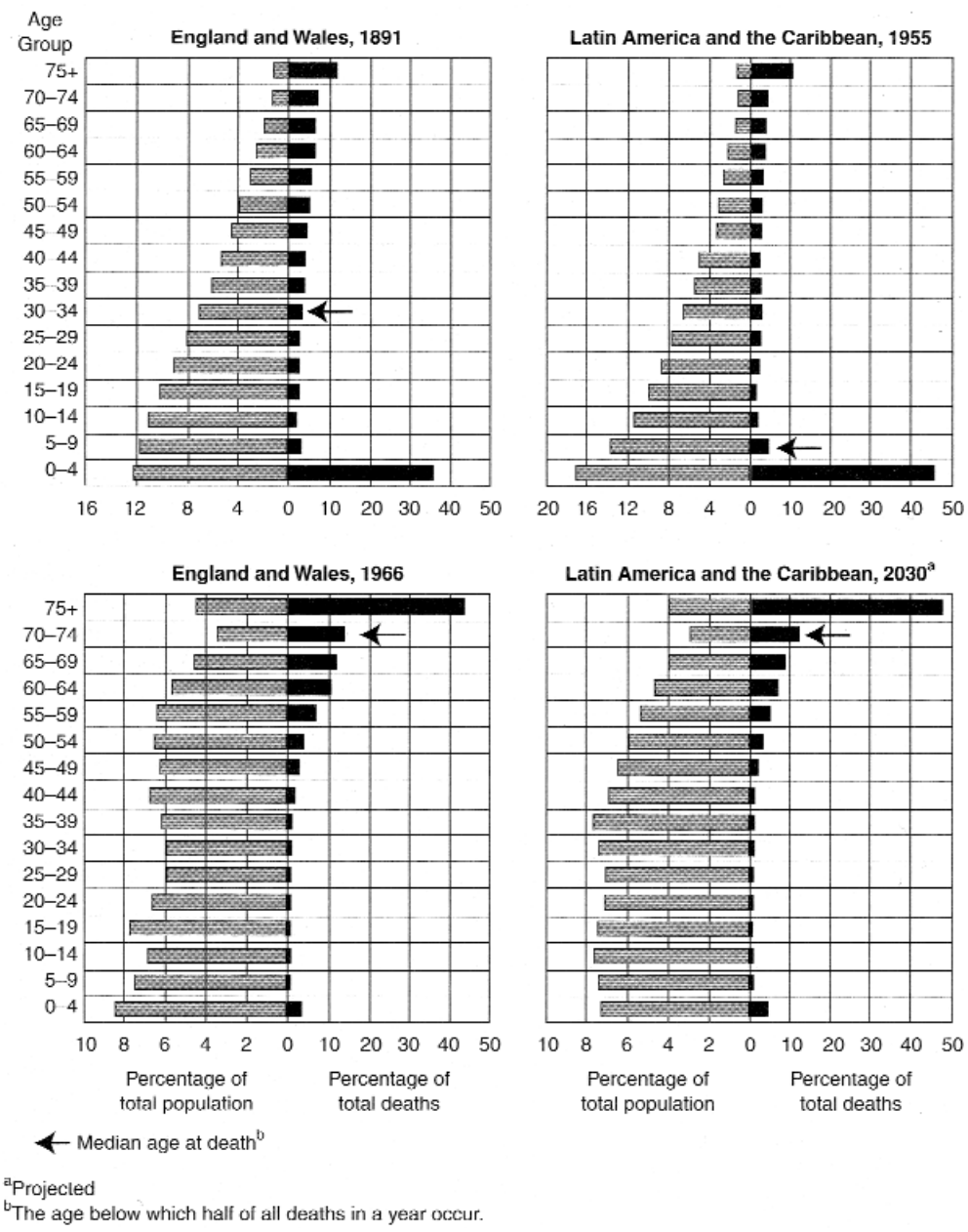
Exhibit 1-3 **The Demographic and Epidemiologic Transitions**

Changes in the pattern of disease proceed in two steps. The first is the demographic transition, when mortality from infectious disease and undernutrition decline with a marked drop in under-five mortality plus a reduction in fertility; the second is the epidemiologic transition, with a change in disease pattern. The population grows older, and noninfectious diseases become the main causes of ill health. Health patterns in the developing world over the next three decades will be profoundly influenced by these transitions.

It is commonly assumed that when a country is going through its demographic transition, the changes in its health indicators are primarily a function of declines in mortality. In fact, both the age structure and the cause of death structure are strongly influenced by the rapid decline in fertility. When fertility is high, the age structure of a population is heavily skewed toward the young. Because birth rates remain high and larger numbers of women enter the reproductive ages every year, the base of the population is continually expanding. When birth rates start to fall rapidly, the absolute number of babies born each year may remain unchanged or even decline. The graphs that follow show the shape of the age structure of the population on the left and the percentage of total deaths in 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 under-fives to the elderly.

(continued)

Exhibit 1-3 **Continued**



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

Source: A. A. Hyder, G. Rotllant, and R. H. Morrow, "Measuring the Burden of Disease: Healthy Life Years," 1998, *American Journal of Public Health, 88,* pp. 196–202. Reprinted with permission.

is 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 and often provide only the patient's chief complaint and probably the treatment dispensed. The main value of most such records is limited to establishing the fact of us-

ing a facility. There are usually strong biases in terms of those 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 health care facilities, functional disability (measures of activity that is less than usual), and time spent away from work (absenteeism, work days lost) have been used to assess the magnitude of morbidity from various conditions. A common 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). There is considerable literature on a wide variety of instruments used to measure such functional capacity, especially in clinical medical literature, that is not directly useful for population-based morbidity assessment.

Data about morbidity presented in the literature are often based on self-perceived or observed assessments, and frequently 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 variation between reported and measured prevalences 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, often interpreted to indicate that wealthy individuals and low-mortality populations *report* higher rates of morbidity (Murray & Lopez, 1996a).

Measurement of individual preferences for different health states in order to determine relative severity of disability has been done by a variety of methods (Torrence, 1986; Kaplan, 1990; Murray et al., 2002). Factors that influence preference include the type of respondent, the type of instrument used to measure the response, and the time from entry into the disabled state. Individuals who are in a particular state, healthy individuals, health care 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 and thus affect the value of that state to the individual. Healthy people may have different valuations for health states than people who are disabled, and the valuation by the disabled may change depending on time and adaptive processes. 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, 1994b).

Instruments used to extract such preferences involve visual and interview techniques (Torrence, 1986; Murray & Lopez, 1996a). Two alternative scenarios are often presented to the subject and the point of indifference sought (as in standard gamble techniques). Despite much work, there is no consensus or accepted standard method. 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, 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. There has been repeated concern about and work on their reliability and validity. These measures are not discussed further, because they have been primarily used in clinical assessments 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 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 disability from various causes 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 expressing extent of disability.

In general, three components of morbidity 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 since, by the definition of disease given earlier, they will have signs or symptoms. However, 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 trait), the CDR is likely to be less than 1.

The second component is the extent or severity of disability—how incapacitated the person is as a result of the disease. The extent of disability is expressed on a scale, usually from 0, which means no disability, to 1, which is equivalent to death. The assessment of severity can be quite subjective, particularly because there are so many different types and dimensions of disability. A number of methods have been tried to achieve comparability and obtain consensus (Murray et al., 2002). 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 1-1). 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 1-1). 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 go to much finer distinctions have been equivocal. The need to become more refined for purposes of health program decisions is a national or local decision.

The third component is the duration of the disability. The duration is generally counted from onset until cure and recovery or death. Sometimes there is continuing permanent disability after the acute phase is completed, and thus 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 is an important feature. This level of precision needs to be guided by the purpose of collecting the information. The ultimate reason for data is to guide decision making—to make it better and more efficient at helping reduce the burden of disease on populations. The level of precision

Table 1-1 Examples of Disability Classification Systems

Ghana Health Assessment Team, 1981 Couples at an increased risk:		
Class	**Severity**	**Equivalent to (max)**
1	0	Normal health
2	0.01–0.25	Loss of one limb's function
3	0.26–0.50	Loss of two limbs' function
4	0.51–0.75	Loss of three limbs' function
5	0.76–0.99	Loss of four limbs' function
6	1	Equivalent to death
Global Burden of Disease Study, 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's syndrome
6	0.50–0.70	Major depression, blindness, paraplegia
7	0.70–1.00	Psychosis, dementia, migraine, quadriplegia

Source: Data from Ghana Health Assessment Team, "A Quantitative Method for Assessing the Health Impact of Different Diseases in Less Developed Countries," 1981, *International Journal of Epidemiology, 10*, pp. 73–80; and C. J. L. Murray and A. Lopez (Eds.), *The Global Burden of Disease* (Geneva, Switzerland: World Health Organization, 1996).

depends on the decisions to be taken; even rough estimates may be helpful. Though disconcerting to some, the time, human, and monetary cost of further precision needs 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.

Composite Summary Measures of Population Health

This section focuses on the main approaches used for 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 approach, examines methodological differences, makes explicit the value choices that each entails, and outlines the advantages and limitations of each.

Rationale for Composite Measures

Rationing of health care resources is a fact of life everywhere; choices about the best use of funds for health must be made (World Bank, 1993; Hyder, Rotlland, & Morrow, 1998). The global scarcity of resources for health care is a challenge for every country, rich and poor (Evans, Hall, & Warford, 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 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, but the health sector has had no coherent basis for determining the comparative value of different health outcomes. To make decisions about whether to put money into programs that reduce mortality in under-fives, as compared with those that reduce disabling conditions in adults, a common denominator is needed. In recent years, work has been carried out to develop composite indicators combining morbidity and mortality into a single measure that may serve as a common denominator. The common unit of measure is time lost from healthy life.

The most important reason for attempting to capture the complex mix of incommensurable consequences resulting from disease into 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 must also be so expressed. *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 the use of these indicators must be carefully examined and looked at from all viewpoints. Not only are there problems of trying to put so many dimensions together, which inevitably leads to distortion, but also there are serious issues concerning the reliability and validity of the information on which these are based. Thus, all the problems associated with determining cause of death, counting the number of cases of disease, and assessing the extent of disability from a condition will lead to great uncertainties when they are added and multiplied together. The development of a single indicator with a specific number provides deceptive substantiality to what may be composed of fragile data. Continuing vigilance in how 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 what underlying assumptions and limitations are associated with these approaches. But with all 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 within and across populations. They can 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 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

Precursors of composite indicators have been discussed in the literature for decades and generally were developed to assist prioritization of health issues. Usually these 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 this to a dollar value. These measures are thus more economic measures than disease-burden measures.

Types of Composite Summary Measures

Two types of composite summary measure have been developed: *health gaps* (healthy life lost), such as healthy life years (HeaLY) or disability-adjusted life years (DALY), and *health expectancies*, such as disability-free life expectancy (DFLE) or health-adjusted life expectancy (HALE). Both types use healthy lifetime lost through disability and death as a common measure of the impact of mortality and nonfatal health outcomes. These two types of measures are complementary (Figure 1-1).

In Figure 1-1, 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; the thin line is the survivor curve of those in full health to each age *n*.

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 age *n*, when all died.

Health expectancies are summary measures that estimate expectancy of life in a defined state of health. Examples include disability-free life expectancy, active life expectancy, and health-adjusted life expectancy.

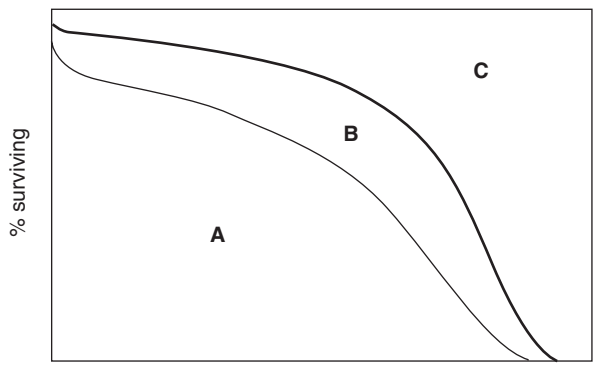


Figure 1-1 Survivorship Curve of a Hypothetical Population, Showing the Areas of Health Expectancies. *Source:* C. J. L. Murray et al., *Summary Measures of Population Health* (Geneva, Switzerland: World Health Organization, 1999). Reprinted with permission.

These 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 1-1, 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. (Full health has a weight of 1.)

Health gaps (healthy life lost) are summary measures that estimate the *difference* between actual population health and some specified norm or goal. In Figure 1-1 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 are between 0, or no disability (full health), and 1, or complete disability (equivalent of death). Note that this 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.

A major advantage of healthy life lost summary measures as compared with health expectancies is that they provide a common denominator for population health and for the outcomes in randomized trials and cohort and other health services studies, as well as for economic evaluations of interventions and monitoring of health system outcomes.

Composite Indicators

A number of composite summary indicators for burden of disease assessment have been developed. We shall focus on four: three of the health gap type—the healthy life year (HeaLY), the disability-adjusted life year (DALY), and the quality-adjusted life year (QALY)—and one of the health expectancy type, namely, the health-adjusted life expectancy (HALE).

In addition to measures of morbidity and mortality per se, these composite indicators all incorporate several social value choices either explicitly or implicitly: the choice of life expectancy tables, valuing future life as compared with present, 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"), 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

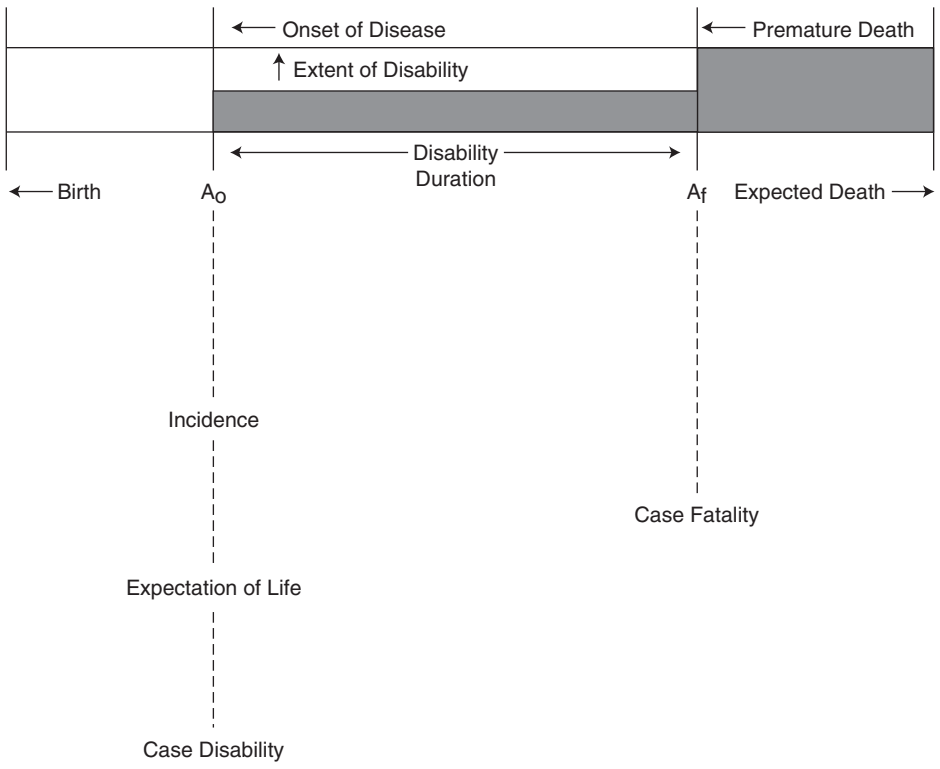
The *healthy life year* (HeaLY) is a composite measure that combines the amount of healthy life lost due to morbidity with that lost due to death (loss of life expected had the disease not occurred) (Hyder, Rotllant, & Morrow, 1998). We discuss the healthy life year first because it is conceptually straightforward, 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 incorporating several additional features.

The measure of loss from death is based on the years of life that would have been expected 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 what choice to make for the basis of 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 for Ghana, but considerations of equity as well as those concerning 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 much more challenging than measuring that from death, and many approaches have been used (Murray & Lopez, 1994b. In order for a measure to be used in a composite measure, it must have comparable dimensions to that for life lost due to death. The HeaLY includes three components: case disability ratio (CDR, comparable to the case fatality ratio), extent of disability, and duration of disability. The CDR and duration can be determined objectively, but assessment of the extent, which ranges from 0 to 1 (from no disability to that 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 1-2). For the purpose of estimating healthy life lost or gained, disease is defined as in the introduction of 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 with infection or some biological characteristic (such as AS hemoglobin) 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 (for example, 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.

The onset of disease usually will be dated from the start of symptoms or signs, as determined by the individual afflicted, a family member, a medical practitioner, or as the result of a lab test. There are several different patterns of disease evolution; Figure 1-3 illustrates healthy life lost from disability and premature death due to typical cases of cirrhosis, polio, and multiple sclerosis 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),



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

Figure 1-2 The HeaLY Model: Loss of Healthy Life from Disability and Death.

progression to another disease state (such as chronic hepatitis progressing to cirrhosis), and death. The latter includes death directly caused by the disease and death indirectly brought on by the disease as a result of disability.

The definition of variables and formulas to calculate HeaLYs are described later in this section and summarized in Table 1-2. 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 countries. However, like other choices in this method, if sensitivity testing indicates that the average age is not satisfactory, then estimates may be based on age distributions instead. 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 (like DALYs) is based on normative expectations of what should be achieved under optimal circumstances. Women in Japan, who have the highest global expectation of life, approximate this expectation, taken from the West model with an expectation of life at birth of 82.5 years for females (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 they have been labeled as diseased.

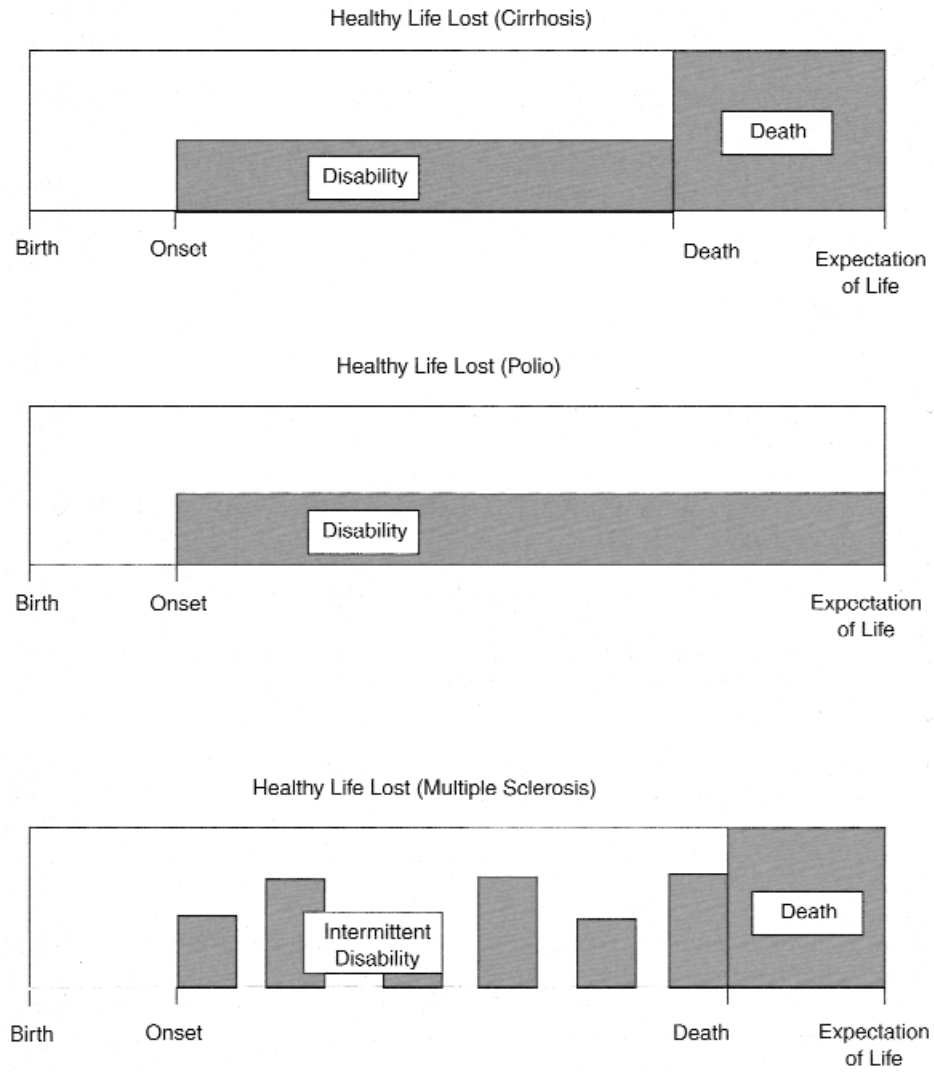


Figure 1-3 Different Patterns of Healthy Life Lost. *Source:* A. A. Hyder, G. Rotllant, and R. H. Morrow, "Measuring the Burden of Disease: Healthy Life Years," 1998, *American Journal of Public Health, 88,* pp. 196–202. Reprinted with permission.

However, there are some conditions, such as sickle cell trait or HIV positivity, and risk factors for which 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. If the disability is temporary, then D_t is the duration of that disability until recovery (see Table 1-2). 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 [$D_t = E(A_o)$]. On the other hand, if the disability is permanent and the disease 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 severity (see Table 1-1). These scores represent an estimate of the *average* disability suffered by typical cases of the specific disease over its course. The Ghana scale, for example (see Table 1-1), is simple and has been used for HeaLY calculations; similar scales may be developed in countries interested in doing burden of disease studies.

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 is a prospective view of

Table 1-2 Variables for Estimating Healthy Life Years

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

the event (disease onset) and its consequences because cases are followed 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 optimum 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. In a steady state, a random sample of 1,000 from such a population has the potential of 41,250 years of healthy life (Morrow & Bryant, 1995; Hyder, Rotllant, & Morrow, 1998). Each year this population would experience events 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 1-1. Discounting future life or adding productivity, dependency, or age weighting would affect these denominator numbers.

HeaLYs measure the *gap or loss* between the current situation in a country and an ideal or standard, as defined by the selected expectation of life. In recent work the standard used is 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 that due to disability would be counted. In other words, because the population under study is the ideal, and assuming stability of the population, 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 only 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 in its spreadsheet form is that the effects of different kinds of interventions can be readily explored to determine their expected gains in healthy life. Interventions may usefully be divided into two broad categories: those that are used to prevent the initiation of the disease process and those that are used to 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 effects of treatment strategies are to interfere with the natural history of the disease process, thereby reducing the case fatality and/or case disability ratios or extending life, providing a later age at death for conditions such as diabetes and AIDS. The spreadsheet, available upon request to the authors, also incorporates the proportion of the population that are or 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 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 (Murray & Lopez, 1994a, 1994b 1996a, 1999; Murray et al, 2002).

DALYs are calculated as two separate components and directly include three social value choices. The two components are (1) the loss of life from death, assessed by evaluating all deaths in a year for life expectancy at the time of death to estimate years of life lost (YLL) for each disease category, plus (2) the loss of healthy life from disability, estimated using years of life lived with disability (YLD) based on the incidence, the average duration of the condition (to remission or death), and a severity weight using an average health state weight. Thus,

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

The social value choices include life expectation tables, discount rates for future life, and weighting for life lived at different ages, as discussed later. The DALY is described in detail in Murray and Lopez (1996a). DALYs, with their separate components of YLLs and YLDs, were estimated for three disease categories, both sexes, five age groups, and eight regions of the world and published in the Global Burden of Disease study for the year 1990 (Murray & Lopez, 1996a) and for the year 2000 (Murray et al., 2001). Some of these figures are presented in the section entitled "Comparisons and Trends in Disease Burden."

The calculation for YLL uses the age distribution of deaths by cause for a year to estimate standard expected years of life for each disease. The loss of life is obtained by comparison with a model life table based on best achievable low levels of mortality, such as in Japan, reflecting high life expectation at birth of more than 80 years (Coale & Guo, 1989).

For disability, the DALY uses estimates of incidence, duration, and severity to calculate the time lived with disability across age groups. This information is based on the expectation of a proportion of cases in most conditions experiencing some form of disability over time. The onset of this disability, an estimate of severity at each stage, and the period of time spent in each stage are used to generate YLDs. A description of the severity scale used in one version of DALYs has been given previously in the section on measurement of disability (see Table 1-1).

Note that an important difference between the HeaLY and DALY is 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), as illustrated in Figures 1-1 and 1-2. 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, but when there is an increasing incidence, such as with HIV in many parts of the globe, the DALY approach can greatly understate the true situation (Hyder & Morrow, 1999).

Once the years of life lost to mortality and morbidity have been estimated, they are discounted, usually at 3% per annum. This social time preference has been used for most estimates; recently, DALY results discounted at 0% have also become available.

DALYs are age-weighted according to an arbitrary exponential curve designed to give the most value to life lived as a young adult Hyder, Rotllant, & Morrow, 1998; World Bank, 1993). Weighting by age was the most controversial component of the DALY and caused great dissent from other health professionals. See the section "Valuing Life Lived at Different Ages" later in this chapter. Recent DALY listings of Global Burden of Disease (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 the final result, but this depends on the purpose for making the estimates and has been challenged (Anand & Ranaan-Eliya, 1996; Barendregt, Bonneux, & Van Der Maas, 1996; Barker & Green, 1996; Hyder, Rotllant, & Morrow, 1998).

The calculation for DALYs can be expressed in the form of an integral that was first published in the World Bank literature (Murray & Lopez, 1994a). 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. However, 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 (Morrow & Bryant, 1995; Bobadilla, 1998; Hyder, Rotllant, & Morrow, 1998). Recent DALY formulations allow for this.

Quality-Adjusted Life Year

The quality-adjusted life year (QALY) was introduced in 1976 to provide a guiding principle for selecting among alternative tertiary health care interventions (Zeckhauser & Shephard, 1976). The idea was to develop a single measure of quality of life in order to compare expected outcomes from different interventions—a measure that relies on weighting the range of possible health states and the duration of time spent in each state to compute their equivalency with healthy life.

Since its introduction, a 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 and use QALYs (Kaplan, 1990; Nord, 1993). A central notion behind the QALY is that a year of life spent in one health state may be preferred over a year spent in another. Measures that focus exclusively on duration of life fail to reflect these types of value choices. Multiplying time spent in a particular health state with the value given to that state forms the basis of comparing outcomes from treatment options.

In general a quality-adjusted life year takes into account both quantity (duration or amount of time) and the quality of life generated by health care interventions. It is the arithmetic product of life expectancy and a measure of the quality of the remaining life years. A QALY places a weight on time in different health states. A year of perfect health is worth 1, whereas a year of less than perfect health is worth less than 1. Death is considered to be equivalent to 0; however, some health states may be considered worse than death and have negative scores. QALYs provide a common currency to assess the extent of the benefits gained from a variety of interventions in terms of health-related quality of life and survival for the patient. When combined with the costs of providing the interventions, cost-utility ratios can be constructed; these indicate the additional costs required to generate a year of perfect health (one QALY). Comparisons can be made among interventions, and priorities can be established based on those interventions that provide the most QALYs per net expenditure.

Particular effort has gone into researching ways in which an overall health index might be constructed to locate a specific health state on a continuum between 0 (death) and 1 (perfect health). It has been especially useful in distinguishing different types and levels of disability, impairment, and handicap. The use of QALYs is exemplified in the QALY Toolkit (Gudex & Kind, 1988) in which eight degrees of disability are combined with four levels of distress to categorize patients into one of 29 possible health states, with a weight assigned to each health state. Each s valued and compared with others, and QALYs are derived by summing the time and value product for the progress of an individual through these states as treatment is given. The nature of this construct allows the use of QALYs for individual decision making, with potential for application in policy decisions.

The QALY was not originally developed as an indicator of disease burden in a population but rather as a differentiating indicator for individual choices among tertiary health care procedures. It was used for assessment of individual preferences for different health outcomes from alternative interventions (Morrow & Bryant, 1995). But the idea has generated many alternative formulations and methods for assessment and has been put to a variety of purposes.

Perhaps the most important use of QALYs has been as a common denominator to measure utility in cost-utility analysis to assist in resource allocation among alternative health interventions (Torrence, 1986; Kaplan, 1990; Nord, 1992). *Cost-utility analysis*, in this instance, has been considered as a special form of cost-effectiveness analysis with a common unit of measure, the QALY, for gains by alternative health interventions (Torrence, 1986). Interventions can be ranked in terms of cost *per QALY*, and money can be allocated to those that have the lowest result. In general this approach is acceptable to and understood by health

policy makers, though there are notable exceptions (Exhibit 1-4).

The Health-Adjusted Life Expectancy

Health-adjusted life expectancy (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 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 in its *World Health Reports* (WHR) for annual reporting on population health (WHO, 2000).

During the 1990s, disability-free life expectancy (DFLE) and related measures were calculated for many countries (Mathers et al., 2001; Robine et al, 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. The term disability-adjusted-life year, or DALE, was replaced by health-adjusted life expectancy (HALE) for the WHR 2001 and will be used henceforth.

Health expectancy indices combine the mortality experience of a population with the disability experience. The HALE is calculated using the prevalence of disability at each age 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, this translates to a measure of the total disability burden in a population. Comparison of the

Exhibit 1-4 **Oregon: Application of the Quality-Adjusted Life Year**

The best-known application of 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 one of the first attempts worldwide to explicitly ration health services. A coalition including consumers, health care providers, insurers, and business and labor representatives launched a broad and bold reform. It began with a series of experiments in which the decision-making process was based on a cost-effectiveness approach using QALYs 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 x 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 ranked above hospitalization for starvation, and crooked teeth ranked above 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. Though enormous public effort went into the reform and much was accomplished, the explicit cost-effectiveness approach with QALYs as the outcome measure was dropped (Eddy, 1991; Blumstein, 1997; Morrow & Bryant, 1995; Nord, 1993).

The lack of success in developing a satisfactory list of services based on the QALY in Oregon was largely due to two factors: aggregation of scores appropriate for individual choices rather than using a population base, as for HeaLYs and DALYs, and a poorly tested quality of well-being scale. The result was the ranking of many conditions and their treatment much lower than the public as well as public health experts considered appropriate. Although unsuccessful in Oregon, such a method has the potential to be converted to a population-based approach.

various methods and specific indicators is available in the literature (Robine, 1994). Alternative methods are given in the National Burden of Disease Studies manual (Mathers et al., 2001).

As originally designed, this measure does not relate to specific diseases but rather to the average extent of disability among that proportion in 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.

Although the HALE is conceptually interesting and is now being calculated and included regularly in the WHO annual reports, it must be asked 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. The relative ranking of countries by health expectancy (HALE) is virtually identical with that of life expectation at birth.

Summary

Table 1-3 summarizes the main differences among these four summary measures in terms of origin, purpose, level of use, sources of data, and the disciplinary background of the measure's 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 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 in order to be comparable to unidimensional monetary expenditure units for costs.

To construct composite measures of population health, important social value choices must be made. Choices must be made about what 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; these choices 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

Table 1-3 Comparisons of Composite Summary Measures of Population Health

	Healthy Life Years (HeaLY)	Disability-Adjusted Life Years(DALY)	Quality-Adjusted Life Years (QALY)	Health-Adjusted Life Expectancy (HALE)
Origin	Ministry of Health, Ghana 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 on a comparable basis	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
Discipline Base	Epidemiologists, clinicians, national planners	Economists, statisticians	Economists, clinicians	Demographers, economists, statisticians
Social Values Incorporated	Future life discounted	Age weighting; future life discounted	Generally not included	Not relevant

occurred. To estimate the expectation of life in a population, a choice must be made between using a national or a 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 model life tables. On the other hand, a model life table can be selected to reflect the best health state possible, 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 in Ghana from a particular intervention with those in the United Kingdom, even if both costs and lives saved were the same in each country. The reason for this 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 view point 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 life tables (1983). The latter have been revised (Coale & Guo, 1989) and have been used in the Global Burden of Disease study and for HeaLYs (Murray & Lopez, 1996a; Hyder, Rotllant, & Morrow, 1998). 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 should be achievable elsewhere.

The choice of life table in a burden of disease analysis should be determined by the objective of the exercise, and the impact of the choice on the results can be explored in each situation.

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 immediately raises questions as to the basis for valuing human life. Is a day of anyone's life of the same value as that of anyone else? 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 lifewas equal to that of any other life. The valuing of a year of life equally, irrespective of age, has been considered egalitarian (Busschbach, Hesing, & de Charro, 1993; Morrow & Bryant, 1995). But the healthy life approach values individuals in direct proportion to their expectation of life at their current age. Therefore the loss of a healthy child is regarded as costing society more than the loss of a healthy adult.

DALYs assign an exponential function to provide a value chosen so that life lived as a dependent (e.g., infants, children, and the elderly) is given less value than that lived during the productive years. The intrinsic value of life in this system increases from 0 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% of that of a 25-year-old. Paradoxically, the age weighting used in the DALY integral leads to higher valuation of life lived at under age 15 than does the HeaLY formulation, in which life lived at all ages is valued equally. Current formulations of the DALY leave age weighting as an option, and it is not used with the HALE.

Age-related valuing has been justified by showing that individuals value their own life lived at different ages differently. Such values have been reported in the literature for decades, and recent studies report that they are consistent across respondents of different ages (Busschbach, Hesing, & de Charro, 1993). Murray reports studies from many countries that reveal a preference for saving younger lives as compared with older ones (Murray& Lopez, 1994a). But it was not clear how much of the differential valuing of life at different ages was 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 it would be better to explicitly add a productivity factor (or subtract for the societal costs of dependents, such as education) rather than subsuming it under age weighting. See the section "Valuing Life for Its Economic and Social Productivity."

Valuing Future Life Compared with Present Life: Discounting

Discounting is the process of 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. It can be considered the inverse of an interest rate.

This concept 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 benefits take place in the future and are worth less than if they could occur now. This is equivalent to investing money now in order to obtain more in the future. A healthy life year in the present 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), which is a numeric reflection of societal values regarding intertemporal allocation of current resources. There is no consensus on the choice of a discount rate in health, but most agree that it should be lower than that in the private sector. The WDR in 1993 and the Global Burden of Disease studies discounted at 3% a year; in lieu of other information, this rate has come to be used in most international health cost-effectiveness studies. However, the impact of using a range of different discount rates, including zero, should be explored with each study.

The main issue concerning discounting in relation to composite 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)?

Valuing Life for Its Economic and Social Productivity

Whether and how to value economic and social productivity for purposes of health care 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 depend 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. Those under age 15 and over 65 may be considered as dependents and given a lower value. There are many variations for differential valuing, 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 health care provider 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 in terms of productivity have been put forward for valuing life according to productivity. Claims have been made that human life cannot be expressed in economic terms for decision-making purposes; however, efforts to avoid such expression nevertheless result in implicit valuation of life. Barnum has argued for adding productivity to the valuing of human life, stating that it has been ignored in health policy, is easily quantifiable, and does not ignore the welfare of children because the whole population is dependent on adult productivity for quality and sustenance (Barnum, 1987). 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. Of interest in this regard is that in the report of the Commission on Macroeconomics and Health (WHO, 2001), a DALY gained was stated to be worth at least an average annual income per head. Though 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

Decisions based on cost-effectiveness (e.g., cost per healthy life year) 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 needs may be.

In terms of social justice, equity has to do with a fair distribution of benefits from social and economic development. *Equity* is used in different conceptual senses: equal access to health services for all (opportunity equality), equal resources expended for each in-

dividual (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, which will certainly not lead to an equitable distribution in any other sense), care according to biological or socioeconomic need; and, finally, equal health states for all.

Evaluation of the disease burden in low- and middle-income nations reveals the persistence of infectious, childhood, and maternal conditions. These and other conditions of childhood predominate in low- and middle-income countries, and their impact on the poor is severe. Cost-effective interventions, such as immunization, exist for these conditions, and yet effective delivery has not been achieved. UNICEF reports that half the world's poor are children. They are paying an excessively high price for the failures of adults, while diseases and wars continue to threaten the lives of millions of children. It is estimated that more babies are being born into poverty than ever before. Poverty means that a child born in Malawi or Uganda will likely live only half as long as one born in Sweden or Singapore. It also means that one in three babies born in Niger or Sierra Leone will not live to see his or her fifth birthday.

Equity must go beyond equality of access to health care and must entail 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.

Composite summary measures such as HeaLYs and DALYS should be used not only to guide allocation of resources based on cost-effectiveness criteria, but also 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 countries 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 as both denominators and 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 1-year postcensus follow-up on a sample of enumeration areas in order 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, providing birth and death numbers. Underreporting, age misreporting, and other bias in data may have to be addressed (using standard demographic methods) prior to use in burden of disease estimation.

Mortality. Mortality data are required for any burden of disease analysis; 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. Information has to be particularly evaluated for deficiencies in the under-five years and older ages. 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 health care 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 infants and the elderly. Especially in low-income countries, it can be helpful to cross-check with other information, using postmortem interviews and hospital registers to assist in defining causes of death or 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. Often morbidity information is institution-based or restricted to one or two sources, such as hospital inpatient and clinic outpatient records. 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. These are 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, such as diabetes, hypertension, injuries, trachoma, and cataracts, can lead to blindness, 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. Although scales used in other studies may be helpful in making estimates, generally each group will have to construct its own de novo. The process used to construct a severity scale, the type of people participating, and the nature of the condition may all affect the final scale.

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 1-2. 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 (more commonly available than incidence) and knowledge of the average duration of the disease can be helpful.

The case fatality ratio (CFR) is the proportion of those developing the disease who die from it at any time. It is expressed as a decimal between 0 (for nonfatal conditions) and 1 (for universally lethal conditions such as AIDS). The case disability ratio (analogous with 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 that expected at that age had the disease not occurred. Similarly, expectation of life at fatality is that 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 qualitative and quantitative criteria determined a priori. For example, large population-based studies may be given preference over smaller sample-based work if both were available and the quality of data comparable. Better conclusions may be possible by cross-checking with 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 be from a biased population sample. The following are simple 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. These 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. The outcome is often most sensitive to one or more variables, making their precision 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, infectious, maternal, perinatal, and nutritional diseases. This group declines at rates faster than overall mortality rates as socioeconomic conditions improve; it contributes a relatively small share of deaths in the high-income world. Group II, noncommunicable and chronic diseases, accounts for most loss of healthy life in the high-income countries and proportionately increases with the epidemiologic transition (see Exhibit 1–3). Group III consists of injuries, both intentional and nonintentional.

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. The group I to group II ratios vary from 1:1 in the low- and middle-income world to 1:2 in Latin America, 1:5 in China, and 1:17 in the high-income world (Murray & Lopez, 1994a). Group III comprises 5% to 15% of the total burden, indicating that injuries are an important cause of death and disability everywhere. 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 as one moves 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

Generic steps for a national burden of disease study include the following:

- Assessing demographic information, including a census with age, sex, geographic (urban/rural), and selected socioeconomic status information distributions and vital statistics with births and deaths
- Collecting cause of death information for all deaths in a year by age, sex, geographic location, and socioeconomic status as possible according to the WHO International Classification of Disease system
- Defining disability by cause/disease and developing of a severity scale using expert and community input
- Collating information by disease from all sources and assessing reliability/validity, using expert opinion when needed for defining variables for a spreadsheet
- Defining social preferences such as age weighting, discounting, economic and social productivity, and expectation of life and deciding on their usage
- Estimating healthy life lost for each disease condition and by disease groups
- Performing a sensitivity analysis to check robustness of results to critical variables and assumptions
- Considering other variations, including assessment of losses by risk factors; regional, age, and sex breakdowns; and future projections
- Reviewing policy implications on overall mortality and morbidity in the country and for each cause; feeding into cost-effectiveness analysis and further research
- Including other modifications as appropriate to the country setting

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

- Estimating 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.
- Working out the costs of the interventions.
- Developing cost-effectiveness ratios to plan what combination of interventions targeted to which groups will provide a maximum return of healthy life per expenditure for the funds allocated to health.
- Reviewing the 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 adjusting as necessary. (Note that currently no country is assessing its health planning decisions on the basis of equity as proposed in this step.)

These steps 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 1-5). An actual study requires careful planning on the part of those responsible for its conduct and may require many further steps that are beyond the scope of this chapter.

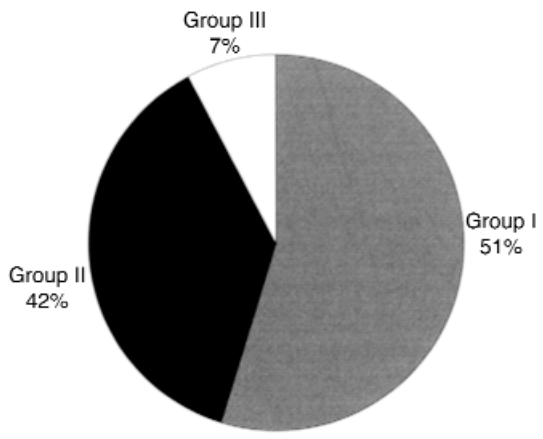
These steps summarize the essentials of applying the burden of disease methods to a country. A very important consideration in this process is time. The national studies in Mexico and the state of Andhra Pradesh in India have taken upwards of three years,

Exhibit 1-5 **The Burden of Disease in Pakistan, 1990**

Pakistan is a developing country in South Asia with a population of 112 million in 1990. A study was undertaken to estimate the burden of disease in Pakistan for 1990 and to calculate the loss of healthy life from a spectrum of common conditions(Hyder & Morrow, 2000). 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, and diarrhea and pneumonia in children caused the greatest loss of healthy life. Sixty-three percent of healthy life was lost from mortality, and 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 a great burden on infants and children.

Though communicable diseases dominate the burden of disease in Pakistan, noncommunicable diseases also take a heavy toll, as evident from a review of the top conditions responsible for loss of healthy life (see accompanying figure and table), and the proportion of loss from noncommunicable conditions can be expected to increase. Injuries need to be recognized as a major public health problem in the country. According to these estimates, Pakistan has a lower overall burden of disease than most countries in sub-Saharan Africa but a higher burden than most in Latin America.



Distribution of Disease Burden in Pakistan, 1990

Loss of Healthy Life in Pakistan: Top 10 Conditions for 1990

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

with two to three fulltime people. The conduct and analysis of such studies must be timely for 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. Hope lies in better definition of the exact data needed for the decisions and in automated collection systems combined with computerization to provide the right data at the right time.

Comparisons and Trends in Disease Burden

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

Comparative Disease Burden Assessments

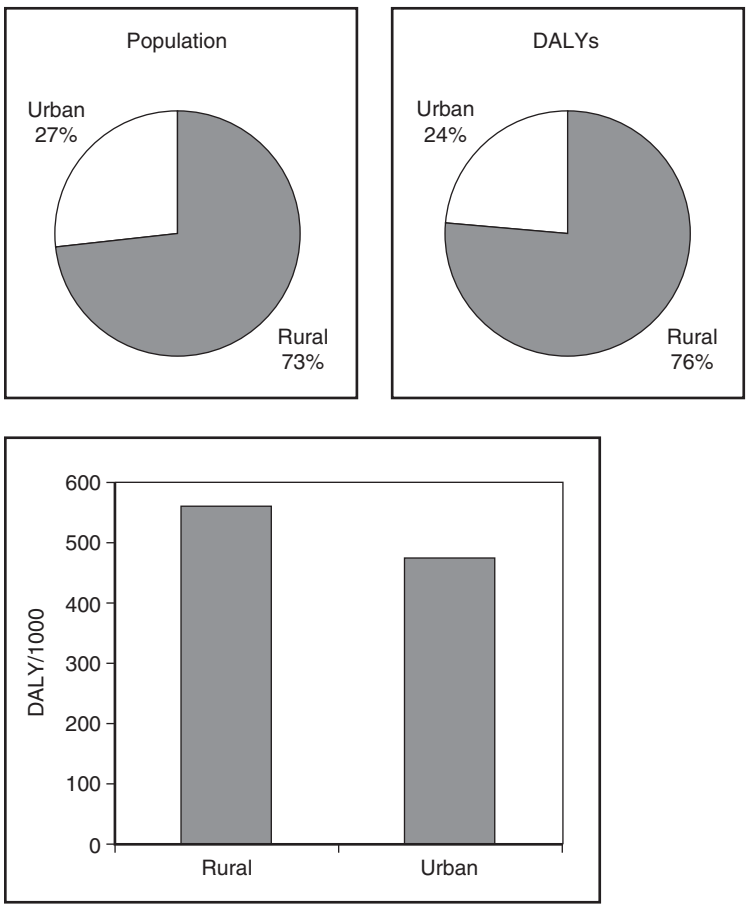
Comparing the burden of disease across populations, time, and place is an important aspect of national burden of disease studies. This subsection uses examples from recent burden of disease studies to illustrate how disaggregated data can help in understanding the distribution of ill health in a country.

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 a meticulous burden of disease study conducted between 1994 and 2001. It had a population of 76 million in 2001, with 27% urban (20.8 million people), and showed a 1:3 ratio of urban to rural disease burden in terms of DALYs lost (Mahapatra, 2001). The burden of disease rates was 19% higher in rural than in urban areas, as measured by DALYs lost per 1,000 population (Figure 1-4).

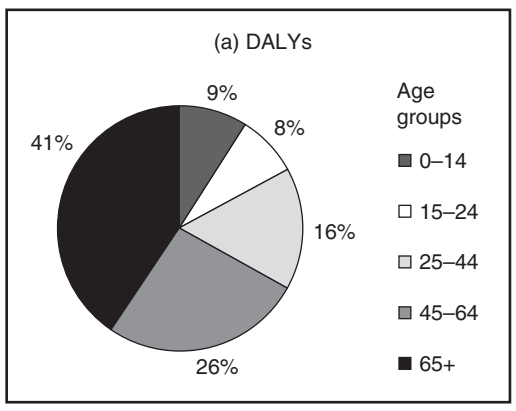
The Burden of Disease and Injury in New Zealand, 1996

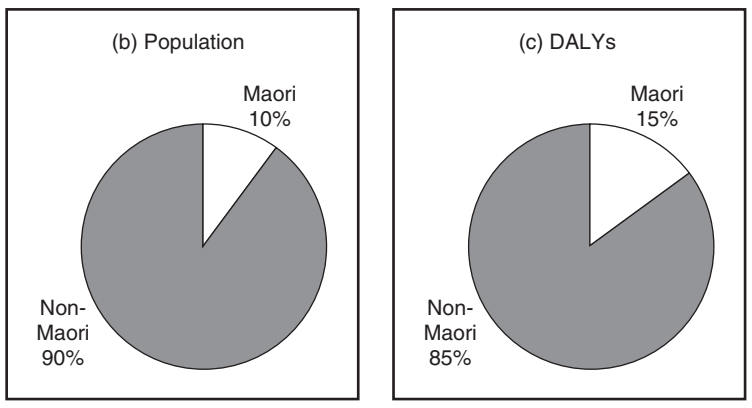
Age and ethnicity are key characteristics of a population that require a disaggregation of the burden of



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

Figure 1-4 Burden of Disease in Andhra Pradesh, 2001, by Region. *Source*: Mahapatra, P. (2001).





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

Figure 1-5 Burden of Disease in New Zealand, 1996, by Age (a) and Ethnicity (b and c). *Source*: New Zealand Ministry of Health. (2001).

disease. The national burden of disease study of New Zealand (population in 1996, 3.6 million) provides a clear example of how the DALYs lost in 1996 were predominantly among the older age group (65+), though they represented about 12% of the population (New Zealand Ministry of Health, 2001). The identification of 15% of the burden in the indigenous Maori population, compared with the 9.7% of the population they constitute, is an important finding (Figure 1-5).

The Burden of Disease in Chile, 1993

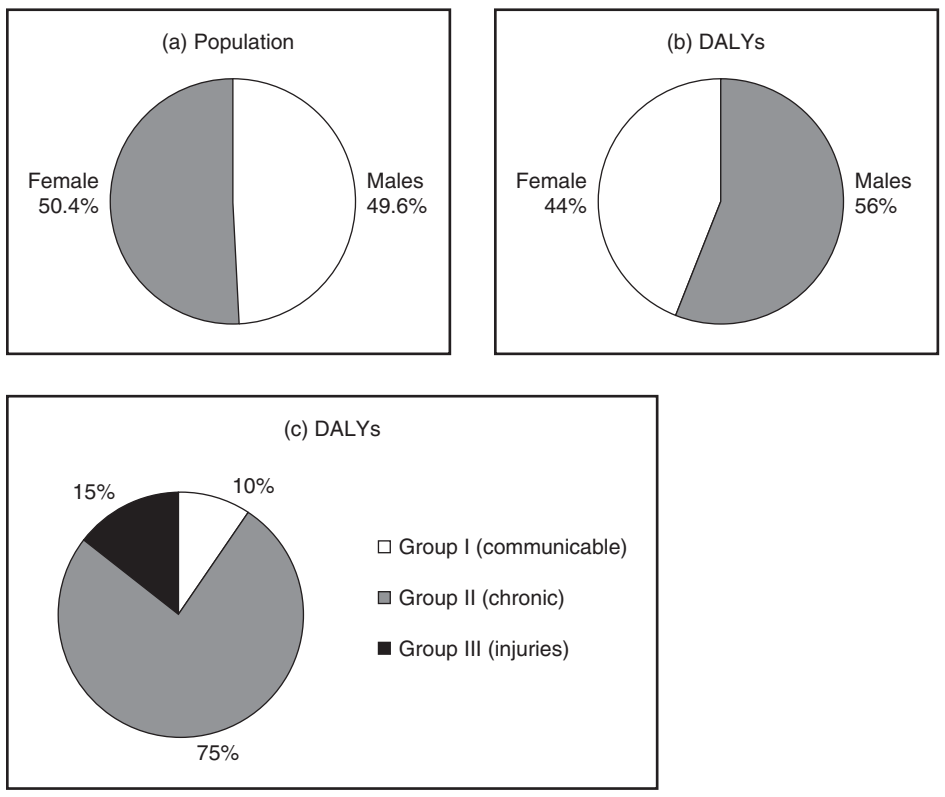
A disaggregated burden analysis by gender can also be seen in the work done in Chile in 1993, where at that time 49.6% of the population was male. The study found that 56% of the DALYs lost were among males (Figure 1-6). The distribution of the burden by major disease groups (see Figure 1-6) showed the dominance of chronic conditions in the burden (Concha, 1993).

Burden of Disease Estimates for South Africa, 2000

HIV/AIDS is ravaging Africa, and 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 could be attributed to HIV/AIDS (Figure 1-7) (Burden of Disease Research Unit, 2003); for a population of 45 million, this means 0.33 DALYs per capita. Such data are important for national decision making.

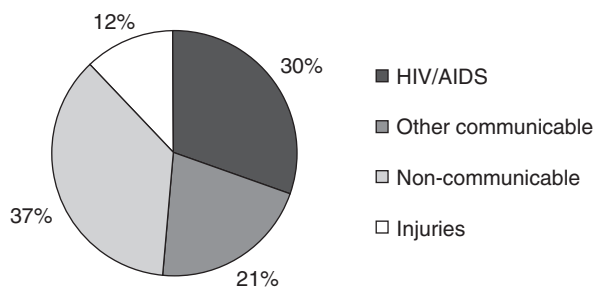
The Burden of Disease and Injury in Australia, 1996

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 1-8) (Mathers, Vos, & Stevenson, 1999). These show the high disability losses for



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

Figure 1-6 Burden of Disease in Chile, 1993, by Gender (a and b) and Disease Groups (c). *Source*: Concha. (1993).



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

Figure 1-7 Burden of Disease in South Africa, 2000, by Disease Groups. *Source*: Burden of Disease Research Unit. (2003).

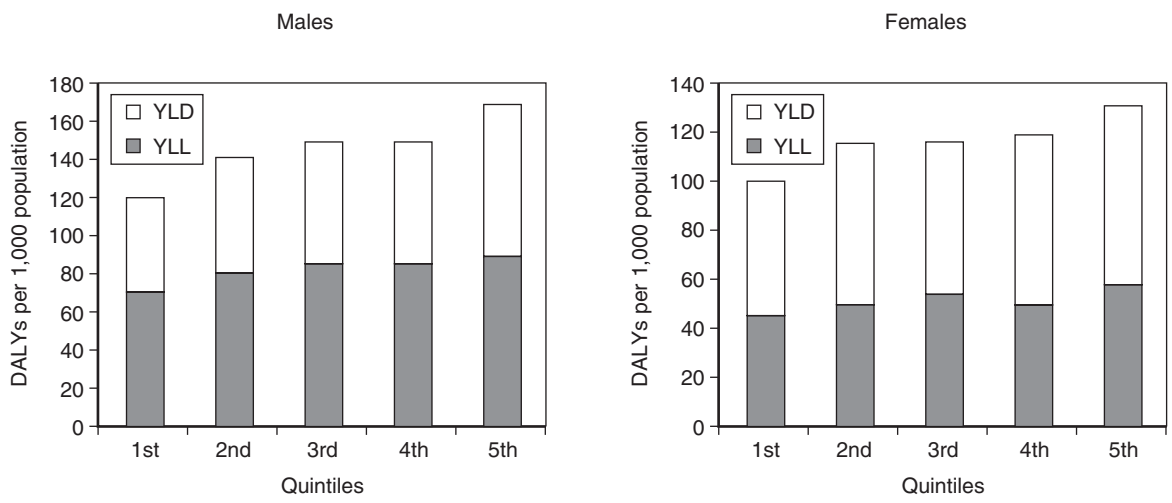
women and the poor. Such explorations of intranational distributions of disease burden are useful in studying the disproportionate impact of ill health on the poor and women.

Burden of Disease and National Income

WHO has categorized member states by income levels into high-, middle-, and low-income nations. The population of the world in 2000 was slightly more than 5 billion people, with 85% in the low- and middle-income nations (Figure 1-9). As may be expected, more than 90% of the global burden is found in the 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

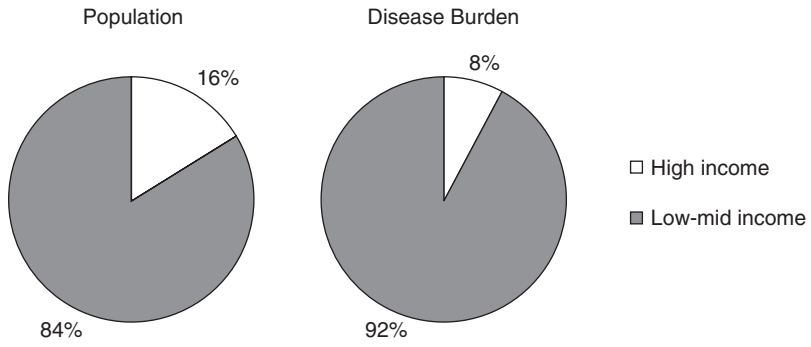
Another way to disaggregate data is to explore the disease burden based on disease groups: group I (communicable, infectious, maternal, and perinatal), group II (noncommunicable, chronic), and group III (injuries and violence). There is great variation in the portions allocated to these groups; for example, group I may be responsible for 12% to 70% of the burden. When the countries are stratified by GNP per capita as a measure of development, an important trend can be seen (Table 1-4). As income rises, the proportion of the burden attributable to group I decreases, while



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.

Note: Total DALYs lost in Australia for 1996 = 2.5 million.

Figure 1-8 Burden of disease in Australia, 1996, by socioeconomic status and gender. *Source*: Mathers, Vos, Stevenson. (1999).



Note: Total global disease burden for 2000 = 1.46 billion DALYs.

Figure 1-9 Global Burden of Disease, 2000, by Income Level of Countries. *Source*: World Health Organization. (2000).

that of group II increases. The effect is progressive, although countries such as Turkmenistan (middle income) still retain a high group I burden. This is consistent with the theory of epidemiologic transition predicting a change in disease profile with economic development.

Intentional and unintentional injuries are responsible for 4% to 40% of the disease burden. Injuries contribute significantly to premature death and disability in low- and middle-income countries. The primary causes within this category also tend to change with development, although causes such as road traffic crashes are ubiquitous.

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 done as a separate 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.

Table 1-4 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 of $635 or less)[a]			
Andhra Pradesh	54	30	16
Guinea	70	23	7
Lower medium-income nations (GNP per capita > $635 ≤ $2,555)			
Colombia	22	39	39
Jamaica	16	60	24
Turkmenistan	51	45	4
Uzbekistan	46	40	14
Upper medium-income nations (> $2,555 ≤ $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.

[a]GNP per capita from the World Bank (1993).

Source: World Bank, *World Development Report 1993: Investing in Health* (New York: Oxford University Press, 1993). Reprinted with permission from Oxford University Press.

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 2000 study presented estimates for 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, in addition to assessment of disability for evaluation of the disease burden using DALYs. The results were based on a variety of sources, including vital registrations systems, special studies, surveys, and expert opinion.

Mortality. Globally, in 2000, ischemic heart disease, cerebrovascular disease, and respiratory infections were the top three causes of death, while 10 causes accounted for 54% of deaths worldwide. One death in 10 was from injuries, with road traffic accidents included in the top 10 causes of deaths. The low- and middle-income world accounted for 98% of all deaths in children, 83% of deaths in persons aged 15 to 59 years, and 59% of deaths in persons aged 70+ years. Of all deaths in the low- and middle-income world, 28% were in children. Thus, an inordinate proportion of the mortality burden at the beginning of this decade is in low- and middle-income countries, even at adult and older ages.

Table 1-5 shows the differences in the 10 leading causes of deaths for 2000 for the high-income and the low- and middle-income world. The presence of perinatal conditions, tuberculosis, HIV/AIDS, and malaria in the low- and middle-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 world, reflecting the success in combating these infectious conditions. It is important to note that noncommunicable diseases such as ischemic heart disease were already prominent causes of premature death in the low- and middle-income world in 2000.

Disability. The Global Burden of Disease 2000 study also evaluated the contribution of conditions to disability in the world. Leading causes of disability in 2000 worldwide are shown in Table 1-6. Neuropsychiatric and behavioral conditions dominate the causes of disability, represented by 4 of the top 10 conditions. However, a diverse spectrum of conditions, such as hearing loss, congenital anomalies, and osteoarthritis, also appears on the list. This has been a unique contribution of the Global Burden of Disease work—placing nonfatal health outcomes in the center of international health policy in recent years. The important, and yet often ignored, impact of these conditions is obvious once disability is counted in estimates of disease burden.

Disease Burden. Based on the estimation of deaths and disability presented previously, the global disease burden for 2000 was estimated using disability-adjusted life years (DALYs). Leading causes of the global burden of 2000 (Table 1-7) indicate the impact of those conditions affecting the low- and 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 continues fighting the infectious diseases, improving the response to chronic conditions, and preparing to meet the increasing impact of injuries, all at the same time.

Age and Disease Distributions. Figure 1-10 presents the distribution of the global burden in 2000 by disease groups and demonstrates the growing relative impact of chronic diseases (group II) over infectious diseases (group I). Comparable figures for loss of healthy life in sub-Saharan Africa, the Middle Eastern crescent, Latin America, and the Carribean are presented in Figure 1-11. It is important to note

Table 1-5 Leading Causes of Deaths in High-Income and Low- and Middle-Income Countries, 2000

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

Source: Murray, Lopez, Mathers et al. (2001).

Table 1-6 Leading Causes of Disability Losses Globally, 2000

Rank	Cause
1	Unipolar major depression
2	Hearing loss, adult onset
3	Alcohol use disorders
4	Osteoarthritis
5	Schizophrenia
6	Perinatal conditions
7	Bipolar disorders
8	Chronic obstructive pulmonary disease
9	Congenital anomalies
10	Asthma

Note: Disability losses are defined by years of life lived with disability—YLDs.
Source: Murray, Lopez, Mathers et al. (2001).

Table 1-7 Leading Causes of Global Burden of Disease, 2000

Rank	Cause
1	Lower respiratory conditions
2	Perinatal conditions
3	HIV/AIDS
4	Unipolar major depression
5	Diarrheal diseases
6	Ischemic heart disease
7	Cerebrovascular diseases
8	Road traffic accidents
9	Malaria
10	Tuberculosis

Source: Murray, Lopez, Mathers et al. (2001).

that communicable diseases still represent a considerable portion of the disease burden in 2000, especially in sub-Saharan Africa due to HIV/AIDS.

The analyses indicate that subregions within middle- and low-income countries are at different stages of the epidemiologic transition. The influx of chronic diseases has added another layer of problems, while the burden of communicable diseases has not yet been eradicated. This double burden is 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.

Other Ways Burden Can Be Measured

Mortality and morbidity alone have been used for decades for international comparisons of disease burden. Child mortality under the age of 5 years is considered a sensitive indicator of the overall health of nations, especially women and children. UNICEF publishes an annual *State of the World's Children* report

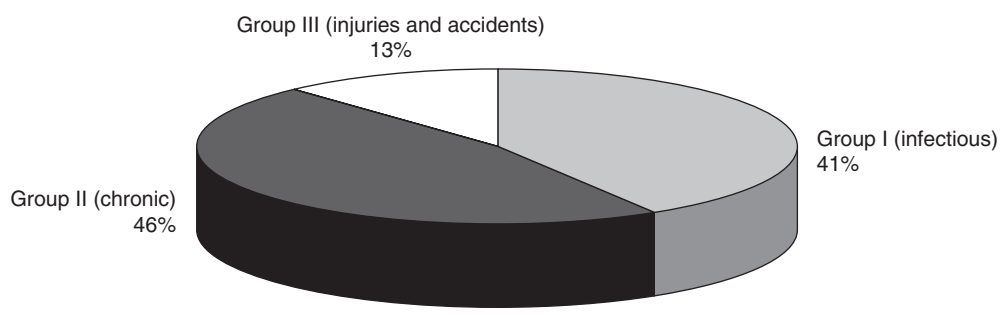


Figure 1-10 Global Burden of Disease 2000 by Disease Groups. *Source:* Murray, Lopez, Mathers et al. (2001).

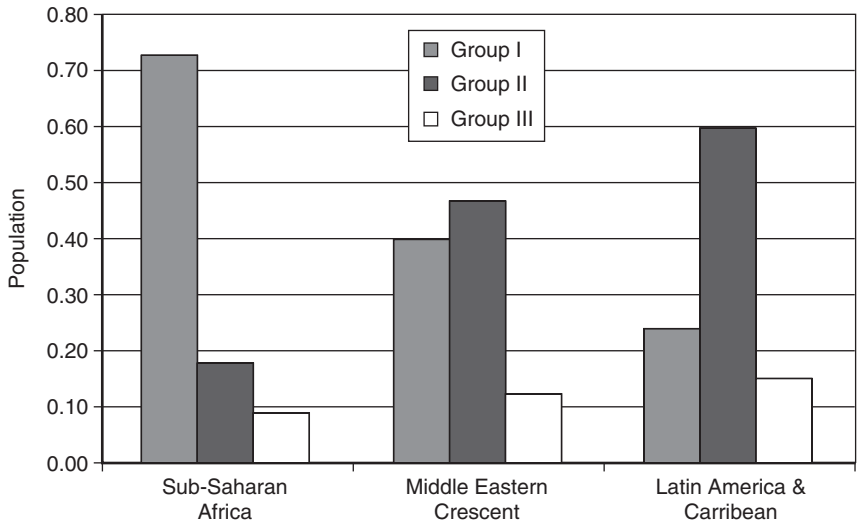


Figure 1-11 Proportion of Disease Burden by Disease Groups in Selected Regions, 2000. *Source:* Murray, Lopez, Mathers et al. (2001).

that includes a ranking of nations based on this indicator (Table 1-8). Gross national income (GNI) per capita is an indicator of national wealth, and the relationship between these variables usually follows an expected sequence in which the country with the lowest GNI per capita has the worst indicators of health. However, as Table 1-8 indicates, 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, Bhutan has a per capita GNI that is higher than Mongolia, but it ranks lower in child mortality and life expectancy. These examples demonstrate that the relationship between health and poverty is complex and needs in-depth investigation. To improve the health of nations, the disparities within societies that are impediments to the empowerment of the poor and needy, especially women and children, need to be addressed as well as absolute poverty.

In 1999 UNICEF reported a new risk index for children in countries worldwide. This proposed index was developed with the intent of measuring children's welfare in a new manner. This national index measures countries on a scale of 0 to 100 and is based on the following factors: mortality rate of children under 5, percentage of children who are moderately or severely underweight, access to primary schooling, risks from armed conflict, and risks from HIV/AIDS. The high-income nations—United States, Australia, New Zealand, and Japan— are in the lowest risk index, whereas the poorest nations of Angola, Sierra Leone, and Afghanistan are in the highest risk category. As a continent, Africa is in the highest risk category. This index does not consider other factors that impinge on child welfare, such as child labor, sexual exploitation, and lack of family support. However, the collation of traditional indicators of child and national health (such as child mortality) with issues of access (to primary education) and emerging threats

Table 1-8 **Health Status Indicators and National Income for Selected Developing Nations (UNICEF, 2004)**

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: United Nations Children's Fund, *The State of the World's Children 2004* (New York: UNICEF, 2004). Reprinted with permission.

(HIV) makes for an innovative approach to measuring the suffering linked to poverty and bringing it to the attention of the global community.

Future Projections

Future projections of disease burden have been attempted, with the intent of providing some basis for health planning. This is a challenging task that requires further data manipulations and the use of assumptions. These assumptions must predict changes in disease prevalence and incidence over time, the effect 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 Global Burden of Disease study for 1990 attempted to project the global burden in the future to the year 2020. These estimates were based on projected changes in the expectation of life, age structure of the global community, disease profiles based on current states, and other relevant parameters (Murray & Lopez, 1996a). In addition, the projections were 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 2020, as shown in Table 1-9. The domination of chronic diseases is obvious, although respiratory conditions still appear to be important. Injuries from road traffic crashes are projected to become the third leading cause of the global disease burden. It is interesting to note that the mortality and disability consequences of war make it the eighth leading cause of projected global disease burden. In addition, the lower ranking of HIV in 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 project a different scenario for the future.

The growing importance of noncommunicable diseases may be a global phenomenon, and their impact on low- and middle-income countries and regions needs to be assessed. Table 1-9 also shows the projected leading causes of the disease burden in the low- and middle-income world for 2020. Here again, four of the top five conditions are chronic diseases and injuries. However, unlike the list for the world, the persistent burden of respiratory infections and diarrheal diseases is evident. The situation in the low- and middle-income world is one in which the triple burden of persistent communicable diseases, prevalent noncommunicable conditions, and increasing injuries will call for an appropriate response.

Burden of Risk Factors

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

Table 1-9 **Projected Leading Causes of Disease Burden in 2020**

Global		Developing Regions Only	
Rank	**Cause**	**Rank**	**Cause**
1	Ischemic heart disease	1	Unipolar major depression
2	Unipolar major depression	2	Road traffic accidents
3	Road traffic accidents	3	Ischemic heart disease
4	Cerebrovascular disease	4	Chronic pulmonary obstructive diseases
5	Chronic pulmonary obstructive diseases	5	Cerebrovascular disease
6	Lower respiratory infections	6	Tuberculosis
7	Tuberculosis	7	Lower respiratory infections
8	War	8	War
9	Diarrheal diseases	9	Diarrheal diseases
10	HIV	10	HIV

Source: C. J. L. Murray and A. D. Lopez (Eds.), *The Global Burden of Disease and Injury 1990* (Geneva, Switzerland: World Health Organization, 1996). Reprinted with permission from the Harvard Center for Population and Development Studies.

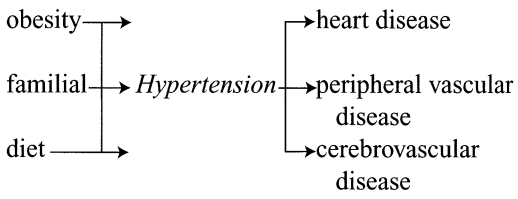


Figure 1-12 Linkages of risk factors with health outcomes.

specific diseases attributable to any one factor may be difficult to estimate. Relationships such as those shown in Figure 1-12 need careful assessment to determine the burden from heart disease that can to be attributed to hypertension in relation to other interacting causal factors. The best way to determine the portion of disease that may be attributed to hypertension is through a randomized trial and careful assessment of disease outcomes over time: Results from such studies have shown a reduction of death and disability from not only cardiac disease but also cerebrovascular and renal diseases.

Because the purpose of risk factor analysis is to assist in decisions concerning 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 healthy life lost attributed to the diseases that the risk factor affects. For the evaluation of an intervention that reduces hypertension, the HeaLY losses from the entire range of diseases that hypertension influences are required.

The Burden of Selected Major Risk Factors

As reviewed in this chapter, a substantial body of work has focused on the quantification of trends in mortality and, more recently, burden of disease.[1] However, reliable and comparable analyses of risks to health, key for preventing disease and injury, have not been quantified as well. Most analyses of the relation 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. This has made comparisons of risk factors on a population health level difficult. As a result, the Comparative Risk Assessment (CRA) project of the Global Burden of Disease 2000 study attempted a systematic evaluation of the contributions of selected risk factors to global and regional burden using a specific model for analysis (Murray et al., 2001).

The model used in CRA for causal attribution of health outcomes was based on counterfactual analysis (Ezzati et al., 2002). Under this analysis, the contribution of one or a group of risk factors to disease or mortality is estimated by comparing the current or

[1]This section is based on a paper by Ezzati et al. (2002).

future disease burden with the levels that would be expected under an alternative hypothetical scenario (referred to as the counterfactual). In this case, the CRA project's estimates of burden of disease and injuries due to risks were based on a counterfactual exposure distribution that would result in the lowest population risk. This involves an evaluation of the effect a risk factor has on the disease or mortality 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).

Twenty-five risk factors were selected by the CRA project based on how likely they were to be among the leading causes of death and disease, the likelihood of causality, the availability of data, and whether or not they were modifiable. For each risk factor, an expert working group did a comprehensive review of published work and other sources to obtain the data on prevalence of risk factor exposure and hazard size.

The results of this analysis show the contribution of the 20 leading global risk factors for the world and for three broad combinations of regions: demographically and economically developed, lower-mortality developing, and high-mortality developing (Figure 1-13). Undernutrition was the single leading global cause of health loss, with 140 million DALYs lost; 9.5% of this was from underweight, 2.4% from iron deficiency, 1.8% from vitamin A deficiency, and 1.9% from zinc deficiency. Although the prevalence of underweight has decreased as a global average, it has increased in sub-Saharan Africa, where its effects are disproportionately large due to simultaneous exposure to other childhood disease risk factors. At the same time, risk factors for noncommunicable diseases—high blood pressure, high cholesterol, and high body mass index (BMI)—are also widely prevalent and causing substantial loses to healthy life. The lower-mortality developing regions (40% of the global population) suffer from risk factors affecting both developed and high-mortality developing regions.

An important finding of this analysis is the key role of nutrition in health worldwide. Approximately 15% of the global disease burden can be attributed to the joint effects of childhood and maternal underweight or micronutrient deficiencies. In addition, almost as much can be attributed to risk factors that have substantial dietary determinants—high blood pressure, high cholesterol, high BMI, and low fruit and vegetable intake.

Some risk factors may have little impact on the total global burden of disease, but they may be very important locally within certain populations and regions. For example, iodine deficiency still affects important parts of the developing world, resulting in substantial disability in those populations. The approach used in CRA was new and innovative, but has not yet been widely used at the national level. With further experience and additional refinements, studies concerned with interventions directed toward risk factors will play an increasing role in improving the health of populations.

Conclusion

The health of populations is the fundamental concern of international 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. Composite 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 are coming to be important tools for assisting health related decision making, but to avoid misuse, it is critical for those using them to understand the underlying assumptions and limitations and also to meet the rather formidable data requirements. These summary measures also could be used to examine the burden of disease among subpopulations according to socioeconomic, cultural, and especially vulnerable groups and to ensure that health-related decisions consider equity as well as cost-effective criteria.

Trends in disease burden provide important clues to the success of ongoing health programs and the need for development of new interventions. At the

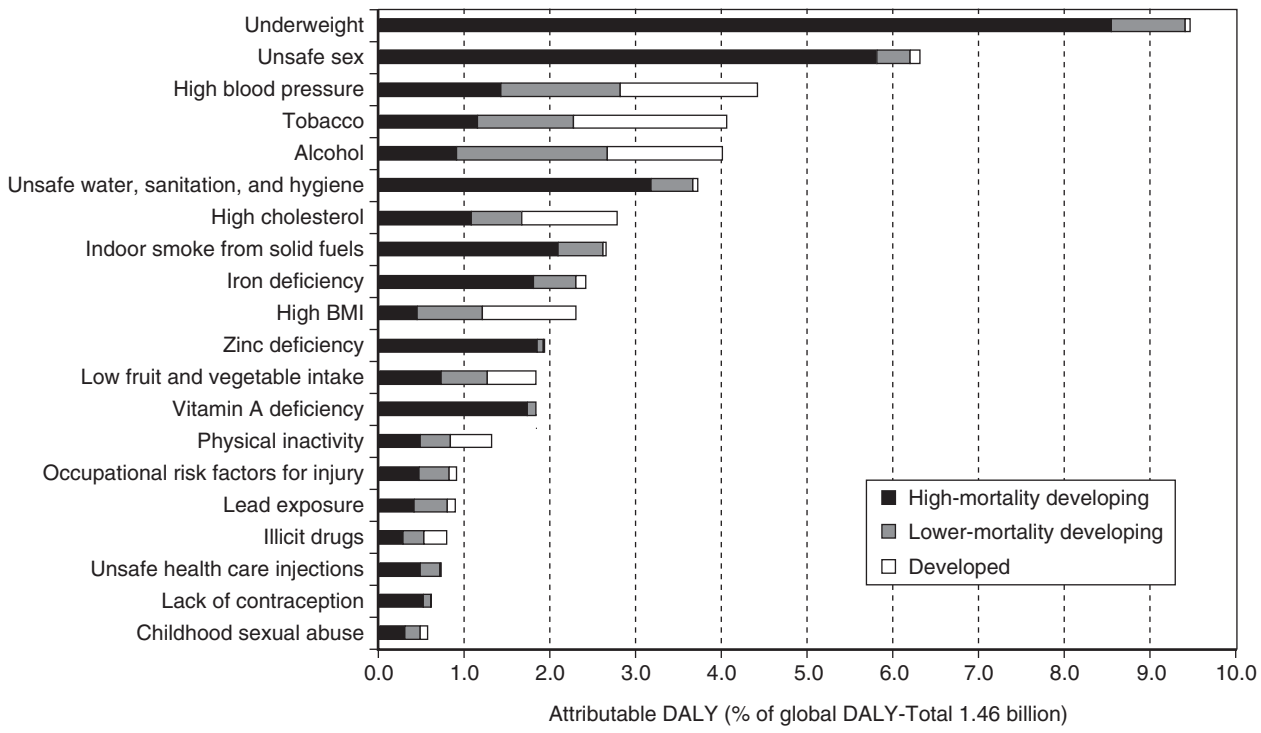


Figure 1-13 Mortality (a) and Burden of Disease (b) Due to Leading Global Risk Factors. *Source:* Ezzati, Lopez, Vander Hoorn. (2002).

same time, they reflect non-health factors that are important to the production or maintenance of health in populations. Intercountry and interregional comparisons allow for measuring progress among nations and can highlight inequalities in health status and examine these in relation to social, economic, educational, and other factors as well.

Health systems across the world are greatly affected by 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 prerequisites for improving equitable global health development.

• • • DISCUSSION QUESTIONS

1. What is the primary purpose of a health system in a country? How can data help achieve this purpose?
2. What are the essential elements of health information, and what types of data are required to assess ill health?
3. What are the relative strengths and weaknesses of composite indicators compared with more traditional indicators of disease burden?
4. In your country, what would be the most appropriate set of indicators to assess the impact of diseases on the population? Why?

• • • References

1. Anand, S., & Ranaan-Eliya, R. (1996). *Disability adjusted life years: A critical review* (Working Paper No. 95-06). Boston: Harvard Center for Population and Development Studies.
2. 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.
3. Barker, C., & Green, A. (1996). Opening the debate on DALYs. *Health Policy and Planning, 11*(2), 179–183.
4. Barnum, H. (1987). Evaluating healthy days of life gained from health projects. *Social Science and Medicine, 24*, 833–841.
5. 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.
6. Bobadilla, J. L. (1998). *Searching for essential health services in low and middle income countries*. Washington, DC: Inter American Development Bank.
7. Burden of Disease Research Unit. (2003). *Initial burden of disease estimates for South Africa, 2000*. South Africa: South African Medical Research Council. Retrieved August 17, 2004, from http://www.mrc.ac.za/bod/bod.htm.
8. 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.
9. Coale, A. J., & Demeney, P. (1983). *Regional model life tables and stable populations*. New York: Academic Press.
10. Coale, A. J., & Guo, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index, 55,* 613–643.
11. Concha, M. (1993). *Burden of disease in Chile.* Retrieved August 17, 2004, from http://www.globalforumhealth.org/Non_compliant_pages/forum3/forum3doc823.htm.
12. Crooper, M., & Kopits, E. (2003). *Traffic fatalities and economic growth* (World Bank Policy Research Working Paper 3035). Washington, DC: World Bank.
13. Eddy, D. (1991). Oregon's methods: Did cost-effectiveness analysis fail? *Journal of the American Medical Association, 266*, 2135–2141.
14. 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.
15. Ezzati, M., Lopez, A., Vander Hoorn, S., Rodgers, A., Murray, C. J. L., & the Comparative Risk Assessment Collaborative Group. (2002). Selected major risk factors and global and regional burden of disease. *Lancet, 360*(9343), 1347–1360.
16. Feachem, R. G. A., & Kjellstrom, T. (Eds.). (1992). *The health of adults in the developing world*. New York: Oxford University Press.
17. 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.
18. 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.

19. Gudex, C., & Kind, P. (1988). *The QALY toolkit*. York, England: Centre for Health Economics, University of York.

20. Hill, K. (2001). Demographic techniques: Indirect estimation. In *International encyclopedia of social and behavioral sciences*. Oxford, England: Elsevier Science.

21. 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.

22. 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.

23. 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.

24. Joint United Nations Programme on HIV/AIDS. (2003). *Global summary of the HIV and AIDS epidemic in 2003*. Retrieved October 13, 2004, from http://www.unaids.org/en/resources/epidemiology/epicorejuly2004.asp.

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

26. King, M. (1990). Health is a sustainable state. *The Lancet. 336*, 664-667.

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

28. 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.

29. Mahapatra, P. (2001). *Estimating national burden of disease: The burden of disease in Andhra Pradesh, 1990s*. Hyderabad, India: Institute of Health Systems. Retrieved August 17, 2004, from http://www.ihsnet.org.in/BurdenOfDisease/APBurdenofDiseaseStudy.htm.

30. 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.

31. Mathers, C., Vos, T., & Stevenson, C. (1999). *The burden of disease and injury in Australia*. Canberra, Australia: Australian Institute of Health and Welfare. Retrieved on August 17, 2004, from http://www.aihw.gov.au/.

32. Morrow, R. H. (1984). The application of a quantitative approach to the assessment of the relative importance of vector and soil transmitted diseases in Ghana. *Social Science and Medicine, 19*, 1039–1049.

33. 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.

34. Mosley, H. W., Jamison, D. T., Bobadilla, J. L., & Measham, A. (Eds.). (1993). *Disease control priorities in the developing world*. New York: Oxford University Press.

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

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

37. Murray, C. J. L., & Lopez, A. D. (1994b). Quantifying disability: Data, methods and results. *Bulletin of the World Health Organization, 72,* 481–494.

38. Murray, C. J. L., & Lopez, A. D. (Eds.). (1996a). *The global burden of disease 1990*. Geneva, Switzerland: World Health Organization.

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

40. 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.

41. 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: World Health Organization.

42. 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.

43. New Zealand Ministry of Health. (2001). *Burden of disease and injury in New Zealand*. Retrieved August 17, 2004, from http://www.moh.govt.nz/moh.nsf/.

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

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

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

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

48. Preston, S. H. (1976). *Mortality patterns in national populations*. New York: Academic Press.

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

50. Sullivan, D. F. (1971). A single index of mortality and morbidity. *HSMHA Health Reports, 86*, 347–354.

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

52. United Nations Children's Fund. (2004). *The state of the world's children 2004*. New York: UNICEF.

53. United Nations Population Division. (2002). *World population prospects: The 2002 revision population database*. Retrieved October 7, 2004, from http://esa.un.org/unpp.

54. 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.

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

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

57. World Health Organization. (1992). *International statistical classification of diseases and related health problems* (10th revision). Geneva, Switzerland.

58. World Health Organization. (2000). *The world health report 2000*. Geneva, Switzerland.

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

60. World Health Organization. (2002a). *The international classification of functioning, disability and health: Introduction*. Retrieved October 7, 2004, from http://www.who.int/classification/icf/intros/ICF-Eng-Intro.pdf.

61. World Health Organization. (2002b). *World report on violence and health*. Geneva, Switzerland.

62. World Health Organization. (2003). *The world health report 2003*. Geneva, Switzerland.

63. World Health Organization. (2004). *The world health report 2004*. Geneva, Switzerland.

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