Chapter **15**

Cost-Effectiveness Analysis for Priority Setting

Philip Musgrove and Julia Fox-Rushby

The economic analyses in this volume focus on activities whose main objective is to improve health. Although the chapters vary considerably, all possess, nonetheless, a common core of definitions, assumptions, and methods of analysis. These are drawn primarily from concepts and applications in the *Oxford Textbook of Public Health* (Jamison 2002), drawing partly on the 1993 *World Development Report* (World Bank 1993). In this chapter, we summarize and explain the common features and some of the variations of economic analysis and point the reader to examples throughout the book.

First is a general discussion of cost-effectiveness analysis (CEA), which is the principal analytic tool used here. Here we explain what such analysis does and does not provide, how it is related to the concept of burden of disease, and how it can be used, along with other criteria, in setting priorities.

Because CEA is applied to specific interventions, the next section describes the several meanings of that term and the way that interventions are classified and evaluated. It is essential to understand what is being analyzed before considering in detail how the analysis is conducted.

Estimating the effectiveness of an intervention requires specifying the units in which that concept is measured. This action in turn requires choices of several parameter values, including, in the analyses reported here, the discount rate applied to future years; the disability weights that describe the severity of diseases and conditions, corresponding to the health losses that they cause; and the life expectancy at different ages, which determines how many years of healthy life can be saved by averting a death or preventing or treating a long-term health problem. We also consider briefly the nonhealth benefits that may result from a health intervention.

The subsequent section deals with the costs of interventions: first with the question of which costs to include in the analysis, and then with the conversion of costs in national currencies to equivalents in U.S. dollars for international comparisons.

Despite the common assumptions and parameter values, the economic analyses differ from chapter to chapter in how comprehensive and how exact they are, including how they deal with a variety of approximations and how the results vary from place to place or according to differences in the assumptions. This section also contains a brief description of the differences in the quality of the basic evidence and in how widely conclusions are applicable.

Estimates of the cost-effectiveness of interventions often describe what happens at the level of the individual patient or beneficiary. In the next section, we suggest two ways to consider costs and outcomes at the population level, allowing for large differences among countries in the size of population; the incidence or prevalence of a disease, condition, or risk factor; and the amount spent or available to spend on an intervention.

In the final section, we indicate how the type of analysis presented in this volume might be improved and how it can be applied to help set priorities among the large number of interventions to which limited resources can be applied.

COST-EFFECTIVENESS AND PRIORITY SETTING

The principal analytic tool throughout this volume is CEA, which compares the cost of an activity, called an *intervention,* with the known or expected health gain. The result is summarized in a cost-effectiveness ratio (CER), as explained more fully below. This ratio corresponds to the concept of

(health) value for money. Favoring activities that are more cost-effective over those that are less so is consistent with the ethical view that "limited resources for health should be allocated to maximize the health benefits for the population served" (chapter 14).

Cost-effectiveness provides the clearest simple way to promote value for money in health: hence, the emphasis on it here. CEA allows comparisons throughout the health sector and not only for the same health outcome. It does not allow comparison to nonhealth outcomes unless these outcomes can be incorporated into costs, and the calculation of the CER by itself makes no pretense of monetizing the intrinsic value of health. To use CERs for choosing what to buy and what not to, decision makers must determine a maximum willingness to pay for units of health gain, unless other criteria are considered to justify buying something with relatively poor cost-effectiveness.

For risk factors, CEA requires estimating the gain in health consequent on introducing an intervention to reduce the risk of acquiring or transmitting a condition. For packages of interventions or elements of the health system, such as hospitals, effectiveness is estimated by judging how much mortality and morbidity would be reduced by providing the whole package or set of services or by operating the facility. With some exceptions, the analyses may describe but do not quantify the nonhealth benefits of an intervention. Apart from the difficulty of obtaining enough data, such quantification requires attaching values to nonhealth outcomes, which is problematic when comparisons are made over large cultural and income differences.

All comparisons are relative, with no absolute distinction between being and not being cost-effective. In assigning priorities among interventions for public funding or for other policy actions, one must also consider the magnitude of health problems to which interventions apply because that affects what is affordable. Calculations of the effect of spending US$1 million or the total cost and health gain in a population of 1 million people offer ways of looking at such choices. Equity, poverty, and risk of impoverishment from ill health may also influence priorities; so do the budgets available—and the decisions of how much to make available—for buying interventions. Finally, the effectiveness of an intervention and, therefore, the degree to which it deserves priority depend on how far it is culturally appropriate or acceptable for the population it is intended to benefit. The identical intervention, technically speaking, may lead to different degrees of use or compliance in different population groups, and information and incentives may be needed to achieve the full potential outcomes.

Cost-effectiveness is only one of at least nine criteria relevant for priority setting in health if the object is to decide how to spend public funds (Musgrove 1999). Cost matters by itself, as do the capacities of potential beneficiaries to pay for an intervention. The other criteria that may affect priorities include horizontal equity (equal treatment for people in equal circumstances); vertical equity (priority for people with worse problems); adequacy of demand; and public attitudes and wants. Two criteria—whether an intervention is a public good and whether it yields substantial externalities—are classic justifications for public intervention, because private markets could not supply them efficiently, just as in other sectors. As noted in chapter 1, the interventions analyzed in this volume are not limited to public or semipublic goods. The emphasis is on value for money—that is, whether an intervention is worth buying, not who pays for it. Nonetheless, when one is choosing which public goods to buy, several criteria become irrelevant, and cost-effectiveness can be used as the chief or even the only consideration. Cost-effectiveness can similarly determine what to include in a mandatory universal public package of health care alongside competitive voluntary insurance (Smith 2005).

Cost-effectiveness can conflict with both kinds of equity—that is, the more cost-effective of two interventions may also lead to a less equitable distribution of health benefits. Equity and cost-effectiveness are compatible when a cost-effective intervention is provided to only part of the population that would benefit from it because everyone in the group suffers from the same problem. Then expanding coverage will generally also promote horizontal equity. These equity effects are reinforced when those who are better off already benefit while the poorer and sicker population does not. Choices about vertical equity—doing more for those in greatest need—are more complicated. Doing very little for people with severe health problems—because the available interventions for those problems are not very effective at reducing their suffering—is not necessarily preferable to doing more for people with less severe problems that are more amenable to intervention. When an intervention is reaching only part of a potential beneficiary population and those not benefiting tend to have more severe illness, then expanding coverage can improve both horizontal and vertical equity. Where possible, chapters consider the equity effects of expanding or changing interventions.

Cost-Effectiveness and Disease Burden

Cost-effectiveness and disease burden are related because effectiveness is the reduction in burden caused by an intervention. This relationship holds true at the individual level. The magnitude of a health problem—the total burden in the population—is irrelevant for marginal changes in resource allocation. However, it matters for large changes from the status quo. Health interventions demand managerial capacity as well as financial and physical resources, and managerial ability may be stretched thin if it has to deal with a large number of interventions. In consequence, it may be efficient to concentrate on relatively few and somewhat less cost-effective interventions, provided they attack substantial burdens, rather than many other interventions that are more cost-effective but affect only small burdens. Moreover, even for a cost-effective

intervention, high prevalence or incidence may make the cost of covering the whole potential beneficiary population prohibitive. The authors of chapter 21 indicate how expensive it would be to protect all at-risk African children from malaria with bednets, even though bednets are highly cost-effective. Conversely, an intervention that costs more per health gain may be affordable and given priority if it treats a manageable burden of disease and corresponds to a small beneficiary group. Priority turns on the available budget relative to the cost of a program; on how divisible a program is (that is, how easily it can be operated at different scales, as a technical or political matter); and on whether interventions are mutually exclusive (Karlsson and Johannesson 1996).

Because of the interaction between cost-effectiveness, disease burden, and available funds, no single threshold of maximum cost per health gain exists below which an intervention is "cost-effective." A rule of thumb, such as that any intervention is worthwhile if it costs less than two or three times income per capita, ignores this interaction and is an inadequate guide to priority setting. However, even an intervention that is considered justified by cost-effectiveness may be infeasible to deliver, for example, if the costs are monetary and come from the public budget but the benefits are nonmonetary and diffused over the population. Economic theory would suggest removing the current budget constraint by raising more revenue until the marginal social cost of the interventions plus the cost of obtaining the revenue equals the marginal social benefit. Although theoretically attractive, this escape from resource limitation may not be possible because of political reasons or because the economic cost of raising extra taxes is prohibitive.

Because so many criteria can affect priority setting and because evidence on cost-effectiveness in low- and middle-income countries is so scarce, health system policies and budgets seldom derive purely from considerations of cost versus outcomes. Even in high-income countries, where more such analyses are available, their effect has been limited, although it is growing (Gabbay and le May 2004; Glick, Polsky, and Schulman 2001; Hoffmann and others 2002; McDaid, Cookson, and ASTEC Group 2003; Sheldon and others 2004; Taylor and NICE 2002). Cost-effectiveness studies are now required by, for example, the U.S. Food and Drug Administration for labeling claims, the National Institute for Clinical Excellence before advising national policy on treatments and care in England, and the Ministry of Health in the Netherlands for new drugs (iMTA 2005).

DEFINITION AND CHARACTERISTICS OF INTERVENTIONS

The object of a CEA—the thing to which it is applied, the costs and outcomes of which are to be analyzed—is an intervention. An *intervention* is an activity using human, physical, and financial resources in a deliberate attempt to improve health by reducing the risk, duration, or severity of a health problem (Jamison 2002, table 2). The term usually refers to an activity undertaken by a health system rather than by an individual. The emphasis on a deliberate, systemic effort means that an intervention is not simply anything that improves health; for example, if more rainfall leads to higher crop yields and better nutritional status, the rain does not count as an intervention. Similarly, although breastfeeding protects infants' health, it is not itself an intervention as the word is used in this volume. In contrast, a program to encourage new mothers to breastfeed *is* an intervention (as described in chapter 27). How effective such a program is, of course, depends on how many mothers it persuades to adopt the practice when they are neither currently breastfeeding nor planning to do so.

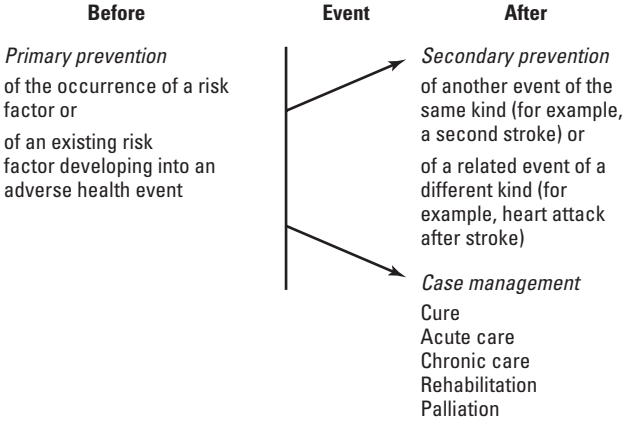
Interventions can be directed against an injury or disease (such as trachoma), a condition associated with or deriving from a disease (such as blindness), or a risk factor that makes the disease or condition more likely (such as the lack of hygiene that leads to trachoma). An intervention may pursue primary prevention at the population level—promoting personal behavior change, controlling environmental hazards, or delivering a medical intervention such as immunization to a large population—or individual action for primary prevention, cure, acute management, chronic management, secondary prevention, rehabilitation, or palliation. Box 15.1 defines these terms, and the figure in the box illustrates how interventions may prevent ill health events or deal with their consequences. Characterizing an intervention fully also means distinguishing the level at which it is delivered (home, primary care facility, district hospital, or referral hospital); indicating whether it involves drugs, immune enhancement, surgery, or physical or psychological therapy; and determining whether it requires a physician or uses diagnostic, laboratory, or imaging procedures. Such procedures are most often evaluated relative to the interventions they screen for or lead to, because they produce no health gain by themselves (although the information they provide can be valuable for reassurance or for promoting behavioral changes).

An intervention in the everyday sense includes such activities as immunizing a child, performing a surgical procedure, or treating an infection with antibiotics. The authors of some chapters use the term only in this sense—for example, in discussing interventions that contribute to meeting the Millennium Development Goals (chapter 9). Authors of other chapters use the term in several other senses as well. It can mean modifying an existing intervention—for example, adding *Haemophilus influenzae* type B (Hib) antigen to the Expanded Program on Immunizations (EPI). Immunization against Hib is itself an intervention, but instead of analyzing it separately, one can use CEA to evaluate the additional cost of incorporating that antigen and the additional health gain that

Box 15.1

Intervention Categories, with Examples

The following figure illustrates how interventions are related to a health event; the definitions of these categories are given below.



Interventions Related to the Occurrence of an Adverse Health Event

Population-based interventions all aim at primary prevention (as defined later), are directed to entire populations or large subgroups, and fall into three categories:

- Promoting personal behavior change (diet, exercise, smoking, sexual activity)
- Control of environmental hazards (air and water pollution, disease vectors)
- Medical interventions (immunization, mass chemoprophylaxis, large-scale screening, referral).

Personal interventions are directed to individuals and can be intended for the following:

- Primary prevention—to reduce the level of one or more risk factors, to reduce the probability of initial occurrence of disease (medication for hypertension to prevent stroke or heart attack), or to reduce the likelihood of disease when the risk factor is already present (prophylaxis for sickle cell anemia).
- Secondary prevention following the occurrence of disease—either to prevent another event of the same kind or to reduce the risk of a different but related event (medication to reduce the likelihood of a second coronary event or a first heart attack after stroke).
- Cure—to remove the cause of a condition and restore function to the status quo ante (surgery for appendicitis)
- Acute management—short-term activity to decrease the severity of acute events or the level of established risk factors, to minimize their long-term impacts (thrombolytic medication following heart attack, angioplasty to reduce stenosis in coronary arteries).
- Chronic management—continued activity to decrease the severity of chronic conditions or prevent deterioration (medication for unipolar depression, insulin for diabetes). Chronic management can include some secondary prevention.
- Rehabilitation—full or partial restoration of physical, psychological, or social function that has been damaged by a previous disease or condition (therapy following musculoskeletal injury, counseling for psychological problems).
- Palliation—to reduce pain and suffering from a condition for which no cure or rehabilitation is currently available (analgesics for headache, opiates for terminal cancer).

Source: Authors.

is expected to result (see chapter 20). The intervention studied is then not Hib immunization as such but the change in the full vaccination procedure. A change in the scale of an existing activity can also be considered an intervention, even if the activity itself is unchanged: that is, one can analyze the change in costs and in outcomes associated with expanding or contracting the coverage of the activity—for example, extending antiretroviral treatment for HIV and AIDS to a larger population (chapter 18) or screening more newborns for sickle cell anemia (chapter 34). In most chapters, the authors assume that expansion affects costs and outcomes linearly, so that the CER does not change. The chapters on vaccine-preventable diseases (chapter 20) and malaria (chapter 21) provide explicit estimates of the differential costs of expanding coverage.

Adding one intervention to another to deal with the same disease or condition is also an intervention, and combinations of interventions can be analyzed to determine which is most cost-effective or how the cost-effectiveness of one intervention depends on the other activities with which it is combined. Examples include successively adding drugs for treatment of epilepsy (chapter 32) or secondary prevention of cardiovascular disease (chapter 33) or combining several quite different interventions to control tobacco addiction or alcohol (chapters 46 and 47, respectively). The analyses of community health and

nutrition programs (chapter 56) and integrated management of infant and childhood illness (IMCI; chapter 63) define "the intervention" as a whole program incorporating several different activities. Generally, even less empirical evidence exists concerning combinations of interventions than for individual activities, but IMCI is an exception; it has been evaluated more thoroughly than most single interventions.

Box 15.2 includes a more detailed discussion, using a hypothetical example of three different ways to deliver immunization, of how CEA can be applied to four of the meanings of *intervention* used here: an existing intervention at its current coverage, changes in the scale of that intervention, the addition of one intervention to another when expanding coverage, and the complete shift from one intervention to a different (and more cost-effective) one.

Depending on the comparison undertaken, the result may be an average cost-effectiveness ratio (ACER) or an incremental cost-effectiveness ratio (ICER). The former compares total costs and total results, starting from zero, whereas the latter compares additional costs and additional results, starting from the current or some other level of coverage of an intervention.

Either shifting completely from one intervention to another or partially replacing one with another may reduce costs while producing more health gain. For example, if spending is high on hospitalization for acute myocardial infarction, a program using a "polypill" (several medications in a single pill) would reduce expenditures by lowering incidence (chapter 33) and would be cost saving, because less hospitalization would be needed. If the status quo is no hospitalization (as is typical at low incomes), a polypill program increases costs but may more than correspondingly increase health gains and therefore be more cost-effective. If the polypill both reduces costs and improves outcomes compared with hospitalization, it is said to dominate a hospital-only strategy. The second figure in box 15.2 illustrates the concept of dominance; table 45.4 and box 45.1 of chapter 45 provide examples of interventions that are dominated by others.

Unfortunately, reliable information on current intervention coverage, costs, and results is not always available even in high-income countries (iMTA 2005) and is extremely scarce in low- and middle-income countries. Studies showing whether an intervention is effective or cost-effective seldom cover the entire potential beneficiary population, and service provision in the private sector is often not recorded. Many chapter authors describe only the ACER of an existing or potential intervention, whereas others explicitly compare alternatives to current practice (for an example, see chapter 16).

Many of the activities analyzed here aim at promoting changes in personal behavior, by informing and persuading individuals to eat differently, to avoid smoking and excessive alcohol, to reduce the risks of sexually transmitted infections, or to practice better hygiene. Such efforts can be considered interventions in themselves, and as such are crucial for controlling HIV and AIDS (chapter 18), promoting better infant and child care (chapters 20 and 27), preventing inherited disorders (chapter 34), encouraging healthful diets and exercise (chapters 44 and 45), and avoiding addiction (chapters 46–48). But they can also be used to improve the effectiveness of other interventions by increasing awareness and demand, combating mistaken beliefs about diseases and risks, or reducing anxiety and stigma. In that sense, information, education, and communication delivered to consumers or providers or both are examples of policy instruments. They can facilitate or promote the use of such interventions as condom distribution, screening for diseases or congenital disorders, prenatal care, or immunization.

Other activities that can be classified either as interventions or as policy instruments include the following:

- *Measures to increase the quality of care,* such as some kinds of staff training or the introduction of better recordkeeping. These activities may simultaneously affect a large number of specific interventions in a health facility (chapter 70).
- *Legislation and regulation* to impose an intervention (for example, limiting the salt content of foods, chapter 45, or requiring that salt be iodized, chapter 28); to limit or prohibit an intervention that is ineffective or dangerous or to reduce unhealthful behavior such as smoking and excessive drinking (chapters 46–47); or to codify how an intervention should be delivered and determine who may provide it, as by licensing doctors, nurses, and health facilities (chapter 71).
- *Economic incentives,* which can take the form of subsidies or taxes (chapter 11) for particular items of consumption other than health goods or services, such as tobacco and alcohol (chapters 46–47) or condoms to reduce HIV transmission (chapter 18), or can be provided through protection of property rights, as for patented drugs (chapter 72).

These activities of informing, mandating, legislating, regulating, and taxing or subsidizing, which are at one remove or more from medical interventions, are also often called functions of the health system (WHO 2000, chapter 2; see also chapter 9 in this volume). Several of these instruments may be used together, such as increased taxes on tobacco or alcohol along with measures to educate consumers and to restrict the times, places, or quantities of consumption. Sometimes the instrument is needed before introducing or expanding an intervention to overcome barriers to its use or to make it cost-effective enough to be worth pursuing. Educating the affected population, for example, is crucial to screening and treatment of cancers and hemoglobin disorders. The need for a particular instrument may vary from country to country even if the

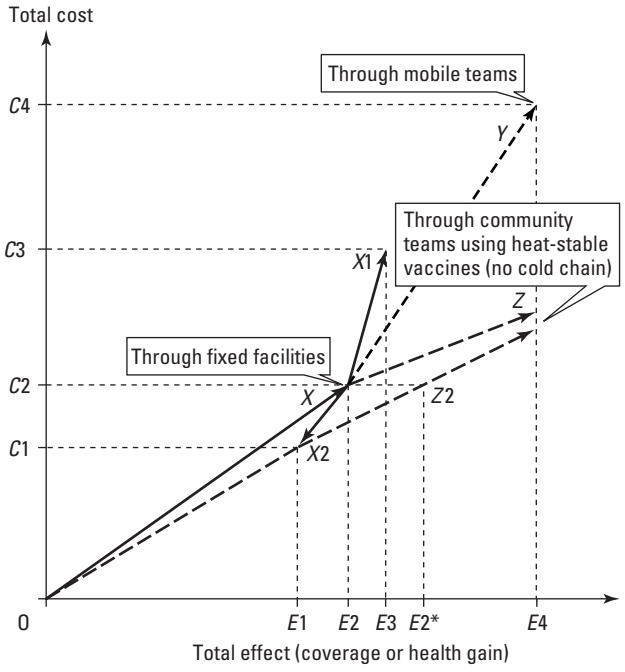
Box 15.2

Average and Incremental Cost-Effectiveness and Intervention Choices

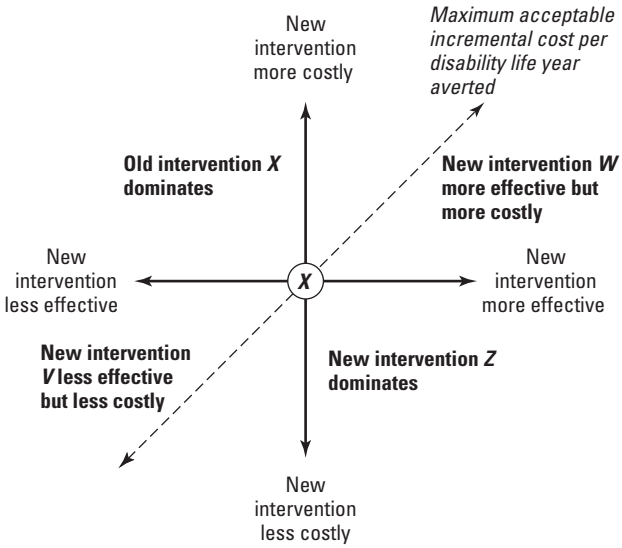
In the figure below, which compares three ways of delivering immunization, point *X* describes the status quo of a current intervention, delivering immunization by means of fixed facilities. At point *X*, the intervention achieves a total effect *E2* (measured as coverage or as disease reduction) at a total cost *C2*. The ratio *C2* to *E2* is the average cost-effectiveness ratio (ACER), shown by the slope of the line *O–X*. Beyond point *X*, expanding coverage becomes very costly, perhaps because the population not yet immunized is dispersed and hard to reach. (Chapter 20 includes estimates of how costs increase as immunization coverage expands but without introducing a sharp increase in costs.) Expansion to point *X1*, which increases the cost from *C2* to *C3*, yields only a small increment *E3–E2* in effect. The slope of the line *X–X1* represents the incremental cost-effectiveness ratio (ICER) of that expansion, which would raise the ACER to line *O–X1*. The line *X–X2* shows the alternative of reducing coverage, which would improve the average cost-effectiveness (to *C1*/*E1*) because marginal costs are rising steeply near point *X*. The ICER of the reduction in coverage is the ratio of *C2–C1* to *E2–E1*.

Raising immunization coverage at an affordable cost may require adopting the alternative of mobile vaccination teams, intervention *Y*. The hypothetical combination of fixed facilities and such teams allows increasing the effect to *E4* (complete or nearly complete immunization) at a total cost of *C4*. The ICER of the mobile teams is shown by the slope of the line *X–Y* and the resulting overall or combined ACER by the slope *O–Y*. Adopting intervention *Y* would be clearly preferable to trying to expand coverage through intervention *X* by building and staffing more fixed facilities.

An alternative even better than *Y* might subsequently be developed, represented by point *Z*—for example, community-based immunization teams that could operate either near or far from fixed facilities because they use heat-stable vaccines that do not require a cold chain. The ICER of turning to that choice, represented by the line *X–Z*, is not only more favorable than intervention *Y*, but it is even better than the current ACER, and preferable to intervention *X* at any coverage level beyond *X2*. The cost-effective choice, therefore, is not to retain intervention *X* at its current level and add *Z* beyond that point but to switch entirely from *X* (or from *X* plus *Y*, if *Y* has already been adopted) to *Z*. Because it costs less but provides a better outcome, *Z* is said to *dominate* both *X* and *Y*. The following figure illustrates dominance of one intervention by another, as well as cases in which neither of two interventions is dominant.



Average and Incremental Cost-Effectiveness and Intervention Choices: Comparison of Three Ways to Deliver Immunization



Comparison of Cost and of Effectiveness between Interventions: Conditions for Dominance

If intervention *Z* is divisible (meaning that it can be operated at any desired scale, such as *Z2*), then it is preferable to *X* at a cost of *C2* because of the additional effect *E2**–*E2*. It can be extended all the way to *E4*, just as with intervention *Y*, provided only that the ICER represented by the slope *X*–*Z* is still acceptable to decision makers choosing how far to expand the intervention. That is, the cost must still appear to be justified by the increased coverage. Under either of these conditions, an obstacle to switching, or to doing so quickly, would exist only if substantial fixed costs accompanied the transition from one intervention to the other, such as recruiting or retraining staff, building health posts in communities, or setting up the system for distributing the new heat-stable vaccines.

Compared with intervention *X*, intervention *Z* is better in both dimensions (lower cost and greater effectiveness), so it is to be preferred, and is said to dominate *X*. However, intervention *X* would dominate any other treatment that is both more costly and less effective and, therefore, falls in the upper left quadrant. An intervention such as *V* or *W* may or may not be considered preferable to *X* (*V* is cheaper but also less effective, and *W* is more effective but also more costly). Whether either such intervention would be selected over *X* depends on the relation of the increased (or decreased) cost to the increased (or decreased) effectiveness. That ratio corresponds to an ICER. If a maximum acceptable, or threshold, value for the ICER is determined, as shown by the dashed diagonal line, then any intervention that falls below the dashed line would be acceptable (preferable to *X*), and those that fall above the dashed line would not be. Uncertainty about the estimates of cost and effectiveness means that, instead of a sharp line as in the figure, the division of preferable from nonpreferable interventions corresponds to a zone of some width that depends on the confidence intervals around the estimates. This kind of comparison can start from an existing intervention such as *X* in the first figure or, when there is currently no intervention, from point *O* in the first figure.

Source: Authors.

intervention that it facilitates is identical, because the legal, regulatory, or financial environment differs.

ESTIMATING EFFECTIVENESS IN HEALTH

Using cost-effectiveness for resource allocation requires health effects to be represented in common units in order to facilitate comparison across interventions, diseases, or conditions. All analyses start with some natural unit: cases of disease or injury, deaths, or numbers of people who quit smoking or adopt some other health-improving behavior. All interventions that avert death are alike in that regard. Preventing a child's death at a particular age, independent of the cause, means that the averted death alone is an adequate measure of outcome. However, when lives are saved at different ages—averting death from malaria at age 2 versus death from beta thalassemia at age 10–15—the outcome is no longer identical, and some measure must account for the difference in years of life saved. These cases provide another natural unit, subject to estimating how much longer a person spared death might live. The choice of life expectancy to assume for such calculations is discussed later.

The unit of time becomes a less natural and more synthetic measure if the future is discounted, as in all these analyses. *Discounting* means reducing the value of each variable in each future year by an amount that increases the further in the future that year is. The discounting procedure reflects inherent uncertainty about the future and preferences for timing of consumption, and it avoids two problems. First, outcomes that potentially generate benefits forever, such as smallpox eradication, appear to have infinite benefits if the future is not discounted and therefore seem to justify any finite cost at all. Second, it makes little sense to postpone interventions forever simply because funds to finance them could be invested today and be worth more tomorrow. Even discounting the future at the low rate of 3 percent annually has a substantial effect—that is, dividing the values for future years by successive powers of 1.03. That means dividing values for year 1 by 1.03; those for year 2 by 1.03 squared, or 1.0609; and so on. At that rate, averting an infant death saves not all the 80 calendar years of life expectancy at birth (or fewer in low-income countries) but at most 30 discounted years.

For interventions that avert mortality, analysis starts by estimating the deaths prevented, uses age at death to yield numbers of life years saved, and then discounts those years as described above. When interventions improve health by averting or reducing nonfatal disability, different disabilities must be compared in severity. As with mortality, age at the time of intervention matters for long-lasting conditions, and so does discounting. In contrast, age is irrelevant for episodes of illness or injury that are self-limited or quickly resolved by intervention, because the duration of ill health does not depend on age, and all ages are treated alike in this analysis. Discounting also makes little difference over short intervals.

Unit of Measurement of Health

The common unit of health loss or gain used here takes into account duration and severity, as well as discounting the future. The disability-adjusted life year (DALY) is a unit introduced by the World Health Organization (WHO) and the World Bank (Jamison and others 1993; Murray 1996; World Bank 1993). As previously discussed, the DALY incorporates assumptions and measurements about severity of nonfatal conditions, age at incidence or intervention, duration with and without intervention, and remaining life expectancy at that age. For interventions directed to risk factors rather than diseases, the analysis incorporates estimates of reductions in diseases that result from changes in the level of risks. Smoking cessation, for example, reduces deaths from both cardiovascular disease and cancer (chapter 46).

Published analyses, particularly in high-income countries, often use not DALYs but quality-adjusted life years (QALYs), an alternative measure of how much a year of life is worth if a person suffers one or more limitations of various kinds and degrees. QALYs can be estimated directly using a valuation method such as the *time tradeoff* (comparing and varying the time spent in one health state with time spent in another state until the quality of life is judged the same in both). Alternatively, a prescored questionnaire such as the EQ5D (a European quality of life measure) could be used. The EQ5D distinguishes three grades—no problem, an extreme problem, and total disability—on each of five dimensions of life quality—mobility, self-care, performance of usual activities, pain or discomfort, and anxiety or depression (Brooks, Rabin, and de Charro 2003). Discounting of QALYs occurs as an additional step, although some concern exists that discounting values derived from the time tradeoff approach is double discounting (Dolan and Jones-Lee 1997).

QALYs allow comparison among interventions and can easily account for comorbidity. Although the concept of DALYs averted by an intervention is similar to that of QALYs gained, no systematic formula exists for converting between DALYs and QALYs except in broad approximations (Fox-Rushby 2002). This gap is partly because DALY disability weights are specific to diseases whereas the QALY system of evaluation is not (it is based on overall health status). Authors sometimes report effectiveness results in QALYs, because they cite studies in high-income countries that often use QALYs. When some interventions are evaluated in DALYs and others in QALYs, ranking interventions according to cost-effectiveness may still be possible (see chapter 29 for examples of the use of both units).

Priority setters sometimes stop exercises in priority setting after concluding that something is or is not cost saving, without asking whether an intervention yielding a different outcome (against a different disease, for example) would be still more cost-effective. In this volume, the intent is to estimate both costs and effects, permitting all such comparisons. Knowing that one intervention achieves the same results as another at lower cost, which would be relevant if they were the only two possible interventions against a common problem, is not enough. Comparing both with another intervention with different effects may also be necessary. For example, a coronary artery bypass graft for myocardial infarction costs, on average across regions, US$37,000 per DALY gained, compared with an average of only US$409 for the polypill discussed earlier. However, both are much more expensive than saving life years for a middle-aged person by treating active tuberculosis (and thereby preventing transmission), an intervention that costs only US$15 per DALY in the absence of HIV infection, or US$102 on average where coinfection makes treatment more difficult. (In only a few cases do chapters deal explicitly with comorbidity, in part because the DALY approach considers conditions only individually.)

Parameter values for effectiveness are required in order to conduct CEA: how to value disability, compared with mortality; how to treat the future; and whether to distinguish people according to age, sex, or other characteristics. Because effectiveness is related to reduction in disease burden, nearly all these parameter choices coincide with those adopted to estimate the existing burden (see Mathers and others 2005 for a full explanation).

Because disease burden estimates discount the future at 3 percent annually, CEA in this volume does the same, for both effects and costs. This method follows the recommendations of the U.S. Public Health Service Panel on Cost-Effectiveness in Health and Medicine (Gold and others 1996) and appears appropriate whenever the benefits of an intervention begin immediately. Constant discounting (using the same percentage rate each year) undervalues interventions for which the benefits appear long after the costs have been paid. Immunization against hepatitis B can prevent liver cancer decades later (chapter 20) but, compared with the costs incurred at the moment of vaccination, appears less cost-effective if the health gain is heavily discounted during that interval. Slow discounting, with the rate falling close to zero for the more distant future, would yield a higher present value of benefits (Jamison and Jamison 2003), but given the absence of consensus on the correct form, the analyses here use constant discounting.

The limitations from a disease or condition in the absence of intervention are measured by disability weights (Mathers and others 2005), despite some controversy as to whether they adequately capture all the disability (see chapter 24 on helminthic infections). These weights range from zero for perfect health to 1.0 for death. Authors have made their own estimates whenever WHO did not provide any disability weight because the corresponding condition was not explicitly included in the burden of disease. For example, weights for anemia caused by hemoglobin disorders (chapter 34) were taken from other

causes of anemia. Note that years lost to early death also are DALYs, since they include the disability weight of 1.0.

When an intervention prevents or completely cures a condition, the postintervention disability is zero. For partially successful interventions leaving residual disability, the disability weight is reduced but not eliminated. WHO has sometimes estimated weights for "treated" as opposed to "untreated" conditions (Murray and Lopez 1996, annex table 3) without specifying the intervention. This distinction is introduced for some consequences of chronic conditions: cancers not yet in the terminal stage, diabetic conditions, major psychological disorders, cataracts, various cardiovascular conditions, chronic respiratory conditions, ulcers, arthritis, cleft lip and palate, edentulism (total loss of teeth), and some burns. Chapters 31 and 33 use these values to describe intervention outcomes.

WHO burden-of-disease estimates used in the first edition of *Disease Control Priorities in Developing Countries* (Jamison and others 1993) incorporated age weights—that is, numbers attempting to describe the relative value of life at different ages. These numbers were calculated to keep the discounted integral from age 0 to 80 the same, as if no age distinction were made. The weights are zero at birth, ignoring health losses from stillbirth prior to live birth; reach a maximum at age 25; and decline almost to zero at advanced age. They are a particularly controversial element in the burden estimates (Musgrove 2000) because they value some years of life more than others, and little evidence suggests what an appropriate weight should be. In consequence, only constant age weighting (treating all years alike) is used in these analyses. Removing age weights makes no difference to an intervention that averts an infant death, but it changes the relative importance of interventions at later ages. Because life is more highly valued at advanced ages, death and disability after age 38 become more important compared with events before that age, and interventions later in life become more cost-effective. Some estimates used here therefore differ from those published previously by WHO even when all the other parameters are unchanged.

The only parameters for CEA that differ from those in the burden of disease concern life expectancy. In estimating burden, people at any age and in all regions are assumed, on the ethical criterion of valuing all lives equally, to have the same life expectancy. The only exception is that at birth males appear to have a biologically determined (not behavior-related) life expectancy of 80 years, which is shorter than the life expectancy of females by 2.5 years. However, applying these expectancies to CEA will overstate the effects of interventions when life expectancy in a population is low. Averting a death at age 5 in Sub-Saharan Africa or South Asia does not confer a high probability of living to age 80 or longer. Competing causes of death reduce the effectiveness of any single intervention, unless it affects so large a population that it actually increases life expectancy. An intervention that completely interrupted the transmission of HIV and AIDS or prevented all deaths from malaria would do that. Given the absence of evidence that any intervention actually has such a substantial effect, it is assumed in this volume that individuals face the same probability of death at each subsequent age as the existing population does.

This assumption makes interventions appear less effective when overall mortality is high than when mortality is low. For example, averting an infant death in Sub-Saharan Africa will save, on average, only 44 to 49 undiscounted life years and should not be credited with saving 80 or more. Cost-effectiveness calculations and estimates of burden of disease are inconsistent in that fully effective interventions appear able to deal with only part of the burden they aim to control. Regional rather than standard life expectancy also makes interventions in a high-mortality region appear more effective relative to outcomes where mortality is lower, when they avert deaths later in life.

Nonhealth benefits of health interventions sometimes should be taken into account, because many health interventions also yield other kinds of benefits. They often make beneficiaries mentally or physically more productive, better able to continue in and learn from school or to work and earn more. This benefit occurs particularly with interventions against nonfatal consequences, as indicated in the chapters on malnutrition (chapter 28), malaria (chapter 21), helminthic infections (chapter 24), tropical diseases (chapters 22–23), psychiatric disorders (chapter 31), and learning and developmental disorders (chapter 49). Interventions that prevent injury or restore work capacity also have such effects (chapters 39–40 and 60), as do interventions against diseases that kill in the prime of life, notably tuberculosis and AIDS (chapters 16 and 18). Nonhealth benefits also occur as time is saved when piped water is made available, as less cleaning is needed when air pollution is reduced (chapter 42), or as property damage is reduced by improved traffic safety (chapter 39).

Several chapters include discussions of the nature and, where possible, the magnitude of nonhealth benefits from health interventions. This factor is important when the health benefits, although substantial, are so costly that interventions do not appear cost-effective on health grounds alone but may be justified by large nonhealth returns. Safe water and sanitation services are the classic example (chapter 41). Because different types of benefits—health gains, increased income, time saved—cannot be compared directly, the only way of combining them into a single expression is usually to evaluate all outcomes in monetary terms. (An exception occurs when some monetary gains can be measured directly—for example, increased worker productivity from better health. Those gains can be subtracted from costs and incorporated into CEA without attributing a monetary value to the health gains.) Most chapters that account for any nonhealth benefits simply offer descriptions of them rather than incorporating them into

monetary outcome indicators. Chapter 7 is an exception; it compares gains in welfare from living longer or in better health with those from higher income.

Estimating the monetary value of all benefits and adding them together for comparison with cost is what cost-benefit analysis does. Interventions are considered justified in absolute terms if the benefits exceed the costs. However, when faced with constrained budgets that cannot finance all interventions whose benefits are greater than costs, policy makers need to establish some minimum acceptable rate of return. This choice is parallel to the need to set a maximum on cost per unit of health gain when choosing according to cost-effectiveness.

Published analyses of health interventions sometimes use cost-benefit analysis, so results following that method are incorporated in some chapters here. The decision to emphasize CEA instead derives from two chief considerations. One is that, for most interventions, the health consequences seem more relevant or more important than any nonhealth outcomes. The other is that, conceptually, it is unclear what dollar value to assign to improved health, as would have to be done in most cases. Two approaches to valuing health, particularly for judging how much a life is worth, are known respectively as (a) the *value of a statistical life,* or the *human capital approach,* and (b) the *willingness-to-pay valuation*, or *contingent valuation*. The former depends on estimating earnings lost from premature death or retirement, and the latter on what people pay or indicate they would pay for care to protect or restore their health.

Both approaches reflect a society's level of income. Although they may be appropriate within a homogeneous society, if applied globally they imply that better health is worth less among poor populations than among those who are better off. Both methods are arguably more appropriate for marginal improvements (like saving travel time for commuters) than for valuing life-or-death differences, although willingness to pay is sometimes used in analyses of policies to reduce mortality. Avoiding monetary evaluation of health benefits sidesteps most of the ethical problems of valuing individual lives and requires fewer assumptions about what benefits are worth. The cost of this simplification is that occasionally substantial nonhealth benefits are not explicitly valued, so interventions may look less justified than they would be if all benefits were analyzed.

DETERMINING COSTS FOR INTERVENTIONS

Whatever outcome measures are used to evaluate an intervention, its costs must be estimated. This need raises several questions about which costs to attribute to the intervention and how some of them should be valued.

Direct and indirect costs should be distinguished, and choices should be made about which, if any, of the latter to include. In addition to the direct costs to the health system of producing an intervention, the U.S. Public Health Service guidelines (Gold and others 1996) recommend including the indirect costs to patients and their families of consuming it. This recommendation means, in particular, the value of time needed for travel, waiting, and undergoing medical tests and procedures, or the value of time used in caregiving, as well as any income forgone during treatment. Externalities, or costs imposed on third parties, such as on the school system or the environment, should also be included. The analyses in this volume generally exclude such costs and report only the direct costs of delivering interventions, partly because published analyses seldom include the various indirect costs, and they are harder to estimate. Walker and Fox-Rushby (2000) found that only 20 of 101 studies included some element of indirect costing. Valuing time according to local wages or income, for example, may underestimate how valuable time actually is to poor people. Estimating such costs, even if time is not valued in money, may show whether time or monetary costs or both account for a relatively low level of use and therefore impede expanding coverage. Applying one or more of the policy instruments discussed earlier, along with the intervention, may then be important in order for it to be cost-effective.

Including such costs also raises a question of interpretation. If an intervention appears low in cost-effectiveness because it requires much travel or waiting time, the fault may lie not with the intervention itself but with health facilities that are located too far from the beneficiary population, are understaffed, or are inefficiently managed. For this reason, cost-effectiveness is estimated assuming a functional health system that does not impose prohibitive time costs on users.

Not only the characteristics of the interventions themselves, but also the capacity to deliver interventions greatly affect cost-effectiveness across many activities. In a complete analysis, each intervention is characterized by how demanding it is of managerial or institutional capacity. This element is difficult to measure directly, but authors often provide at least an intuitive description of how easy or hard delivery of an intervention is or what factors facilitate or impede its implementation. Where capacity to deliver several interventions together is important, authors deal explicitly with the issue, as in the chapters on health facilities (chapters 64–66), resources (chapters 71–72), service management (chapter 73), and whole packages of interventions (chapters 56 and 63).

Dollar values of unit costs need to be calculated for international comparisons. The inputs used to produce an intervention—the time (and training) of human resources; drugs and supplies; and depreciation or rental value of equipment, vehicles, and buildings—are either produced in the country or imported. If the latter, they already have prices in U.S. dollars; if the former, prices in local currency must be converted to U.S. dollars for comparison with other interventions and other countries. The usual distinction between tradable

and nontradable goods is that tradables move from producing to importing countries at relatively constant "world" prices. In fact, the same good may be imported at different prices to different countries or may be imported to one country but locally produced in another, so that it has both an international and a local price. This situation is increasingly true of drugs and supplies, which middle-income countries (Brazil, for example) and some low-income countries (India) now produce and sometimes export.

Prices in local currency can be converted to U.S. dollars by exchange rates or by purchasing-power parity rates (as estimated in World Bank 2003). The former may reflect under- or overvaluation of the local currency, making goods systematically cheaper or more expensive than at world prices, and they may change quickly and substantially in response to changes in a country's trade balance, indebtedness, or capital flows. Nonetheless, they represent what is actually paid for locally produced inputs at any given moment. Purchasing-power parity rates, in contrast, attempt to say what local currency is worth in purchasing power, correcting for systematic price differences. Such rates can be calculated for the country as a whole, for the health sector, or for specific inputs or combinations thereof (Wordsworth and Ludbrook 2004). This calculation means valuing local inputs at external prices, assuming they are equally productive or of equal quality in the particular country as in the countries from which purchasing-power parity rates are derived. A doctor in South Asia or Sub-Saharan Africa is treated as costing just as much as a physician in high-income countries. This approach approximates measuring the real resource cost of intervention by comparing *quantities* of inputs among countries, eliminating *price* differences as a source of cost variation. Estimates of real national income are derived this way, making poor countries usually look less poor in dollar terms than if income in local currency were valued at exchange rates.

Granted that purchasing-power parity rates are reasonable for comparing large aggregates such as income across countries, but they bear little relation to the allocation of resources and budgetary choices within a country. The cost calculations in this volume are, therefore, all based on exchange rates. Exchange rates more accurately reflect what a domestic buyer—or a foreign donor or investor—has to pay for imported versus domestic inputs and, therefore, are more relevant for choices between interventions with high or low imported content. (If exchange rates are artificially fixed, the country pays a cost for that distortion that affects all interventions to the extent that they require foreign exchange.) In general, the more an intervention is produced with local inputs, the more cost-effective it will appear when priced using exchange rates, compared with its cost at purchasing-power parity rates. For decision makers and purchasers in the country, efficiency means choosing interventions according to what they actually cost, not according to what they would cost if prices were more nearly uniform among countries. If, in local currency, physicians are paid little more than nurses are, it may make sense to employ more doctors per nurse—even if at international prices doctors would cost much more and should be replaced by nurses when possible. Of course, the staffing decision turns on the competencies of the two groups as well as on their costs; for certain health problems, more nurses might be the better choice even if they cost more.

Two other reasons besides that of efficiency in buying interventions support basing cost-effectiveness on exchange rate prices. First, authors who have used published costs (which usually involve exchange rates) rather than building up estimates from individual prices and quantities seldom break down costs into imported and domestic components. Local inputs cannot be repriced at purchasing-power parity rates or can only be repriced very approximately. Second, for readers accustomed to dealing with prices converted using exchange rates, real resource estimates may simply appear to penalize the use of local inputs by valuing them at unrealistic prices. The problem with exchange rate prices, in contrast, is that when rates change, so may the relative cost-effectiveness of interventions, as imported inputs become relatively more or less expensive. Cost-effectiveness is not static or intrinsic but depends on prices as well as on quantities and on the results of an intervention—and prices can change individually or generally, through exchange movements. Priorities sometimes need to shift because of such price changes, as well as because of technological changes that make interventions more effective, and so analyses should be kept up to date.

MORE AND LESS COMPREHENSIVE DATA AND ANALYSIS

Several authors (Drummond and others 1997; Gold and others 1996; Sloan 1995) provide similar guidance and recommendations for relatively comprehensive economic evaluation in general or for specific medical procedures. This volume aims at estimating cost-effectiveness for interventions against many different problems in all low- and middle-income regions, for which varying amounts and quality of information are available. It has therefore not always been possible to conduct as complete an analysis as would be desirable. Some degree of modeling is usually inescapable (Buxton and others 1997).

More complete analysis starts by characterizing, in each regional setting where an intervention is relevant (where the health problem causes some measurable burden and the intervention appears feasible), estimates of the quantities of inputs required (Q), the unit costs of those inputs (P), and the effectiveness or health gain (E). Authors were provided regional estimates of unit costs for the major inputs—salaries, facility costs,

fuel and vehicle operation, drugs, representative equipment, diagnostic tests, and buildings (Mulligan and others 2003). The total cost of delivery is the sum of the input costs *PQ*, which is compared with effectiveness *E* and the CER calculated from the total costs and total effects of the proposed intervention or from the changes in those costs and outcomes compared with current practice.

The data on unit costs, quantities, and outcomes may all derive from published literature; what is original is how that information is combined to calculate cost-effectiveness rather than taking the ratios from existing studies. Estimates are built up using prices and physical inputs in the chapters on tuberculosis (chapter 16), vaccine-preventable diseases (chapter 20), malaria (chapter 21), cancers (chapter 29), psychiatric disorders (chapter 31), neurological disorders (chapter 32), cardiovascular disease (chapter 33), hemoglobin disorders (chapter 34), water and sanitation (chapter 41), indoor air pollution (chapter 42), tobacco (chapter 46), alcohol (chapter 47), community programs (chapter 56), family planning (chapter 57), surgery (chapter 67), emergency care (chapter 68), and complementary medicine (chapter 69). Several chapters analyze some interventions more fully and others less fully, depending on the available information.

As indicated in box 15.2, expanding or contracting the scale of an intervention may change the CER because of difficulty in reaching more of the population. The ratio may also vary because of the cost of identifying who would benefit most from the intervention—for example, whether to screen all newborns for sickle cell disease or only those of African origin (chapter 34). And expansion may change the cost-effectiveness because it would require considerable fixed investment. The costs of expanding capacity to deliver an intervention, including physical capital and training of human resources, should be amortized over a reasonable interval (10 years is the standard in this volume) and included in the total costs. Ideally, one would know the complete production function of the intervention, including the possibilities of substituting one input for another to minimize costs in response to differences in prices. However, analysis of this level of complexity is difficult to achieve, so most chapters assume fixed input proportions. *Q*, then, does not depend on *P*, and the CER varies (at most) only with coverage, prices, and outcomes. This result could be an underestimate of the true cost-effectiveness if much substitution is possible (see chapter 16 on tuberculosis).

Approximations are required when the average and incremental CERs have to be taken directly from the literature and when key parameter values are not easily available. Existing estimates of total cost or effectiveness may or may not incorporate the standard assumptions about discounting, disability weights, and life expectancy. Authors then need to judge how to adjust the available estimates for a more consistent analysis.

Local cost and outcome estimates that have not been constructed transparently from inputs and prices provide a less complete basis for secondary analysis. Where such estimates are used, information about how costs are constructed or how results vary with the scale of the intervention is usually not available, but the data may explicitly show regional differences in one or both elements, thereby permitting regionally differentiated recommendations (or may show differences so small that recommendations need not differ regionally). If costs and results refer to only one moment or are specified year by year, they can be discounted at 3 percent. Published analyses often use higher constant rates of 5, 6, or even 10 percent and may specify only total costs and outcomes rather than the respective streams through time. In that case, both costs and health gains occurring in the future are valued less, but conversion to a CER based on 3 percent discounting may be impossible and is at best only approximate. Some published analyses discount costs but not health outcomes, which makes interventions look more cost-effective when costs are spread over long intervals (for examples, see chapter 29). Imported estimates of cost and effect—that is, estimates from other regions, commonly from high-income countries—are often all that is available. Sometimes data on costs and outcomes derive from the same source; in other cases, they come from different sources and even different regions and are difficult to compare directly. More appropriate adjustments to total costs are possible with information on quantities of inputs. In the absence of data on quantities of resources used, differences in average cost can sometimes be calculated using estimates of input proportions. Such an approximation characterizes the analysis for diabetes (chapter 30), in which proportions were known in one region and assumed to be the same elsewhere and costs were estimated from regional cost ratios.

Variation of Results and Uncertainty of Estimates

Variation and uncertainty are two different aspects of cost-effectiveness estimates that also need to be accounted for in setting priorities. Because costs of inputs differ among regions, intervention costs vary even if effectiveness does not—and there are often reasons why the same intervention is more effective in one place than another. Such variation means that a single estimate of incremental or average cost-effectiveness of an intervention is not universally applicable. All estimates should ideally be local, and regional values capture only part of the real variation. For example, the average cost per DALY of chemotherapy for active or contagious tuberculosis, in the absence of HIV and AIDS, is US$15, but that figure varies from US$6 to US$31 across regions, and such wide variation is common in many chapters. Whenever the estimates of cost-effectiveness in different chapters use the same input prices, their results are comparable within a given region. Analyses that draw on published

estimates for price or unit cost information are necessarily less comparable across interventions and introduce another element of variation, even in the same locale. Still more variation arises when costs or outcomes are extrapolated from one country or region to another.

Because the CER depends on many parameters and variables, of which only the discount rate and the disability weights are uniform, good analytic practice calls for sensitivity analysis to see how the ACERs and ICERs change with plausible variation in one or more parameters. Many chapters (such as chapter 26) provide such analyses, varying one value at a time, to sketch the likely range of estimates. This method is one way of dealing with uncertainty (which differs from real, known variation) about the true values of the data and seeing whether the ranking of interventions changes when those values change. Such analyses do not indicate the probability that the true CER falls in a particular interval, only under what input values it would do so. Estimating such probabilities requires knowing or assuming the statistical distributions of the parameters in question and using that information to derive confidence intervals around the point estimates. Guides to CEA recommend these approaches (Gold and others 1996), and the National Institute for Clinical Excellence requires probabilistic sensitivity analysis before approving medical treatments in the United Kingdom (NICE 2004).

Data for estimating probability distributions around mean parameter estimates are seldom available in low- and middle-income countries. Simply having available several different estimates of a parameter is inadequate for deriving a distribution, because the differences may be caused by variation in regional costs or expected life years rather than uncertainty. However, assumptions about the shape of distributions can be applied within modeling exercises to give an indication of the likely distribution of ICERs. Only a few chapters, therefore, include confidence intervals. The analyses for tuberculosis (chapter 16) and malaria (chapter 21) do, but ranges associated with most cost-effectiveness estimates (see chapter 2) reflect other causes of variation, not statistical accuracy.

Although calculations are often reported to several significant digits, such precision is not really feasible given the uncertainties in the original data: "economics is a one- or at most a two-digit science" (Morgenstern 1963). However, even crude findings can be valuable, either as guides to value for money if inaccuracies do not affect the relative order of magnitude of the results or for understanding and exploring the sources of variation and their effect on priorities as well as indicating future research needs (Claxton, Sculpher, and Drummond 2002). These issues arise, for example, when considering whether to expand the EPI or to add new antigens (chapter 20), how far to extend screening procedures (chapters 29 and 34), and when to change drugs in response to vector or parasite resistance (see chapters 21 and 23).

The quality and relevance of evidence can vary considerably, depending on whether information comes from randomized controlled trials or systematic overviews, nonrandomized studies with multivariate analyses and well-defined endpoints, or case studies or expert opinion. For these analyses, the quality of evidence also depends on geographic coverage, as distinguished in chapter 2:

- literature review of one cost-effectiveness study, in one country
- literature review of several studies in different countries in different regions
- literature review of several studies in different countries in the same region
- original analyses starting with price and quantity data in one country
- original analyses starting with price and quantity in one or more regions.

The first three categories differ in how representative published findings are; the latter two categories differ according to the data used in constructing total effects and total costs.

Besides the quality of the evidence at its source, how the results will apply to other settings matters, particularly when the data are limited to high-income countries. The more that outcomes depend on underlying biology, the more the findings will apply to low- and middle-income countries. Outcomes depending more on cultural or environmental factors are less readily transferred and require judgment and evidence as to their applicability elsewhere. Sometimes the only detailed studies refer to high-income countries, as for abuse of substances other than alcohol and tobacco (chapter 48). At the other extreme, in a few cases all or nearly all the information comes from low- and middle-income countries, and there is no need to extrapolate, as for nutritional interventions (chapter 28) and community health and nutrition programs (chapter 56).

COST-EFFECTIVENESS AND POPULATION IMPACT

An intervention CER, whether average or incremental, is based on assumptions about introduction, expansion, contraction, or modification of the activity compared with current (or sometimes "best") practice. Comparison of ratios indicates whether one intervention offers better or worse value for money than another at the individual level but says nothing about how either one affects the whole population. The analysis, therefore, includes, wherever possible, two ways of describing the latter effect. One is to consider a population of 1 million, with a typical regional age and sex structure, and to suppose that the intervention were delivered to all the potential beneficiaries. That number of people is just the prevalence or incidence of

the condition times 1 million. The total cost would then be the unit cost times that number (or the cost of reaching that many people if the unit cost varies with coverage). The total health gain would be the individual effectiveness times that same number (or the overall outcome if that depends on externalities, such as the transmission of communicable disease, that are sensitive to coverage). Standardizing on a population of 1 million allows comparisons among regions and interventions in which the incidence or prevalence may vary greatly.

A second approach standardizes not on population but on expenditure: if an additional US$1 million were devoted to the resources needed for an intervention, how many people could benefit from it and how large would the health gain be? The coverage of the intervention would be US$1 million divided by the average cost, and the total gain in DALYs would be that number of people times the average effectiveness. This approach is applied in relatively few chapters because of the information requirements; its advantage is to facilitate judgments as to where increased spending would be most justified—where it would yield the largest improvement in health, reach the most people, or account for the largest share of burden from a condition. Table 1.3 in chapter 1 provides examples for some interventions to reduce child mortality, prevent or treat HIV and AIDS, reduce smoking prevalence, treat heart attack and stroke, detect and treat cervical cancer, and operate a basic surgical ward. The estimates of DALYs gained per US$1 million vary from less than 100 to more than 100,000—a thousandfold difference in value for money. Annex 26.A of chapter 26 provides both kinds of calculations, per million population, to compare the cost-effectiveness of interventions for improved maternal health in South Asia and Sub-Saharan Africa.

IMPROVEMENTS AND FURTHER APPLICATIONS

What would improve the kind of estimates and conclusions reported in this volume? Most crucially, more and better data are needed in low- and middle-income countries to reduce reliance on extrapolation from high-income countries and on expert judgments. The need for information starts, in some cases, with better estimates of incidence and prevalence, but even where the epidemiology is well known, data on coverage and outcomes of existing interventions are scarce. Evidence of what it would cost to change coverage of existing interventions or add new interventions, and with what results, is particularly scarce and depends heavily on assumptions. This situation is sometimes true even for activities that have been conducted widely for many years and have been extensively analyzed, notably the EPI (chapter 20). Analyses should when possible be conducted at the level of a country or even smaller units, to take full account of all the reasons cost-effectiveness varies from place to place and to develop priorities on the basis of analyses appropriate to local circumstances. The methods used here are intended to help guide such efforts, and they can and should be refined through research to provide more robust help to policy.

Finally, a more concerted approach is needed for clarifying the options facing different decision makers and incorporating the results from systematic literature reviews into analytic models that compare the costs and effects of alternative interventions (Buxton and others 1997; Kuntz and Weinstein 2001). Modeling encourages explicit decision making and can deal comprehensively with the inputs and outcomes of decision options, which allows a range of uncertainties to be reflected. Thus, hypotheses about interventions can be formulated and tested statistically. Specifying models explicitly (as in chapter 16, for example) can also help identify gaps in current evidence and can capture details specific to particular populations and settings.

ACKNOWLEDGMENTS

The authors are grateful to Jo Mulligan for the estimation and explanation of input costs used in this volume; to Sonbol A. Shahid-Salles for help in drawing methodological examples from numerous chapters; to George Alleyne, Joel G. Breman, Mariam Claeson, Anthony R. Measham, and Elinor Schwartz for helpful comments; and to Dean T. Jamison and Anne Mills for overall guidance.

REFERENCES

Brooks, R., R. Rabin, and F. de Charro, eds. 2003. *The Measurement and Valuation of Health Status Using the EQ-5D: European Perspective.* (Evidence from the EuroQol BIOMED research program.) Dordrecht, Netherlands: Kluwer Academic Publishers.

Buxton, M. J., M. F. Drummond, B. A. Van Hout, R. L. Prince, T. A. Sheldon, T. Szucs, and M. Vray. 1997. "Modeling in Economic Evaluation: An Unavoidable Fact of Life." *Health Economics* 6 (3): 217–27.

Claxton, K., M. Sculpher, and M. F. Drummond. 2002. "A Rational Framework for Decision Making by the National Institute for Clinical Excellence (NICE)." *Lancet* 360 (9334): 711–15.

Dolan, P., and M. Jones-Lee. 1997. "The Time Trade-Off: A Note on the Effect of Lifetime Reallocation of Consumption and Discounting." *Journal of Health Economics* 16: 731–39.

Drummond, M. F., B. O'Brien, G. L. Stoddart, and G. W. Torrance. 1997. *Methods for the Economic Evaluation of Health Care Programmes.* Oxford, U.K.: Oxford Medical Publications.

Fox-Rushby, J. A. 2002. *Disability-Adjusted Life Years (DALYs) for Decision-Making?* London: Office of Health Economics.

Gabbay, J., and A. le May. 2004. "Evidence Based Guidelines or Collectively Constructed 'Mindlines?' Ethnographic Study of Knowledge Management in Primary Care." *British Medical Journal* 329 (7473): 1013.

Glick, H., D. Polsky, and K. Schulman. 2001. "Trial-Based Economic Evaluations: An Overview of Design and Analysis." In *Economic Evaluation in Health Care: Merging Theory with Practice*, ed. M. Drummond and A. McGuire, 113–40. Oxford, U.K.: Oxford University Press.

Gold, M. R., J. E. Siegel, L. B. Russell, and M. C. Weinstein, eds. 1996. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press.

Hoffmann, C., B. A. Stoykova, J. Nixon, J. M. Glanville, K. Misso, and M. F. Drummond. 2002. "Do Health-Care Decision Makers Find Economic Evaluations Useful? The Findings of Focus Group Research in U.K. Health Authorities." *Value Health* 5 (2): 71–78.

iMTA (Institute for Medical Technology Assessment). 2005. *Newsletter* 3 (1): 1–3.

Jamison, D. T. 2002. "Cost-Effectiveness Analysis: Concepts and Applications." In *Oxford Textbook of Public Health*, 4th ed., ed. R. G. Detels, J. McEwen, R. Beaglehole, and H. Tanaka, 903–19. Oxford, U.K.: Oxford University Press. Also published as Disease Control Priorities Project Reprint 3.

Jamison, D. T., and J. S. Jamison. 2003. "Discounting." Disease Control Priorities Project Working Paper 4, World Bank, Washington, DC.

Jamison, D. T., W. H. Mosley, A. R. Measham, and J. L. Bobadilla, eds. 1993. *Disease Control Priorities in Developing Countries*. New York: Oxford University Press.

Karlsson, G., and M. Johannesson. 1996. "The Decision Rules of Cost-Effectiveness Analysis." *Pharmacoeconomics* 9 (2): 113–20.

Kuntz, K. M., and M. C. Weinstein. 2001. "Modelling in Economic Evaluation." In *Economic Evaluation in Health Care: Merging Theory with Practice,* ed. M. Drummond and A. McGuire, 141–71. Oxford, U.K.: Oxford University Press.

Mathers, C. D., A. Lopez, C. Stein, D. Ma Fat, C. Rao, M. Inoue, and others. 2005. "Deaths and Disease Burden by Cause: Global Burden of Disease Estimates for 2001 by World Bank Country Groups." Disease Control Priorities Project Working Paper 18, World Bank, Washington, DC.

McDaid, D., R. Cookson, and ASTEC Group. 2003. "Evaluating Health Care Interventions in the European Union." *Health Policy* 63 (2): 133–39.

Morgenstern, O. 1963. *On the Accuracy of Economic Observations*. Princeton, N.J.: Princeton University Press.

Mulligan, J., J. A. Fox-Rushby, T. Adam, B. Johns, and A. Mills. 2003. "Unit Costs of Health Care Inputs in Low and Middle Income Regions." Disease Control Priorities Project Working Paper 9, World Bank, Washington, DC.

Murray, C. J. L. 1996. "Rethinking DALYs." In *The Global Burden of Disease*, ed. C. J. L. Murray and A. D. Lopez, 1–98. Cambridge, MA: Harvard University Press.

Murray, C. J. L., and A. D. Lopez, eds. 1996. *The Global Burden of Disease*. Cambridge, MA: Harvard University Press.

Musgrove, P. 1999. "Public Spending on Health Care: How Are Different Criteria Related?" *Health Policy* 47: 207–23.

———. 2000. "A Critical Review of 'A Critical Review': The Methodology of the 1993 World Development Report, *Investing in Health*." *Health Policy and Planning* 15 (1): 110–15.

NICE (National Institute for Clinical Excellence). 2004. *Guide to the Methods of Technology Appraisal*. London: NICE. http://www.nice.org.uk.

Sheldon, T. A., N. Cullum, D. Dawson, A. Lankshear, K. Lowson, I. Watt, and others. 2004. "What's the Evidence That NICE Guidance Has Been Implemented? Results from a National Evaluation Using Time Series Analysis, Audit of Patients' Notes, and Interviews." *British Medical Journal* 329 (7473): 999.

Sloan, F. A. 1995. "Introduction." In *Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Technologies*, ed. F. A. Sloan. New York: Cambridge University Press.

Smith, P. C. 2005. "Statutory Packages of Health Care Alongside Voluntary Insurance: What Treatments Should Be Covered?" York, U.K.: Centre for Health Economics, University of York.

Taylor, R., and NICE (National Institute for Clinical Excellence). 2002. "HTA Rhyme and Reason?" *International Journal of Technology Assessment in Health Care* 18 (2): 166–70.

Walker, D., and J. A. Fox-Rushby. 2000. "Critical Review of Economic Evaluations of Communicable Disease Interventions in Developing Countries." *Health Economics* 9 (8): 681–98.

WHO (World Health Organization). 2000. *World Health Report 2000*. "Health Systems: Improving Performance." Geneva: WHO.

Wordsworth, S., and A. Ludbrook. 2004. "Comparing Costing Results in Across Country Economic Evaluations: The Use of Technology Specific Purchasing Power Parities." *Health Economics* 14 (1): 93–96.

World Bank. 1993. *World Development Report: Investing in Health*. New York: Oxford University Press.

———. 2003. *World Development Indicators*. Washington, DC: World Bank.