
Appendix C

Priority Setting for Health Service Efficiency: The Role of Measurement of Burden of Illness

Gavin Mooney and Andrew Creese

The need to set priorities arises from the fact that not all illness can be eradicated nor all needs met. This failure to be able to meet all needs arises not principally because of the limitations of technology—the technology is currently available to eliminate many of the most important diseases, such as poliomyelitis and measles—but because of the scarcity of resources. Policymakers in the health sector have to manage resources in ways that maximize health outcomes, whether this means redeploying resources, allocating limited new resources, or cutting back on the use of existing resources. They must also get the most out of whatever they have available, which is likely to mean changing the mix of resource allocations.

Priorities are about change. Setting priorities to achieve best possible value for the resources available should be based on considerations of both benefits and costs. Using scarce resources in any way means, by definition, giving up the opportunity to use them in some other way; providing benefits here means forgoing them elsewhere. Priority setting means developing analyses and procedures to ensure that the policies that get priority (that is, those which get a higher call on extra resources) are the ones that provide the greatest benefits per additional dollar spent. If the dollars could have been better spent elsewhere, then they should have been spent elsewhere.

In this appendix we illustrate first how, conceptually, information on the burden of illness can contribute to the process of priority setting. We then identify some of the practical problems entailed in deriving appropriate information on both the costs (briefly) and the outcomes of health interventions. We also consider here the usefulness and limitations of the dollar cost per disability-adjusted life-year (DALY) gained. Finally, using the cost and outcome information summarized for individual chapters, we give examples of how a cost-effectiveness strategy may be used in setting health priorities at sector, project or program, and clinical levels of the health system.

The Burden of Illness and Priority Setting

The notion of illness as a social and economic burden is very old. Quantitative estimates of society's losses from bubonic

plague epidemics and natural disasters were made in the seventeenth century by the English physician William Petty (1699). The epidemiological and economic tally of diseases on a national or global basis has been documented more recently in an empirical work by Walsh and Warren (1979). Accounts of the costs of individual health problems, such as road traffic accidents or, more recently, acquired immunodeficiency syndrome (AIDS), are regularly published in journals concerned with health policy. (See, for example, Henke and Behrens 1986; also see the subsequent debate on the cost of illness: Shiell, Gerard, and Donaldson 1987; Behrens and Henke 1988; and Hodgson 1989.) The motivating factors, sometimes implicit, sometimes explicit, behind such analyses appear to be an assessment or reassessment of priorities. Measuring the burden of illness is thus seen as an ingredient in the rational setting of priorities.

The reason for attempting to measure the burden of illness is thus to allow a better (that is, more efficient) use of scarce resources in reducing the effect of illness on a population, a group of individuals, or even single individuals. In some instances, such as when an epidemic or an important new disease manifests itself—and AIDS is a classic case—awareness of the burden of illness in itself forces a reassessment of expenditure priorities. But even in this instance there is a need to assess the benefits and costs of different policy reactions and also to review the priority status of the new problem with the same criteria used to measure illness problems that have been more long standing.

One cannot examine issues of efficiency, however, without looking at both inputs and outputs—costs and benefits. The burden of illness, however measured, is not a particularly useful concept if it is assessed separately from the question of the policies and resources associated with addressing that burden, that is, questions of how effective and how costly different forms of treatment, care, or prevention are in dealing with the illnesses being considered.

Looking solely at the disease or illness side of the equation, and not simultaneously at the resource or input side, does not permit one to say anything conclusive about the assessment or

evaluation of priorities. Conceptually, it is necessary to give some consideration to the relationship between the burden of illness and the effect that different treatment, care, or prevention interventions have upon it.

The need to consider the disease and resource sides of the efficiency equation together is not, of course, an argument against measuring the burden of illness; it is only an argument against the belief that the epidemiology of illness in itself is a basis for priority setting and against the idea that, in general (there are exceptions), measuring the total burden of an illness is a valuable thing to do. Thus if, for a particular country the burden of morbidity and premature mortality from childhood infections was greater than the burden of adult respiratory disease, this in itself would tell us nothing about the relative resource allocation priority of these two problems. Assuming for the moment that adult and infant lives are weighted of equal importance, considerations of the cost and effectiveness of available technology for altering the course of the disease must still be introduced before the overall cost-effectiveness of interventions can be ranked. Furthermore, in most instances it will not be the total costs and the total benefits (in the context of the burden of illness, the latter will normally be estimated as a reduction in that burden) that are relevant—or indeed the average costs and benefits.

The prime concern is with assessing change. If resource inputs in one program are increased, to what extent is the burden of illness in that program reduced? If inputs are increased again, how much more is the burden of illness reduced? Conversely, if resource allocations to a program are decreased, to what extent will the burden of illness increase? Economic thinking of this sort has clearly established that what is relevant in the setting of priorities are the marginal benefits and the marginal costs and consequently the marginal effect on the burden of illness of an increase or decrease in the resources deployed in that program.

Thus the prime objective of efforts to estimate the burden of illness is best seen in the context of attempting to estimate the reduction in the burden of illness through the application of some treatment or preventive regime which inevitably involves the use of scarce resources. Such efforts are a means toward allowing the quantification of the effectiveness of a particular policy on a particular disease and help in answering the question, For illness x , does treatment A do more good (reduce more the burden of illness x) than treatment B?

Thereafter the issues of operational efficiency can be addressed: how best, with regard to the cost per unit of output, can the burden of particular illness x be reduced? Here the relevant techniques are cost-effectiveness analysis and cost-utility analysis, the latter having the advantage over the former of being able to consider more than one type of output (for example, both mortality and morbidity reductions). This is discussed in more detail later.

A comparison of the burden of illness across different diseases, if such is possible, leads into still more interesting questions of the relative efficiency of using resources to deal with the effects of different illnesses. The clear implication here is

that if there is a need to choose between spending Y on program C and spending the same amount on program D, then for the sake of efficiency (what is called allocative efficiency) the investment should be in the program in which the benefit is larger. The question then is, where can an increase in resources be deployed to decrease the burden of illness to the greatest extent? It should be noted that the question in the other direction is also relevant: where can a cut in resources be made so that the increase in the burden of illness is minimized? Here the relevant techniques are cost-utility and cost-benefit analyses.

The ideal with cost-benefit analysis is to operate with the three key rules for allocative efficiency:

- If for a particular program costs are greater than benefits, then that program should not be implemented.
- If benefits are greater than costs, proceed with the policy. But further, and ideally, these rules should be applied at the margin.
- In other words, a policy should be pursued up to that point where the marginal benefit equals the marginal cost—but not beyond that point.

Such rules are made in recognition of the scarcity of resources (we cannot do everything) and the importance of efficiency (we accept that society should attempt to provide as much benefit as possible with what resources are available; for good reviews of economic appraisal, see Mills 1985 and Drummond and others 1986).

These issues can be summarized in the following five points:

- Measuring the burden of illness is an important ingredient in rational priority setting.
- Rationally set priorities are obtained by a process of weighing costs and benefits, and benefits are obtained largely through a reduction in the burden of illness.
- Priorities are set on the margin: it is the costs and benefits of change that matter. Accepting this leads to such questions as, if resources are increased, where can they be used to reduce the burden of illness most?
- Priorities are not a function of total costs or total benefits, which means consequently that rational priority setting has no interest in the total burden of an illness unless it is practical both technologically and economically to eliminate that illness—and such instances occur very seldom.
- There are two relevant forms of efficiency in priority setting: operational efficiency when the priority questions relate to how; allocative efficiency when they relate to whether and how much.

Measuring the Effectiveness of Interventions

Our main concerns in this appendix pertain to measurements of the burden of illness. Because of the importance attached to the costs of reducing the burden of illness, however, we briefly look first at some issues of cost measurement.

Information on Costs

Several key principles are associated with all costing. First, what we seek to measure are the so-called opportunity costs, that is, the benefits forgone in the best alternative use of resources. Where markets work well, market prices can often be used in estimating costs. In the health care sector and in developing countries generally, however, the frequent market failings or distortions mean that "shadow pricing" is required.

Second, the relevant cost is always the cost of the change being considered, and this can normally be defined as the marginal cost. If, for example, it is expected that a hospital will have to deal with an extra hundred births next year, the relevant cost relates to the extra use of resources for staff, equipment, and other resources for these births. (It should be noted that this cost may have no similarity to the existing average cost per birth in the hospital.)

Third, the cost should normally include all resource use, no matter on whom it falls. Thus it is not just health service or public sector costs that are relevant but also costs falling on private agencies, the patients themselves, their relatives, and so forth.

Fourth, payments for sickness benefits, pensions, and the like are not costs as such but rather transfers from one group in the community (normally the working population) to another (here, ill people and the elderly). These redistributions of resources are not costs from society's point of view. They are called "transfer payments."

At a more practical level, one of the great difficulties in making estimates of costs according to the above principles is the paucity of existing data. What are often available from accounting data are average costs—and yet it is not these that are required. Because they are available, however, there is a great temptation to use them. We would counsel against this and suggest that crude marginal cost estimates are better than precise average costs.

If the use of average costs is to be rejected and marginal costs calculated, how is this best done? The answer is, quite simply, to ask the appropriate people for their estimates. Thus, in extending care to take account of an extra hundred births next year in a particular hospital (as in the example above), the starting point is to ask the hospital manager or obstetrician what facilities and resources will be needed to cope. An estimate can be made of the extra time of doctors, of nurses, of auxiliary staff, of equipment, of food, and the like—and then each of these resources costed. That then gives the relevant cost figure.

The lack of adequate, readily available, marginal cost data in health care (and not just in developing countries) is perhaps just as big a problem as the lack of good outcome measures. It is normally easier, however, to overcome the problems on the cost side and get a sufficiently accurate estimate of the relevant marginal costs.

Information on the Effectiveness of Health Interventions

It is clear that the measurement of the burden of illness is difficult. This is true for three reasons: first, the effects on

health status and illness are multidimensional, involving physical pain, physical impairment, mental disability, mortality, and so on; second, health status is a value-laden concept; and third, the appropriateness of one particular measure is likely to vary, depending on why it is being used. Infant mortality could be a reasonable basis for comparing the effect of child immunization programs across different countries. It would not be a suitable measure for the effectiveness of an antismoking campaign among schoolchildren.

There is also a hierarchy of measurement which has to be noted. If all that is of interest is to answer the question "Is x more effective than y ?" then an "ordinal" ranking is all that is required (that is, we can rank the relevant change in the burden of illness as greater or less). If we want to go further and say that a quantified amount more is obtained, then "cardinal" scaling is necessary.

In most contexts, cardinal scaling is necessary in priority-setting exercises because it is not enough to be able to say that x is more effective than y —especially if x is also more expensive than y ; we need to know how much more, the issue of cardinality.

HEALTH CARE ACTIVITIES. The most basic methods used in measuring health care outputs are activity measures, such as numbers of cases treated, numbers of consultations, and proportion of population vaccinated, which do not directly measure health at all. Of course, it is reasonable to assume that the more patients who are treated the greater will be the benefit with regard to reduction in the burden of illness. But that assumption requires various other assumptions about the effectiveness of intervention, which it would be preferable to measure more directly. For example, to couch the effectiveness of a family planning campaign in terms of the proportion of women reached in the campaign may be a poor measure of the effect it has on family planning *per se* or, more explicitly still, on the number of unwanted pregnancies conceived.

HEALTH INDEXES. The simplest methods which incorporate some assessment of health status involve using estimates of mortality or life-years lost. It is clearly the case, however, that these estimates then ignore morbidity and any other aspects of the burden of illness. Of course, there may be some situations in which it is possible to justify such ignoring—for example, in certain instances in which mortality and morbidity are highly correlated. Generally, however, such measures are of rather limited value. In the field of clinical or individual health status measurement, several different types of index exist. Such indexes are important, because the objects of such measurement are those on which population-based health status measures should be based. For assessments of levels of physical and social functioning of individuals, see, for example, the Duke-UNC Health Profile (Parkerson 1981); the Sickness Impact Profile (Bergner and others 1981); the Index of Well-being (Kaplan and Bush 1982). For a general review see Hall and Masters 1986. For population-based measures, the review method of Walsh and Warren (1979) is worth noting. Still,

the authors made no attempt to aggregate the morbidity and mortality components of health and simply presented ordinal rankings of the main diseases, first by their morbidity and second by their mortality.

More recent attempts to combine both types of information in a single aggregate have counted both avoided disability and avoided mortality as the number of days of a "normal" life gained. This has provided a common yardstick with which morbidity or spells of temporary incapacity can be arithmetically combined.

The Ghana Health Assessment Project Team's calculations of "healthy days of life" are of particular interest in the context of attempting to use burden-of-illness data in the setting of priorities, particularly with respect to the need to be cautious in using such data. In this study, an index was developed for measuring days of healthy life lost to selected diseases which involves the assumption that days spent being dead, being permanently disabled, and being temporarily disabled are equally valued. That seems a difficult assumption with which to agree, but provided the sensitivity (see the section on uncertainty below) of such assumptions is tested, then such apparently gross assumptions may be defensible. The point is that they ought to be tested—by, for example, determining what difference it makes if the weight attached to being disabled is 0.5 compared to a weight for death of 1.

Certainly such a method is valuable, provided its limitations are recognized and provided it is not used to rank priorities in terms of the total burden of illness. Barnum (1987), for example, makes the very relevant point that weightings should be applied to estimates of lost healthy days, first, to reflect the time dimension (that is, the discounting of losses in the future) and, second, to reflect productivity loss. Even here, however, there are problems because the implication of the productivity loss measure is that anyone older than fifty-seven years has a zero value. (Here we have a variant of the human capital method of estimation without any attempt being made to avoid the problems of zero-weighting retired people.)

Barnum states: "The results [provided by this approach] illustrate that weighting and discounting, and their interaction, potentially ... affect the priorities and strategies that evolve from an epidemiological analysis of the health sector" (1987, p. 838). This interpretation, however, gives the impression that the commentator is assuming that the total burden of different illnesses in itself provides some basis for setting priorities. As we have argued above, it does not, except in some very restricted circumstances. (It may be that such total measures will have more relevance in setting research priorities in situations in which the relative size of a problem is the only basis for setting priorities because we know nothing about either the different costs of research in different areas or the different probabilities of success and therefore have to assume that neither varies with the illness. Such assumptions may be an approximation of reality in setting research priorities; that is unlikely to be true, except infrequently, in the case of health care policy priorities.)

A further problem revealed by the method used by the Ghana Health Assessment Project Team is not only that it

deals with totals but also that it deals with averages. We have already considered the need to concentrate on the margin, and in some instances marginal benefits or marginal costs may turn out to be closely approximated by average benefits or average costs, respectively. But they may not.

Thus, and for example, in estimating the value of a death prevented, the relevant formula in the calculations of the Ghana Health Assessment Project Team, when including an allowance for length of survival, considers only the average age at onset and the average age at death. It follows that the value of a death prevented is then always calculated on the basis of the average years of life extended by the program. Yet if there is considerable variability about these averages, then priorities may be wrongly set. Thus, for example, if we consider the priority to be attached to screening women over the age of twenty for breast cancer and use only average figures, the average increase in life expectancy is small. But if we then look at specific age groups or risk groups, the position will be better for some, worse for others. Thus using average figures is likely to lead to a misallocation of scarce resources. We have in this example two lessons to be learned: the danger of using total burden-of-illness data; and the danger of using average burden-of-illness data.

A more recent study, in which program cost information is juxtaposed with an aggregate measure of effectiveness, is that by Prost and Prescott (1984) on onchocerciasis. The authors estimate the cost-effectiveness of prevention measures for onchocerciasis using the alternative measures of effectiveness reproduced in table C-1.

Given the emphasis here on added benefits and on the sensitivity of the results to different measures of effectiveness, this type of empirical work is potentially very useful. As Prost and Prescott themselves state, however, "the relative cost-effectiveness of onchocerciasis control is very sensitive to the choice of effectiveness measure" (1984, p. 801). It is thus clear that there remain problems in improving such measurement of burden of illness to allow relevant measures of effectiveness to be designed.

We accept that this is difficult, but it is what is required for rational priority setting, and no amount of concern about lack of data or about the problems involved in such development will make the basic requirement change. It is, in our judgment, much better to attempt to adopt this methodology in some form or other even if we get no closer than a crude approximation than to adopt what are clearly wrong or inappropriate measures.

A related approach, hitherto used only in industrial countries, entails the adjustment of the quantity of additional days of life by a factor designed to capture (and make comparable) the dimension of quality. In comparing renal dialysis with kidney transplantation, for example, as options for patients with end-stage renal failure, it is clear that a simple comparison of the dollar cost per case would fail to capture the superior quality of outcome of successful transplantation over dialysis. This factor is apparent to all—clinicians, patients and their families, and potential patients, that is, the public. The outcome resulting from transplantation is clearly "better" than

Table C-1. Cost-Effectiveness of Onchocerciasis Control
(U.S. dollars)

Unit of measure	Cost
Per year of healthy life added	20
Per productive year of healthy life added	20
Per disability-adjusted year of healthy life added	150
Per discounted productive year of healthy life added	150

Source: Prost and Prescott 1984.

that from dialysis. Although both options prevent premature death, the difference in the quality of survivors' lives necessitates adjustment to the number of years gained to reflect this, so that the outcomes are comparable.

This approach, it has to be emphasized, tells us only about the relative burden of one disease as compared with others. Its primary use is in attempting to rank for the purposes of priority setting the costs per disability-adjusted life-year (DALY) gained on the margin of different programs. Thus if one program has an extra cost of \$10,000 per DALY gained and another program has an extra cost of \$100,000, it would be rational, if there were no other considerations, to invest in the first program, because the number of DALYs gained would be greater.

The fundamental problem with DALYs as with all such measures of health status is in getting the appropriate weights for mortality and for all the possible forms of morbidity. Questions here relate to whom to ask to do the valuing; how one life is to be compared with another—normally assumed to be the same; how to allow for uncertainty; and many other issues, including whether the only output of health services is improved health status. Because this last point is a concern with all the methods of measuring the burden of illness in this appendix, it will be considered later in a more general context. (For a critique of DALYs see Loomes and McKenzie 1989).

Two contributions to the field of development of DALYs are particularly noteworthy. We will discuss, first, some of the work of the "father of DALYs," George Torrance, from Canada and, second, the work of Alan Williams, from England (see Torrance 1985; Williams 1985).

One of Torrance's key contributions to the field is with respect to methodology and in various papers he has provided much guidance for researchers in how to measure health states in practice. Thus he gives the main steps in developing health status measures:

- Identify the relevant health states for which preferences are required.
- Describe the health states.
- Select the subjects whose preferences will be measured.
- Determine the type of preferences required (ordinal, cardinal).
- Determine the measurement instrument to be used.

Although we cannot discuss all these steps in detail, it is worth noting, regarding the last, that there are various ways of

tackling the question of how to measure burden of illness using health status or DALY measures. All are concerned with attempting to quantify different health states—such as unconsciousness, severe physical impairment, moderate pain—on a scale stretching from perfect health (given a weight, say, of 1) to death (weighted, say, as 0). Thus a year of life with, for example, significant physical impairment and considerable pain might be thought to be only 80 percent as good as a year of perfect life. In such a case, the DALY for this health state would be 0.8.

Various instruments are available to assist in the attempt to measure DALYs. These include the following:

- *The rating scale* normally consists of a line on a page with a scale from, say, 0 at one end to 1 at the other, the end points being defined as death and perfect health, respectively. Other health states are then placed at different points on the line, a point right in the middle being equated with a health state or DALY of 0.5.
- *The standard gamble* involves a choice of the certainty of a health state Y as opposed to the probability of a health state X (where X would normally be preferred to Y). If X were perfect health (weight as 1) and the probability which made the valuer indifferent in this choice were 75 percent, then the DALY for Y would be 0.75.
- The weights of *the time trade-off* are determined by offering choices of different lengths of life in different health states and attempting to get "indifference" across different choices.

The actual use of DALY data linked to costs is provided in table C-2, based on work by Williams (1985). Essentially what this means is that given an additional amount of money, say, £14,000, (approximately \$25,000) to spend on the listed programs, spending it on pacemakers would give twenty DALYs, whereas spending it on hospital hemodialysis would give only one DALY.

It is perhaps superfluous to add that the development of DALYs can be difficult. However it is done, attaching weights to different morbidity states in relation to death, so that, ideally, mortality and all forms of morbidity can be placed on a single index, involves value judgments. It is also the case that we know of no "fully fledged" DALY applications in developing countries.

SOME COMMON PROCEDURAL POINTS. Whatever methodology is adopted for assessing the burden of illness, there are five issues that need to be handled with care.

- Determination of the purpose
- Discounting over time
- Other outputs
- Uncertainty
- Equity

Although determining the purpose may seem an obvious point, it is worth stressing that how the burden of illness is best measured or valued is a function of why it is being measured or

Table C-2. Costs and Consequences of Selected Medical Procedures
(pounds sterling)

Procedure	Present value of extra cost per DALY gained
Pacemaker implantation for heart block	700
Hip replacement	750
Valve replacements for aortic stenosis	950
CABG for severe angina with left main disease	1,040
CABG for moderate angina with three-vessel disease	2,400
Kidney transplantation (cadaver)	3,000
Heart transplantation	5,000
Home hemodialysis	11,000
CABG for mild angina with two-vessel disease	12,600
Hospital hemodialysis	14,000

Note: CABG (Coronary artery bypass grafting).

Source: Drummond 1987; Williams 1985.

valued and in what circumstances. Calculations concerning the burden of illness will almost always be used as an estimate of some output measures, and output measures have to be or ought to be related to the purpose or objective of the exercise. It is also clear that if the wrong measure is used it is quite likely that a distorted answer will be obtained. (For example, if breast cancer treatment programs are related solely to percentage of survival over, say, five years, then all aspects of quality of life—pain, dignity, losing a breast, and so on—will be ignored and given a zero weight. Yet it seems clear that women suffering from breast cancer will value more than just survival.)

For calculations of the burden of illness (and also the resource costs of interventions), the value attached at different points in time is not constant. As Barnum states: "Neither the individual nor the community is indifferent as to when the effects of disease occur.... A healthy day of life in the present has a greater intrinsic value to the individual than a day in the future (1987, p. 834)." The way to handle this phenomenon is through "discounting" future benefits and costs at some positive rate of discount; such discounting results in a weighting over time which gives more weight to current effects, less to those of the near future, and still less to those in the distant future. This means, for example, that preventive programs may seem to do rather badly as a result of discounting. This is because they often involve costs now (which are therefore not discounted) and benefits in the future (which are discounted). What rate of discounting to use is problematical, and it is normal procedure to use a range of rates, usually between about 3 and 10 percent.

Although it can generally be agreed that the decrease of the burden of disease on the sufferer is the prime output of any health care system, other outputs are present and relevant for the setting of priorities. For example, if infectious diseases are cured in some people, others who would otherwise have become infected will benefit. Again, nonsufferers may benefit knowing that others' suffering is reduced—what Culyer (1976)

has called the caring "externality." Information is also an output. For example, informing patients about their state of health even if it is not changed or indeed cannot be changed may provide benefit to the patient. Being able to pass difficult decisions to the doctor may also sometimes be of benefit to some patients.

It is difficult to say what weight will be attached to these other outputs. It is clear, however, that their importance is likely to vary both across different diseases and across different patients. Thus, although it is appropriate in assessing priorities to concentrate on the output side on reductions in the burden of disease, these other forms of output may sometimes alter the priority rankings or weightings.

Benefits in the future may be uncertain, and in such cases an adjustment should be made to reflect their expected value. For example, the reduction of infant mortality may lead to greater benefits as a result of a health education campaign concerned with hygiene to reduce childhood diarrhea. Often, however, it will not be possible to state precisely what all the potential effects of an intervention for the treatment or prevention of some disease will be.

Where uncertainty exists, sensitivity analysis should be used in handling it; that is, a range of values should be put in for a particular parameter to see what the effect of the different values is—how sensitive the result is to the change in values. Where the result does change, it may be necessary to devote some effort to trying to reduce that particular uncertainty.

Although we accept that equity is an important goal in most health care systems, we are focusing in this appendix on efficiency. Still, it is important to recognize that equity and efficiency goals can sometimes conflict. Such a conflict may mean that minimizing the burden of illness is not the goal or at least that such a goal is constrained by concerns for equity. For example, although it may in some instances be efficient to concentrate highly specialized facilities in the cities, this is unlikely to provide an equitable system with regard to geographical access.

It is also the case that if equity is concerned with access or use rather than with health per se, then factors other than purely burden of disease have to be taken into account. In other words, if a society values the fact that individuals have equal access to health care irrespective of whether they then use it to obtain effective care, then such a set of values cannot be directly contained within burden-of-illness calculations.

Certainly in many—but admittedly not all—equity measures there will be some need to assess the relative burden of disease across different groups in society. Such cases present the few occasions in which the burden of disease itself, as opposed to its reduction, is the relevant policy measure with which to operate. Whether that is the relevant measure of equity to use is something that cannot be resolved in this appendix (but see for more discussion Mooney 1992). Other factors such as access may become relevant.

METHODS FOR PUTTING MONETARY VALUES ON OUTCOMES. We have seen above that health status measures are more widely

usable if they aggregate the relevant components (mortality and morbidity) in a single numeraire, such as healthy days or DALYs. In an analogous fashion, the usefulness of health outcome data for priority setting is substantially increased if an acceptable monetary yardstick can be found, to allow direct comparisons between the value of inputs used in improving health and the value of these improvements. When this is possible, not only can cardinal comparisons be made between competing claims on resources, but the more fundamental cost-benefit questions can be asked and answered.

There are three principal methods for putting monetary values on health outcomes:

- The “human capital” method
- The “willingness-to-pay” (for risk reduction) method
- The “implied values” method

The oldest and simplest of these methods in practice is the human capital one. In this method it is assumed that the objective function that we are trying to maximize through improved health is gross national income in that the measure of value is an individual’s output, normally assumed to be equal to the gross labor costs of employment or in some instances simply the earnings of the individual. Thus if a person is unable to work because of illness, we would, using this method, estimate the burden of that illness as being the work output lost, which is equated with the gross labor costs of employing the individual over the relevant time period. If a person dies as a result of illness, the burden is equated with the present value of the gross costs of employment over what would otherwise have been his or her expected working life span.

There are some clear problems with this method. Unless adjustments are made, it means that no weight is attached to retired people, housewives, children (as children), and others not gainfully employed. Also it will give different values to high earners and low earners, which may well be deemed an inequitable basis on which to set health priorities. Further, gross labor costs are at best an approximation of the value of an employee’s output. It is also assumed that there is no value to health beyond the capacity it provides to produce output relevant to the gross national product, a somewhat restricted view of the goal of health services.

The willingness-to-pay method, most often applied to the saving of life or, more precisely, the reduction in risk of death, adopts a different value stance. Here the nature of the social welfare function—that is, what it is that is to be maximized from a societal perspective—is based on individuals’ values with respect to their willingness to pay for reductions in risk of death (or injury or illness). Thus it is assumed that it is legitimate to ask potential victims or potential sufferers how much they are prepared to pay for a reduction in, say, the risk of death from perhaps 3 in 10,000 to 2 in 10,000. If the response on average to such a question were \$5, then the value of a “statistical” life would be \$50,000 (that is, $10,000 \times \$5$).

This strategy has some advantage in that the question is put to the potential victim, whose values, it can well be argued, are

the ones that should be allowed to count. Also, the question posed as a probability does seem appropriate. (For example, to ask an individual what he or she is prepared to pay to avoid certain death is almost certainly an unanswerable question.) Whether it is possible to obtain valid answers to such questions, however, remains unclear. It is possible to study the behavior of individuals in risk situations and elicit their implied values (for example, in their willingness to pay for safety devices on their cars), but many of these situations are so far removed from the sorts of choices relevant to health care valuations that the values emerging may not be very useful. Additionally the studies that have been conducted in which this strategy was used yield a very wide range of values—but ones which are normally much higher than those based on the human capital method.

Despite the practical problems of the willingness-to-pay method, it has considerable theoretical advantages in that the valuation basis of individuals’ willingness to pay for reductions in the risk of illness and death seems more defensible than that in the human capital method. Of course, if there are equity objections to the method on grounds that priorities in health care should not be based on individuals’ ability to pay (on which willingness to pay is inevitably based), then its application has to be handled with care. The use of the method to date has been very restricted and has related more to willingness to pay to reduce the risk of dying than to reduce the risk of having a nonfatal illness and injury.

The third method of evaluating the burden of illness, the implied values method, is somewhat similar to that of willingness of individuals to pay to reduce risk, except that now it is a question of determining what the implied willingness of health care and other health inducing organizations is to pay for various health outputs or reductions in the burden of illness. The basis of the method is simple: if a decision is made, at the margin, to spend \$1 million to save a life, then by implication the value of that life must be at least \$1 million, otherwise the investment would not be made. If a decision is made not to spend \$2 million to save a life, the value of the life is then by implication less than \$2 million.

In this process of estimating the implied values of life, ideally one would wish that for similar outputs the willingness of the health care system to pay at the margin of each program would be the same (the condition for an efficient solution). What limited information exists, however, suggests that there is a very wide range of values for like outputs. That does not mean that attempts to make the values explicit should not be pursued. The point is that the aim might first have to be to sort out the inefficiencies implicit in the fact that there is a range of values rather than in the short run to use the values per se in the assessment of the burden of illness. Even then, however, the use of a mean value in the short run would be a possible strategy.

One of the clear advantages of this implied values method is that it does not involve any change in the value system, because the implied values would simply reflect those of the existing system. The method is also relatively easy to apply.

It must be obvious from what has been said that none of the methods outlined is ideal in both principle and practice. The human capital method is simple but tends to treat people like machines, where their only value is as workers. It may be argued, however, that estimates made on this basis can provide at least minimum values of life and sickness avoided. The willingness-to-pay method is, theoretically, to be preferred but has not yet been widely applied even within the mortality field, where it is most frequently found. It also requires substantial investment in data. The implied values method at least provides a basis for improving technical efficiency and is relatively simple to apply. (For a fuller discussion of valuing life, see Mishan 1981 and Linnerooth 1982).

Cost-Effectiveness Comparisons for Priority Setting

At the present stage of development of methods of priority setting it is suggested that a simple, sensible way to proceed is to identify the marginal costs of similar outputs across different programs and adjust the allocation of funds to try to get such marginal costs closer to equality. In other words, if (a) some form of DALY measure can be devised and (b) the cost per DALY gained can be identified on the margin of each existing program, we can then attempt to reallocate resources from programs in which the marginal cost per DALY gained is high to those in which it is low.

The reviews of contemporary empirical experience of cost and outcome relationships contained in chapters of this collection constitutes an important piece of stock taking. Epidemiologic, technologic, and economic characteristics of the main diseases and the principal current interventions are presented in a broadly similar format, which allows estimated average costs per average number of days of healthy life gained to be compared (see figures 1-7 and 1-9).

What conclusions is it possible to draw from these reviews? Of equal importance, what conclusions is it not possible to derive from these data? In the first place, the very existence of such a quantity of information on such a range of interventions is clearly to be welcomed. Too many studies have argued for greater priority in funding for one specific disease or intervention, in the absence of any explicit comparisons. Such studies are the antithesis of an economic way of dealing with the situation, in which the necessity for making trade-offs between activities, in the face of overall resource limitations, is taken as a starting point. A galaxy of alternative patterns of resource use exists, even in the poorest country—in the target groups (for example, adults or children), in the intervention strategy (preventive or case management), and in the type of disease or health problem.

The scope of these reviews, however, is still very modest when compared with the huge quantity of health-related actions coexisting in any country at a given moment, or even in a single small general hospital. The range of available health interventions, differing in input mix, location of treatment, type of patient, type of illness, timing of intervention (primary or secondary preventive, curative, or caring), is so large as to

encourage classification, rather than enumeration. The interventions for which cost per DALY gained have been compared are a tiny and nonrepresentative fraction of those available. Indeed, their best use lies more in the illustration of the method of cost-effectiveness in priority setting than for any realistic debate on priorities at a global level. For a full review of priorities, more information is needed and on a more local basis.

Although numerically insignificant, however, the interventions evaluated in the chapters of this collection do have an epidemiologic significance beyond their mere number. They include interventions of known effectiveness against some of the main sources of mortality. Many of these interventions might thus be expected to be prominent among health priorities even if the total number of cost- and outcome-documented additional health interventions were dramatically expanded.

Even for those interventions which are considered, there remains some unevenness in the relevant types of cost and health outcome data presented. In two particular areas this shortage of information may be a critical limitation. First, the sensitivity of the estimated costs per DALY gained is not, in all cases, subject to appraisal. Point estimates, or even “greater than” estimates, are of limited value when there are important margins of uncertainty surrounding them. As indicated above, sensitivity analysis is important in narrowing down the areas in which further information is required and in avoiding overdogmatic priority ranking where the state of available knowledge should indicate caution. Second, the data presented are, in all cases, estimates of the average cost per DALY gained. As emphasized above, such information may lead to inappropriate resource allocation decisions. If studies are conducted to establish the relationship between average costs and marginal costs, then no problem arises. But there are few, if any, such studies for the interventions reviewed in this compilation. By comparing costs and output for health interventions operating at differing scale, we can identify the effects of output variation on total and marginal costs. Once again, in too little of the available empirical work have output variations in relation to costs been assessed.

So, in the absence of empirical information about the relationship between average and marginal cost, what analytical use can be made of the available data? One route to follow is to proceed on the assumption that marginal costs are close to average costs. This is a very special and potentially dangerous assumption. The most casual observation of health care facilities in developing countries suggests that chronic overuse (for example, multiple occupants of hospital beds, “floor patients,” long lines at hospital clinics) coexists with equally chronic underuse (for example, less than twenty consultations per month at a health post with a staff of four health workers, and infant immunization rates of under 20 percent). A bold simplifying assumption of equality between marginal and average cost thus seems more sanguine than intuitive.

Without the benefit of either a simplifying assumption or some empirical basis to speculate about the relationship, at current output levels, between marginal and average cost,

restraint should be used in applying such data to a review of priorities. This is a disappointing conclusion. If marginal and average costs per DALY were roughly equivalent, if the data incorporated allowances for uncertainty, and if these interventions were taken as in some sense representative of technological options in health care, then the data in figures 1-7 and 1-9 could be interpreted as revealing the following:

- Globally, interventions aimed at children should receive higher priority, whether for case management or prevention, than those aimed at adults.
- Although the average cost per DALY for preventive interventions targeted on children is approximately half of that of case management interventions, the ten most cost-effective activities (at \$20 per DALY or less) are a mixture of both preventive and curative actions.
- For adults the overall mean cost for preventive interventions is still lower than for case management, although the differences are now much less.
- The ten most cost-effective interventions for adults entail a mix of preventive and curative actions.
- Some service set providing integrated cure and prevention, rather than discrete vertical programs, would appear to be the most appropriate delivery mechanism.
- The optimal mix of interventions will change as demographic and epidemiologic profiles differ or shift, and thus it needs to be kept under continuous review.

These tempting conclusions are not strictly possible. The data fitted into the cost-effectiveness apparatus are simply not good enough—in quantity and in quality—to warrant such conclusions. This does not mean that the exercise is worthless. If we can reach such conclusions, for a project, country, or region, they are clearly of consequence. That we cannot yet do this—although we may be close—gives urgency to the need to accelerate and improve the collection of relevant data.

Additionally the emphasis on looking at marginal change will normally mean that collecting even crude data on marginal costs at a local level will be better than adopting national or international average cost data. The message is clear. Better to have approximate estimates of local marginal costs than precise, more generalized, average costs.

Concluding Comments

Priority setting is about choice. It is about arranging things in such a way that those policies and programs that are considered most worthwhile stand a better chance of being implemented than others that are considered less worthwhile. In other words, not all needs can be met because resources are scarce. Disease cannot be eliminated; it can only be reduced. So priority should be given to those areas in which the burden of illness can be reduced most per dollar spent. Indeed, we should continue to set priorities according to incremental or decremental changes until it is agreed that no further movement can reduce the burden of disease even more. Clearly, if there are

more or fewer overall resources available, that changes the position—but not the principle. Again if other outputs are deemed relevant (for example, reassurance or information), as we believe they should be, then benefits other than reduced burden of disease must be taken into account.

The link between priority setting and efficiency is crucial in the context not only of the burden of disease per se but of the debate about priority setting more generally. Let us restate clearly what our views are on this matter:

- The need to set priorities arises from the fact that not all illness can be eradicated nor all needs met; this is not just a statement about technology but about the scarcity of resources.
- Priorities are about change. Decisionmakers and policymakers have to try to redeploy resources, allocate some new (but limited) resources, and cut back on the use of existing resources in such a way as to get the most out of whatever resources they have. That means changing deployment.
- Priorities should be based on both benefits and costs. Using scarce resources in one way means, by definition, giving up the opportunity to use them in some other way; providing benefits here means forgoing them there. Priority setting means trying to ensure that those policies that get priority (that is, what gets a higher call on resources) are those providing greatest benefits per dollar spent. If the dollars could have been better spent elsewhere, then they should have been spent elsewhere.

These three statements are central to priority setting. They are very neatly summed up by Shiell, Gerard, and Donaldson (1987) in their critique of studies on the cost of illness: “the total ‘costs of illness’ can only indicate the benefits of treatment options if an intervention is capable of totally eradicating or entirely preventing the disease in question. This is only likely to be possible in the case of a very few infectious diseases. The most pertinent questions facing policymakers usually relate to scale; that is, by how much should an existing program be expanded or contracted. The answer to this question requires a marginal analysis which compares the expected change in benefits with the costs of the intervention which brings that change about.”

From this appendix a number of important conclusions emerge on priority setting in the context of the burden of disease. First, the emphasis of efforts on priority setting ought to be firmly “on the margin”: what can be bought with a few dollars more? what shifting of resources from one program to another on the margin can provide the maximum reduction possible of the existing burden of disease? if cuts have to be made, where should this happen to minimize any increase in the burden of disease? Second, developing some common measure of marginal changes in the burden of disease across different diseases is the key to progress in this area. Third, efforts to measure the total burden of any disease ought to be resisted because total burden is not the basis for setting priorities. Fourth, averages are likewise to be resisted except where it can be shown that

they are reasonable approximations for marginals. Fifth, care must be exercised to ensure that all relevant factors are accounted for—other nonhealth outputs, equity considerations, uncertainty, discounting. Finally, whatever measures are adopted, sensitivity analysis should be applied to determine how robust the results are to different assumptions.

Notes

Comments from Howard Barnum (World Bank), David Evans (World Health Organization), Karen Gerard (University of Sydney), Richard Morrow (World Health Organization), David Parker (United Nations Children's Fund), Gerald Rosenthal (REACH project), and Carl Stevens (Reed College, Portland, Oregon) were particularly influential in shaping our thinking about this appendix. Remaining errors of fact and interpretation are our own. We are also indebted to Anne Hastrup for her secretarial assistance in preparing and revising the manuscript.

References

- Barnum, Howard. 1987. "Evaluating Healthy Days of Life Gained from Health Projects." *Social Science and Medicine* 24(10):833–41.
- Behrens, C., and K.-D. Henke. 1988. "Cost of Illness Studies: No Aid to Decision Making? Reply to Shiell et al." *Health Policy* 10:137–41.
- Bergner, Marilyn, R. A. Babbitt, W. B. Carter, and B. S. Gilson. 1981. "The Sickness Impact Profile: Development and Final Revision." *Medical Care* 19:787–805.
- Culyer, A. J. 1976. *Need and the National Health Service*. London: Martin Robertson.
- Drummond, M. F. 1987. "Economic Evaluation and the Rational Diffusion and Use of Health Technology." *Health Policy* 7:309–24.
- Drummond, M. F., A. Ludbrook, K. Lowson, and A. Steele. 1986. *Studies in Economic Appraisal in Health Care*. Oxford: Oxford Medical Publications.
- Ghana Health Assessment Project Team. 1981. "A Quantitative Method of Assessing the Health Impact of Different Diseases in Less Developed Countries." *International Journal of Epidemiology* 10:73–80.
- Hall, J., and G. Masters. 1986. "Measuring Outcomes of Health Services: A Review of Some Available Measures." *Community Health Studies* 10(2):147–55.
- Henke, K.-D., and C. S. Behrens. 1986. "The Economic Cost of Illness in the Federal Republic of Germany in the Year 1980." *Health Policy* 6:119–43.
- Hodgson, T. A. 1989. "Cost of Illness Studies: No Aid to Decision Making? Comments on the Second Opinion by Shiell et al." *Health Policy* 11:57–60.
- Kaplan, R. M., and J. W. Bush. 1982. "Health Related Quality of Life Measurement for Evaluation Research and Policy Analysis." *Health Psychology* 1:61–68.
- Linnerooth, Joanne. 1982. "Murdering Statistical Lives . . ." In M. W. Jones-Lee, ed., *The Value of Life and Safety*. Amsterdam: North Holland.
- Loomes, G., and L. McKenzie. 1989. "The Use of QALYS in Health Care Decision-Making." *Social Science and Medicine* 28(4):299–308.
- Mills, Anne. 1985. "Survey and Examples of Economic Evaluation of Health Programs in Developing Countries." *World Health Statistical Quarterly* 38(4):402–31.
- Mishan, E. J. 1981. *Cost Benefit Analysis*. London: George Allen and Unwin.
- Mooney, G. H. 1992. *Economics, Medicine, and Health Care*. Brighton, Engl.: Wheatsheaf.
- Parkerson, G. R., S. H. Gehlback, E. H. Wagner, S. A. James, N. E. Clapp, and L. H. Muhlbaier. 1981. "The Duke-UNC Health Profile: An Adult Health Status Instrument for Primary Care." *Medical Care* 19:806–28.
- Petty, Sir W. 1699. *Political Arithmetic or a Discourse Concerning the Extent and Value of Lands, People, Buildings, etc.* London: Robert Clavel.
- Prost, A., and N. Prescott. 1984. "Cost-Effectiveness of Blindness Prevention by the Onchocerciasis Control Program in Upper Volta." *Bulletin of the World Health Organization* 62:795–802.
- Shiell, A., K. Gerard, and C. Donaldson. 1987. "Cost of Illness Studies: An Aid to Decision-Making?" *Health Policy* 8:317–23.
- Torrance, G. W. 1985. "Measurement of Health State Utilities for Economic Appraisal—A Review." *Journal of Health Economics* 5:1–30.
- Walsh, J., and K. Warren. 1979. "An Interim Strategy for Disease Control in Developing Countries." *New England Journal of Medicine* 301:967–73.
- Williams, Alan. 1985. "Economics of Coronary Artery Bypass Grafting." *British Medical Journal* 291:326–29.

Source: Dean T. Jamison, W. Henry Mosley, Anthony R. Measham, and Jose Luis Bobadilla (eds.). *Disease Control Priorities in Developing Countries*. New York: Oxford University Press for the World Bank. 1993.