Large organizations responsible for health care face formidable problems in gathering and deploying data relevant to their principal tasks, which are to monitor the health of the communities they serve, and to decide where resources could most effectively be used to improve things. In the short run, health improvements come by choosing and supporting the right health care activities. In the long run, they also come by choosing and supporting the right research activities.

The ambitions of the Global Burden of Disease Study [1] are to contribute to all three of these important objectives. Broadly speaking, the argument is that if we knew the impact of each disease or injury upon people’s life expectancy and upon the (health-related) quality of their lives, and if we knew the incidence and prevalence of each disease or injury, we could use this information to monitor population health, to establish priorities between interventions, and to guide research priorities.

In what follows, I am going to consider these general claims from three different viewpoints. In the first phase I will abstract from practical difficulties and ask whether, in principle, calculating the global burden of disease is the best way to approach each of these problems. I shall conclude that it is not. This will lead into the second stage, in which I shall ask whether there is anything useful that could be extracted from the global burden of disease calculations as they are actually performed, that would help with the three major problems that have to solved. I shall conclude that there is a little, but that it is not worth the cost, and that resources could, and should, be better targeted on discovering those things that we really need to know. This will lead into my third and final section, where I aim to be rather more constructive, by offering an alternative strategy that would achieve the above objectives more straightforwardly.

Although I shall direct my remarks specifically to the Global Burden of Disease Study (henceforth GBD) as conducted under the impressive leadership of Murray and Lopez, much of what I have to say applies in principle to any burden of disease (or cost of illness) approach, applied at any level (global, national, regional or local). But in the second section I will be considering some rather idiosyncratic elements in the Murray–Lopez approach that other calculators of the burden of disease have omitted (for reasons good or bad). And because I regard one particular element in the GBD enterprise, the disability adjusted life year (DALY), as a particular variant of the more general concept of the quality adjusted life year (QALY), my concern here is not with the use of that general concept (which I support wholeheartedly) but with the specific features of the DALY that make it less useful than other types of QALY that might have been used [2].
IS IT IN PRINCIPLE THE RIGHT APPROACH?

First of all it has to be said that combining life expectancy and health-related quality of life into a single measure as the operational definition of health is undoubtedly the best way forward for the purposes in hand. The use of life expectancy alone is a poor second best, and no other morbidity measure is sufficiently generic to enable the broad comparisons to be made that are required in this field. So the use of a QALY-type concept, as the unit with which to measure both levels of health and the benefits of health care, is in principle the right one.

But how should it be used in tackling our three fundamental policy issues? In monitoring the performance of health care systems it is suggested that the GBD be measured periodically for the community in question, to see whether things are getting better or worse. The well-known scientific difficulty here is that of attributing any overall changes in health to any specific causes. It is well-known that a major predictor of the health of a community (and especially among poor communities) is its level of material well-being [3]. This also appears to be true of individuals within a community [4]. So changes in material well-being are likely to obscure anything generated by the health care system. If changes due to the health care system itself are the focus of interest, then it might well be a better strategy to concentrate on measuring more directly the impact on people’s health of those changes. To do this we do not need to know the GBD, and a great deal of effort could then be redirected from measuring broad aggregates to measuring selective impacts. It suggests that, even at the monitoring stage, concentrating on ‘diseases’ as the target of policy interest is mistaken. What we need to measure is what impact different interventions will have, not what impact different diseases have.

This is even more clearly the case, as Murray and Lopez concede [5], when it comes to priority setting across health care activities. Here we need to know the costs of each activity and its effects on people’s health. We do not need to know the GBD, but the marginal impact of a health technology upon it [6]. Priority setting is to be driven by a comparison of incremental gains with incremental costs. That is where information-gathering resources should be directed, because that is the data we so desperately lack. Health technology assessment is a more urgent, more focused, and more useful field of endeavour for priority setting purposes than calculating the GBD.

Is there however a better case to be made for using GBD calculations as a guide to priority setting in research? The argument might be that research is typically directed at diseases, and so research should be concentrated where the disease burden is greatest. I can see some limited validity in this argument, in that we might not want to devote vast amounts of research money searching for a way to eliminate a particular disease if that disease were an insignificant contributor to the overall burden of disease. But before pouring money into diseases that do impose significant burdens, we also need to know what the research would cost, what specific contribution that research would make to reducing the overall burden if it were successful, and what are the chances of it being successful. Again, it is incremental gains and incremental costs associated with particular research activities that we need, not broad aggregates focused on diseases that may or may not be tractable subjects for research. Very little research aims at eradicating diseases completely. It is much more common to find research directed at relieving particular symptoms, or arresting the progression of a condition. Very few interventions are free of side-effects, and the tracing of these has to be built into any research protocol. They are best captured directly as negative impacts on the length or quality of life of the study population, rather than as increased risks of other diseases, which then have to be translated into such impacts by more indirect methods outside the study in question. So there seems no role for GBD calculations that could not be better fulfilled in other ways.

So my conclusion is that we do not need measures of the Global Burden of Disease for ANY of the broad purposes stated, and that the resources devoted to calculating it should, in the interests of global health, be redirected into measuring the cost effectiveness of particular activities (on as wide a scale as can be afforded).
WHAT PRACTICAL ASPECTS OF THE MURRAY–LOPEZ CALCULATIONS WOULD BE USEFUL FOR THIS INCREMENTAL APPROACH?

Before addressing this question, it is necessary to sketch out the main steps in the Murray–Lopez approach to calculating the GBD. The four key elements are:

1. Calculate the years of life lost due to each disease
2. Calculate the loss of quality of life of those living with the disease
3. Apply a set of weights to reflect the social value of people at different ages
4. Apply a discount rate to reflect the fact that benefits delivered sooner are better than benefits delivered later

I will consider each in turn, but in reverse order.

The use of discounting is much misunderstood, since for many people it is associated with the rate of return on capital, and seen as an economic variable dependent on the productivity of capital in the private sector of the economy, and designed to ensure that the public sector does not invest in activities that have lower returns than are obtainable in the private sector. In my view, this is a mistaken interpretation of what the discount rate is there for. In the public sector it is an important policy statement, indicating the willingness of a community to sacrifice benefits now for the sake of bigger benefits in the future. This is what is called the ‘time preference’ interpretation of discounting, as opposed to the ‘opportunity cost of capital’ interpretation that I mentioned earlier. A low discount rate, such as the 3% used in GBD calculations (which is also the rate recommended by the Washington Panel on Cost-Effectiveness in Health and Health Care [7]) means that the present generation does not need much inducement to make sacrifices now for future gains, whereas a high discount rate (say 10%) implies a very strong sense of urgency about getting the benefits from health care. In general, one would expect rich countries to have low rates of time preference, and poor countries to have high ones (because if the current generations do not survive, there will be no future ones!). Note that these discount rates are ‘real’ ones, from which price inflation has been purged. There is a dilemma here for Murray and Lopez, because they could either have applied different discount rates to different countries depending on local conditions, or, mimicking the situation of central decision-makers (such as the World Bank), apply the same discount rate to everyone, in the interests of equity and comparability. They chose the latter course, though it would be perfectly feasible to redo all of their calculations using different discount rates for different countries. Comparability over time within a country would be maintained, but not comparability between countries. Despite these interpretative difficulties, which are insurmountable, I see no problems in carrying over this element in their calculations to an approach based on estimating incremental benefits rather than total burdens.

The age weights that they have applied in order to reflect the differential social value of people at each year of age are reproduced in Figure 1. They imply that it is not of much social value to prevent a newborn child from dying unless that child is going to live into adulthood. They also imply that prolonging the life of an old adult for 10 years is not as valuable as prolonging the life of a young adult for 10 years. Unpleasant though these implications are, they are probably true. One can, of course, challenge the precise formula, which is probably inappropriate in rich countries with rather low birth rates and long life expectancies, where people remain fit and active well into their seventies, but here we face the same dilemma as with the discount rate. Should it be tailored to individual circumstances, or applied uniformly to maintain comparability? But there is an additional element here that was not present with the dis-
count rate element, and that is whether age-weights should be an integral part of the calculation of the global burden of disease, or whether they should be kept distinct from it, and treated as an ‘optional extra’. I think it would be better to keep the measurement of the GBD as easy to understand as possible, so in the main calculations it would be better to take the age-weights as being 1 at all ages. If someone then wants to use different age-weights, they may do so, and they can specify their own. Comparability is then maintained in the ‘base case’, and the figures are not contaminated by problematic considerations that are difficult to disentangle.

Calculating the loss of quality of life experienced by those currently suffering with each disease is a much more complex matter. The need here was for a measurement scale running from 0 to 1 that would reflect the severity of this part of the burden of disease. In measuring the GBD this is measured as departures from good health, so that 0 represents good health and 1 represents a burden as bad as being dead. In QALY calculations, which focus on the benefits of health care rather than the burdens of disease, the scale runs the opposite way, with 1 being good health and 0 equivalent to being dead. QALY calculations also permit the existence of states that are regarded as worse than being dead (i.e. which have negative values). But these differences are not matters of principle, but matters of appropriateness in each context.

But because of the focus on diseases as the units of analysis, this led to a rather strange method being adopted to derive health-related quality of life (HRQOL) valuations. First of all, some 22 ‘indicator conditions’ were assessed for their impact on HRQOL by panels of experts. This was done by asking the experts to compare how many people would need to be cured of that condition compared with some stated number of people in some other situation, for the outcome to be equal in terms of HRQOL gained. The scores derived in this way from the experts were then averaged, and the indicator conditions ranked accordingly in order of seriousness. The 22 conditions were then grouped roughly into triplets, each triplet having a defined range of average scores. Seven such ranges were identified, labelled disability level 1–7, and the midpoint of each range was taken as a point score. The resulting 7-point category scale was then used by experts to assess the impact on HRQOL of hundreds of other conditions, both treated and untreated. The mean values of these assessments were then taken to be the DALY burden associated with each disease. The judgmental burden placed on these experts must have been considerable, since they needed to know both the lifetime sequelae of each condition in a descriptive sense, and the valuations that people would attach to such time profiles of disability, etc. They would also need to know all this for the condition when treated, so they must also assess the effectiveness of treatments, or at least of the typical treatments. It is a very complicated and roundabout route, fraught with problems. But fortunately there are much more straightforward ways of getting into this territory once the need to calculate the GBD has been abandoned, which I will return to in the next section of this paper.

The final element in the calculation is the estimation of the years of life lost due to each disease. This was done by comparing the age at death of people with the disease with the standard life expectancy of someone of that age and sex. The natural thing to do here would be to take as the ‘standard’ the actual life expectancy of the people in question, but in calculating the GBD this is not what was done. Instead, the standard was taken to be the life expectancy of the longest living major group of people for which we have good data, which was taken to be Japanese women, with a life expectancy at birth of 82.5 years. A model life table with that life expectancy at birth for females was then used (and a slightly worse one was used for males, whose standard life expectancy at birth was taken to be 80 years) as the standard comparator for calculating life years lost through each disease. The argument justifying this procedure was essentially an equity one, because Murray and Lopez did not want the calculations to show that preventing a 40-year-old woman in a rich country from dying (whose life expectancy would still be large at that age) would be more beneficial (in terms of reducing the GBD) than preventing a 40-year-old woman in a poor country from dying (where her remaining life expectancy at that age would have been small). So here we have a strong political judgement entering into the calculations, based on a postulated aversion to inter-country inequalities in life expectancy at each age.

At a moral level this may well be commendable, but it raises two issues of principle. The first is
GUEST EDITORIAL

Figure 2. Implied weights for $e(0) = 50$ for different regional patterns

whether it is right to incorporate this ‘fiction’ into what most people regard as a ‘factual’ matter, namely the years of life lost due to a disease. Would it not be better to keep the factual estimates separate from the moral judgements?

The second issue of principle is whether that particular way of incorporating a moral judgement about aversion to global inequality in life expectancy is a sensible one. It is equivalent to applying an equity weight greater than 1 to each year of life lost in a country where life expectancy at a given age is less than the standard. This implied equity weight will be larger the greater the ratio of the standard life expectancy to the actual life expectancy. If the actual were 20, but the standard 40, then each actual year carries an implied equity weight of 2. When these figures are 30 and 50, respectively, the implied equity weight is 1.6. But when applied to the model life tables that are used in the GBD calculations [8], this process generates some very peculiar results concerning intergenerational equity, both within countries and between countries. Figure 2 shows the implied equity weights for four different regional model life tables, each of which has life expectancy at birth equal to 50 years. It will be seen that, within each country, years of life gained by those who have survived into old age are accorded much higher equity weights than the years of life gained by younger adults, and for ‘East’ and ‘South’, this feature is quite marked. It is hard to imagine what justification there might be on equity grounds for introducing such a bias into the calculations.

But this oddity becomes even more strange when we compare these age profiles of equity weights between rich and poor countries (because high and low life expectancy is closely related to the real income per head in each country). This is shown in Figure 3 for the case of ‘East’ only, and it indicates that old people in richer countries are given more weight on equity grounds than young adults in poorer countries! This is surely an unintended and unwanted side-effect of the Murray–Lopez method of incorporating aversion to inequality in life expectancy into their calculations, and it suggests to me that they should not be averse to abandoning it if something better can be done. I will suggest something better shortly.

It should be noted, however, that when it comes to estimating the impact of particular interventions, the standard life table is abandoned and actual local life tables are used instead. This raises two further questions. The first is: why bother to calculate the global burden of disease, when it has to be abandoned just when its policy value would be greatest? The second is: does equity not matter any more when prioritizing health interventions? It seems to me that in both contexts the ‘years of life lost’ should be the actual years of life lost, estimated as best one can. But equity is still important, and needs to be introduced separately, but explicitly, in both cases. This requires a different method from adopting as the standard the pseudo-Japanese model life table.

IS THERE A VIABLE ALTERNATIVE STRATEGY?

The problems addressed by the GBD enterprise will not go away. They are important and will surely become ever more pressing. The intellectual strength of the enterprise is its explicitness, which
enables others to identify weaknesses and check conclusions. No such undertaking is ever going to be perfect, and those critics who reject such enterprises because they have faults need to look at what happens already and apply the same critical criteria to current practice. We must beware of allowing unattainable perfection to prevent us from recognizing and supporting feasible improvements. It therefore seems to me to be incumbent on critics to say what they would do instead, and why they think their proposals would be better (though not, of course, perfect). That is what I shall attempt to do in this final section.

Starting at a broad strategic level, the most important change that is required is to replace ‘disease’ as the system’s keystone with ‘intervention’. Once this is done, things become much simpler, because we can concentrate on incremental changes and abandon the attempt to calculate comprehensive aggregates on a global basis. But this switch of focus is of interest has other advantages too, because it will simplify the valuation process, reduce the role of experts, and facilitate empirical verification of their judgements. These are considerable advantages, and the sooner the enterprise changes course, the sooner we will be able to enjoy them.

Starting with the monitoring of population health, it was argued above that this is not strictly speaking necessary for priority setting purposes, either for health care or for research. But there may be a different role for such monitoring, namely to assist in the identification of new threats to health, by relating population health to risk factors (such as diet, or smoking behaviour, or environmental pollution). But if the objective is to measure health status as quality adjusted life expectancy in QALY (or DALY) terms, there is no great advantage in pursuing that by the roundabout route of measuring the global burden of disease. It would be much more straightforward to do so by introducing some simple generic measure of HRQOL into any survey of a representative sample of the general population, so as to generate population norms that could then be analyzed by whatever background factors are considered important. The information would be more accurate, less reliant on the remote estimates of ‘experts’, and could be tailor made to local policy concerns. If the Murray–Lopez estimates of life expectancy are indeed superior to any previously available in many countries, they can and should be used in such an enterprise. But the rest of the elaborate machinery becomes redundant.

When it comes to priority setting for health interventions, the GBD has been acknowledged to be irrelevant, but the DALY method of calculating the benefits of interventions is still in play as a potential candidate. As with DALYs, orthodox QALYs are an amalgam of gains in life expectancy and gains in HRQOL, but when focused on an intervention the two items are not rigidly separated, as in the DALY calculations, but measured together by comparing two profiles (as in the example in Figure 4, first published in 1985). This has the great advantage of not assuming that years of life gained are necessarily healthy years of life. In principle, we have here a much more straightforward conceptual framework within which to operate. We still face formidable information requirements, of course, but they become simpler to address in certain key respects. We still need to know the QALY impact of treatments, so we shall still need expert assessments (based on empirical evidence wherever possible) of the differential impacts upon life expectancy and HRQOL in descriptive terms. But we should not rely on experts to value these differences. That is better done by the lay public. Generic measures of HRQOL that have the property of separating description from valuation, and which have used the general population as a source of valuations, have existed for many years. They were recently reviewed by the Washington Panel on Cost-Effectiveness in Health Care, and four were recommended as worthy candidates for such a role [9].

Copyright © 1999 John Wiley & Sons, Ltd.  
Health Econ. 8: 1–8 (1999)
So instead of getting small panels of experts to do all of this work, it would be better to adopt a descriptive system cast in terms that make sense to ordinary people (in a way that the technical names of diseases do not), and a valuation method to which they can respond and which focuses on the actual differences in people’s HRQOL. The weakness of the person trade-off method, as employed latterly in the Murray–Lopez DALY valuations, is that it muddles the valuation of the health states with people’s views about distributional issues. Distributional considerations are important, but they need to be kept separate from the measurement of HRQOL. If samples of the general public each country were used as the source of values (as they could be with systems that focus on health states rather than diseases), and if the very same valuation experiments were conducted on the experts (and on policy-makers), we would have a valuable opportunity to verify the capacity of the latter to make judgements on behalf of the former.

Turning to the distributional considerations, it would be feasible to generate an explicit set of equity weights to replace the implicit ones used to GBD calculations, by adopting the concept of a ‘fair innings’ and carrying through its implications as described in Williams 1997 [10]. This starts from the proposition that if people are averse to inequalities in age at death, both between and within communities, and if we found out empirically the strength of such aversion, and if we set as a reasonable expectation a certain span of years (e.g. seventy) then a set of weights can be derived that have much more plausible characteristics than those implied by the Murray–Lopez method. In Figure 5 a very simple set of weights has been generated by dividing the number 80 (roughly speaking the Murray–Lopez ‘fair innings’ postulate) by people’s actual life expectancy at each age. It will be seen that such weights decline monotonically with age, and even in a poor community will fall below 1 for older people, because even in such communities a few people enjoy the ‘fair innings’.

But the ‘fair innings’ approach has a further potential advantage over the Murray–Lopez method, in that it need not be restricted to life expectancy, but can be generalized to cover quality adjusted life expectancy as well. Distributive justice, expressed in ‘fair innings’ terms, should surely differentiate between people who live to a good age in good health and people who live to a good age but suffer pain and disability most of their lives. To do this, however, we need lifetime profiles of HRQOL. But if the simple survey mentioned above were carried out, we would have those population norms. Moreover, for some populations such norms already exist (see Williams 1997 [10] again). The important thing to remember is that this supplementary data is only required for the generation of equity weights, it is not needed for the main calculations. But these equity weights would nevertheless be a valuable addition to the repertoire of variables to be used in economic evaluations of health care, and if aversion to inequality on people’s life-time experience of health is important, the data needed to estimate them will also be important.

On the subject of guiding research priorities, there is little to add. It will require the measurement of the likely impact of the research on people’s health, if successful, and this can be measured, as described above, without any need for GBD calculations.

CONCLUSIONS

The Global Burden of Disease Study is currently diverting a great deal of highly skilled resources into making calculations that are not needed for the problems they purport to address. The focus on diseases as the central concept is mistaken, and should be replaced by a focus on interventions. This will facilitate considerable simplification of the methods used to elicit health status valuations, and also make it possible to bring lay opinion to bear on matters that are dangerous to leave to
The experts will have quite enough to do in assessing the effectiveness of interventions in descriptive QALY terms, without taking on the valuation task as well. With such a strategic reappraisal, the global burden of disease enterprise will be brought into a more fruitful relationship with health technology assessment, and the redeployed resources will make a more valuable contribution to pressing policy problems there than by continuing to pursue the calculation of aggregates that have too many idiosyncratic elements in them to be capable of straightforward interpretation or application.

REFERENCES


3. See for instance the data cited in the World Development Report, pp. 34–35 (see [1] above) and especially Figure 1.9.


