Progress and Directions in Refining the Global Burden of Disease Approach: A Response to Williams

Christopher J.L. Murray
Alan D. Lopez

Williams [1] calls for a strategic reappraisal of the Global Burden of Disease enterprise. This is a timely suggestion, as use of the burden of disease approach is expanding world-wide; many countries have completed or are undertaking national burden of disease assessments[2], work on the Global Burden of Disease assessment for the year 2000 is underway [3], an independent International Network of Burden of Disease has been established [4] and the World Health Organization is routinely publishing burden of disease results in the World Health Report [5] and promoting its development and application [6]. Since the publication of the first results of the GBD study in the World Development Report 1993 [7], there has been a growing debate on many technical aspects of developing and implementing summary measures of population health [8], in which a number of important conceptual, empirical and ethical aspects relevant to the future evolution of the GBD have been identified. Williams' essay summarises several previously discussed concerns and raises an original argument about distributional issues. For us, this provides a useful opportunity to clarify a number of misconceptions about summary measures of population health and the GBD and to focus attention on the critical issues for future development of this area.

We believe that there is an urgent need to improve the empirical epidemiological basis for periodic assessment of the Global Burden of Disease. There also are a number of key conceptual debates on the construction of summary measures of population health that need further exploration. Some of this empirical and conceptual work is underway at the World Health Organization and in the broader scientific community [9]. We hope that our reply to Williams will stimulate further attention to these key empirical and conceptual challenges. Rather than responding to Williams point by point, we have organised our reflections in seven areas. First, we return to the general aim and goals of the GBD enterprise and emphasise the distinction between the GBD and particular summary measures of population health such as Disability-Adjusted Life Expectancy (DALE) or Disability-Adjusted Life Years (DALYs) used to distil a large body of health information. Second, we dispute Williams' claim that summary measures of population health have no policy relevance, in a brief discussion of the multiple uses of these measures. Third, to facilitate a dialog on the uses and construction of summary measures of population health, we review the two basic families of summary measures (health expectancies and health gaps) and their key characteristics.[10] Fourth, we clarify the link between summary measures of population health and quantifying the benefits of health interventions. Fifth, we review – with a special emphasis on the new directions of work at WHO – one of the key inputs to all summary measures of population health, namely valid and reliable descriptions and valuations of health states. Sixth, we consider Williams' attempts to apply the logic of his fair innings argument [11] to summary measures of population health, and we demonstrate that many health gaps already incorporate a fair innings principle. Finally, we address the advantages and disadvantages of including a range of social values (such as age, discounting, distribution by socio-economic group) directly in summary measures of population health.

Aims and Goals of the GBD

The GBD was initiated by the World Bank and the World Health Organization in an attempt to provide timely information on (a) levels of ill-health from premature mortality and from non-fatal health outcomes, and the contribution of different diseases, injuries and risk factors to these levels of ill-health; and on (b) short to medium-term projections of premature mortality and non-fatal health outcomes. The GBD is an ongoing enterprise. Preliminary results for 1990 [7] were followed by definitive results and extensive documentation of methods, databases and assumptions [12]. Estimates for the GBD in 1998 have now been published in the World Health Report 1999 [5]. A major revision of the GBD for the year 2000 has been launched. Three goals articulated for the GBD [12] remain central: (i) to decouple epidemiological assessment of the magnitude of health problems from advocacy by interest groups of particular health policies or interventions; (ii) to include in international health policy debates information on non-fatal health outcomes along with information on mortality; and (iii) to undertake the quantification of health problems in units that can also be used in economic appraisal. The specific objectives for GBD 2000 are similar to the original objectives: (i) to develop internally consistent estimates of mortality from over 100 major causes of death, disaggregrated by age and sex, for the world and major geographic regions; (ii) to develop internally consistent estimates of the incidence, prevalence, duration, and case-fatality for nearly 500 sequelae resulting from the above causes, disaggregated by age, sex and region; (iii) to describe and value the health states associated with these sequelae of diseases and injuries; (iv) to calculate summary measures of population health and the contribution of different diseases and injuries to population health; (v) to estimate the contribution of major physiological, behavioural, and social risk factors to summary measures of population health by age, sex and region; (vi) to develop alternative projection scenarios of mortality and non-fatal health outcomes, disaggregated by cause, age, sex and region.

As suggested by its aims, goals and objectives, a major preoccupation of the GBD has been to improve the comparability, validity and reliability of the descriptive epidemiology of non-fatal health outcomes and mortality attributed to different diseases, injuries and risk factors. An equally important focus of the GBD has been the improvement of methods to develop alternative projection scenarios of this body of descriptive epidemiological information. Methods including software programs were developed in the GBD and continue to be enhanced through work in various national burden of disease studies and other investigations to improve the cause of death attribution in vital registration data, and to develop internally consistent epidemiological estimates of incidence, prevalence, duration, and mortality for diseases, health states linked to diseases and risk factors. The creation and maintenance of databases on the descriptive epidemiology of major conditions is probably the most formidable, time consuming and resource-intensive task of the GBD enterprise. A variety of measures have been used to analyse the patterns of descriptive epidemiology in the GBD database, including death numbers, probabilities of deaths (between birth and 15 years of age, between ages 15 and 60 and between ages 60 and 70),

years of life lost to premature mortality, prevalence of severity adjusted disability, and years lived in various states of disability. In addition, two summary measures of population health, disability adjusted life expectancy (DALE) and disability adjusted life years (DALYs) have been used to describe the broad patterns.

Williams, like others, incorrectly equates the GBD with one summary measures of population health used extensively in the presentation of the GBD results by cause, namely DALYs[0.03,1] [13]. While the title of his essay refers to a strategic reappraisal of the GBD, he presents no comments on the primary activity of the GBD, which is the development of comparable, valid and reliable epidemiological information on a wide range of diseases, injuries and risk factors. Equating the GBD with one summary measure, DALYs, is a common mistake for those whose only exposure to the literature on the GBD is the World Development Report 1993 or critiques of the GBD. Williams' comments, in fact, more generally pertain to the usefulness of summary measures of population health and technical and normative issues arising in the construction of summary measures. Our response, therefore, is primarily focused on the uses and construction of summary measures of population health.

Uses of summary measures of population health

Summary measures of population health are measures that combine information on mortality and non-fatal health outcomes to represent population health in a single number [14]. Efforts to develop summary measures of population health have a long history [15]. In the past decade, there has been a marked increase in interest in the development, calculation and use of summary measures. Measures such as active life expectancy have been applied widely, especially in the United States. Related summary measures such as Disability-Free Life Expectancy (DFLE) and Impairment-Free Life Expectancy are now commonly calculated [16]. The volume of work from the members of the Reseeau de Esperance de Vie en Sante (REVES) is one indication of the activity in this field. [17]. As noted above, another type of summary measure, Disability-Adjusted Life Years (DALYs), has been used in the Global Burden of Disease Study [12] and a number of National Burden of Disease Studies [2]. Reflecting this rising interest in the academic and policy communities, the United States' Institute of Medicine convened a panel on summary measures and published a report that included recommendations to enhance public discussion of the ethical assumptions and value judgements, establish standards, and invest in education and training to promote use of summary measures.[14]

Interest in summary measures relates to a range of potential applications of them. At least eight uses are worth highlighting here:

1)   *Comparing the health of one population to the health of another population*

  Such comparative judgements are essential to evaluations of the performance of different health systems. Comparisons may allow decision-makers to focus their

5

attention on those health systems with the worst performance. In addition, comparative judgments provide the possibility of analyzing the key contributors to differences in health between populations.

2)   *Comparing the health of the same population at different points in time.*

Monitoring changes in health status over time is essential for the evaluation of health system performance and progress towards stated goals for a given society.

3)   *Identifying and quantifying overall health inequalities within populations.*

4)   *Providing appropriate and balanced attention to the effects of non-fatal health outcomes on overall population health.*

In the absence of summary measures, conditions that cause decrements in function but not mortality tend to be neglected relative to conditions that primarily cause mortality.

5)   *Informing debates on priorities for health service delivery and planning.*

When a summary measure is combined with information on the contributions of different causes of disease and injury or risk factors to the total, such information should be a critical input to debates on the identification of a short-list of national health priorities that will consume the attention of senior managers in public health agencies and government leaders.

6)   *Informing debates on priorities for research and development in the health sector.*

The relative contributions of different diseases, injuries and risk factors to the total summary measure is also a major input to debate on priorities for research and development investment (World Health Organization 1996).

7)   *Improving professional training curricula in public health.*

8)   *Analyzing the benefits of health interventions for use in cost-effectiveness analyses.*

The change in some summary measure of population health offers a natural unit for quantifying intervention benefits in these analyses.

Broad interest and use of summary measures in the policy arena demonstrates the recognition of their value at the practical level for many of these purposes.

Williams focuses on three uses of summary measures: monitoring population health across countries, identifying intervention priorities and identifying research and development priorities. In brief, he argues that summary measures of population health are not useful for any of these purposes. Rather, he argues

that what is required is information on incremental gains in health and the costs of all possible health intervention strategies. His argument is a specific variant of a long-standing thesis in economics that if choices at the margin are made correctly then the current absolute position and ultimate outcome do not matter. Williams implies that the broad public interest in levels of population health simply is misguided and should be redirected to interest in the cost-effectiveness of interventions.

Williams fails to provide a cogent argument why we should not be interested in summary measures of population health for the comparison of population health across place or time. His only argument is that levels of population health are determined by many factors, which makes the attribution of changes in health to particular causes a complex task. This is certainly true, but it in no way decreases our interest in the levels of health achieved in different communities.

What is the use of summary measures of population health decomposed into the contributions of diseases, injuries or risk factors to prioritising investments in different interventions? Neither we, nor our colleagues, have ever claimed that resources should be directed toward health problems solely on the basis of their relative contributions to premature mortality and non-fatal health outcomes. If available resources are to be allocated to minimise the burden of disease or maximise healthy lifespans, allocating resources to interventions proportionate to the size of the problems they address would be logically inconsistent. Williams is constructing a strawman when he implies that we advocate using the GBD alone to select funding priorities. While nearly everyone agrees that one very important input to prioritising resources for interventions is the cost-effectiveness of interventions, nevertheless, debates on priorities for health action can be informed by summary measures of population health decomposed into the contributions of different diseases, injuries and risk factors. If there are fixed assets other than fungible dollars that limit the feasible combinations of interventions that can be delivered – real world examples include the attention of senior Ministry of Health decision-makers or the political commitment of government leaders – then these assets should be devoted not just to the most cost-effective interventions but to those cost-effective interventions with the potential to effect substantial improvements in population health status. To estimate the benefits of an intervention and the total benefits that can be achieved through maximum application of an intervention, a valid assessment of the epidemiology of the disease, injury or risk factor addressed by the intervention is required. In addition to situations where priorities need to be established on the basis of cost-effectiveness and the potential maximum change in population health, summary measures decomposed into causes will be essential to monitor the implementation and impact of specific interventions.

Information on the contributions of diseases, injuries and risk factors to summary measures of population health is also necessary for the health intelligence function of governments. If a disease, injury or risk factor is not yet recognised as a major problem, there will be no attempt to formulate intervention strategies or even to assess the marginal benefits and costs of alternative intervention

strategies. For example, one consequence of the publication of the GBD results for 1990 has been to place dialog on intervention strategies for depression in low and middle-income countries firmly in the policy arena. As one result, efforts to analyse the cost-effectiveness of interventions for depression are underway. If decomposition of summary measures into the contributions of diseases, injuries and risk factors is necessary, this has significant implications for the design of summary measures, favouring health gaps over health expectancies [10].

Williams argues that as with health interventions the only information required to prioritise research and development resources is exhaustive information on the cost-effectiveness of all possible research projects. A panel of the Institute of Medicine tried to operationalize this approach in the 1980s for investments in new vaccines for childhood infectious diseases [18]. The WHO Ad Hoc Committee on Health Research Relating to Future Intervention Options [19] also considered estimating cost-effectiveness of various research and development projects on the basis of the following highly simplified cost-effectiveness ratio:

$$Cost - Effectiveness = \frac{C}{seB}$$

where C is the present value of the cost of the research and implementing the research product, eB is the present value of future health benefits from implementing the product of the research product, which can be broken into the effectiveness of the future intervention and the magnitude of the problem it addresses, and s is the probability of success of the research product as a function of time and dollars invested.[20] This approach was not implemented because a review of past efforts suggested that *ex ante* it was nearly impossible to predict the probability of success of a particular research project, the effectiveness of the research product or the costs of the project. In the extreme, if we have no real information on future costs, effectiveness and probability of success, priorities should be established on the basis of the expected future burden of problems that the R&D may address.

A less extreme view was adopted by the WHO Ad Hoc Committee, which argued that while one of the dominant considerations for R&D priorities should be the expected future burden, the scientific community collectively could make some informed decisions on expected probability of success for broad areas of research. The Committee concluded that it would be useful to carry out research to identify what the characteristics of a R&D product would need to be in terms of costs and effectiveness to make it attractive in the future. Considerable debate remains on the relative importance of information on the magnitude of health problems and scientific judgement on the probability of success and effectiveness for prioritising research funds. For example, in the United States, a panel of the Institute of Medicine recommended burden of disease as one important input to priority setting [21] sparking a vigorous debate in the government [22]. Whether or not there is any information content in the predictions of the probability of success of research projects and the effectiveness of the products of this research, it is clear that information on the contribution of diseases,

injuries and risk factors to summary measures of population health have a much more direct input into R&D prioritisation than into intervention selection.

The argument made by Mooney [23] and Williams – that only incremental changes through intervention, and not the level of population health, matter – appears strange when extended by analogy to national income and product accounts. As Jamison [24] writes on this analogy, "The most natural comparison [to the GBD] is to the development of National Income and Product Accounts (NIPAs)." (pxix) Original pioneering efforts on NIPAs was followed by the codification of international standards in the System of National Accounts [25]. Despite codification, debate has continued unabated on the conceptual and empirical basis for national accounts – should environmental degradation be included in capital depreciation, should household production be included, etc. National accounts measure the level of economic activity in a country. Williams and Mooney must surely argue that this is unnecessary; they must argue that resources wasted on measuring national accounts could be better spent on calculating the incremental gains in national income that could be achieved through various policy options or interventions. Today, the myriad uses of national accounts have so enriched the field of macro-economics that little energy is spent on questioning their utility. Summary measures of population health are, for the health sector, a natural analogue to national income and product accounts.

A typology of summary measures of population health

Further discussion of summary measures will be facilitated by introducing a basic typology of the available options. Summary measures can be divided into two broad families: health expectancies and health gaps [10]. The survivorship function shown in bold in Figure 1 can be used as a heuristic to illustrate the basic differences between health expectancies and health gaps.[26] The x-axis is the age of a cohort and the y-axis is the percent of the cohort in various states (such as alive and fully healthy, alive and in a health state less than full health or dead). The area under the survivorship function is divided into two components, A which is time lived in full health and B which is time lived at each age in a health state less than full health. The familiar measure of life expectancy at birth is simply equal to A+B. A health expectancy is generally of the form:

Health expectancy = $A + f(B)$

where $f(.)$ is a function that weights time spent in B by the severity of the health states that B represents – in most cases a set of health state valuations are used to weight time spent in health states worse than perfect health, but for some measures arbitrary zero or one weights may be used [27]. Many health expectancies have been proposed including Active Life Expectancy, Disability-Free Life Expectancy, Impairment Free Life Expectancy, Disability-Adjusted Life Expectancy, Health-Adjusted Life Expectancy, Years of Healthy Life, and health capital [28].

These health expectancies can be differentiated by several key attributes. First, as with standard life tables, health expectancies can be calculated for a period or for a cohort. The more common period method calculates the health expectancy for a hypothetical birth cohort exposed to currently observed event rates (e.g., rates of mortality, incidence and remission) over the course of their lifetime. Second, health expectancies can be calculated using the original Sullivan prevalence method [29], the double decrement method [30] or the multi-state method [31]. Third, perhaps the most important variation across health expectancies is the implied definition of states worse than ideal health. Many health expectancies are linked to a particular health status measurement instrument; for example, the U.S. National Center for Health Statistics Years of Healthy Life is linked to two questions collected on the National Health Interview Survey. Active life expectancy is a measure linked to activities of daily living (ADLs). Fourth, health expectancies can also be distinguished by the method used to assign values to time spent in health states worse than ideal health. Fifth, other values can be incorporated into health expectancies. Cutler and Richardson, in the calculation of health capital (a type of subjective period health expectancy), includes individuals' discount rates for future health [28].

In Figure 1, a third line is shown at the far right of the graph. This represents a normative goal of survival in full health for the population. In the specific example shown, the normative goal has been set as survival in full health until age 100. By selecting a normative goal for population health, the gap between this normative goal and current survival, area C, quantifies premature mortality. In the specific example shown in Figure 1, this is the familiar measure of potential years of life lost where the potential limit to life is 100. A health gap is generally of the form:

Health gap = $C + g(B)$

where $g(.)$ is a function that weights time spent in B by the severity of the health states that B represents. Note that because health gaps measure a negative entity, namely the gap between current conditions and some established norm for the population, the weighting of time spent in B is on a reversed scale as compared to the weighting of time spent in B for a health expectancy. More precisely, full health is 1 in a health expectancy, whereas death or a state equivalent to death is 1 in a health gap. Because health gaps measure the distance between current health conditions and a population norm for health, they are clearly a normative measure.

Since Dempsey, there has been an extensive development of various mortality gaps [32]. Years of life lost measures are all measures of a mortality gap, or the area between the survivorship function and some implied target survivorship function (area C in Figure 1). A variety of health gaps have been proposed and measured [33] and many others can be derived logically. Health gaps can be distinguished on the basis of four dimensions. First, what is the implied population health target or norm? Health gaps measure the difference between current conditions and a selected target. The explicit or implicit target is a

critical characteristic of any health gap. Second, as with health expectancies, a key issue is how health states worse than ideal health are defined and measured. In Disability-Adjusted Life Years, for example, health states are multi-dimensional and are based on both observations and self-perceptions of performance in different domains. Third, as for health expectancies, health gaps can be distinguished by the method used to value time spent in health states less than ideal health. Fourth, other values such as discounting future health or age-weights have commonly been included in gap measures [34]. Health gaps that include explicit equity weights have also been proposed [35].

In the GBD, two summary measures of population health have been proposed and used. As a simple summary for comparative purposes across populations, we developed a health expectancy, Disability-Adjusted Life Expectancy (DALE) [12]. For purposes of attributing levels of ill-health to various diseases, injuries and risk factors, DALYs have been used extensively in the GBD. In general, health gaps can be decomposed into the contribution of various causes in a more intuitive and easily communicated fashion than health expectancies. One property requested by many users of this information is additive decomposition, whereby the contributions of various different causes (e.g., diseases and injuries) can be aggregated. Additive decomposition can be achieved for health gaps in a straightforward fashion but cannot easily be achieved for health expectancies.

Williams argues that it would be preferable to measure a health gap using local life expectancy. But if local life expectancy is used in calculating health gaps, the implied normative goal for population health is not explicit or, if it is made explicit, has no intuitive appeal. Most disturbingly, Williams offers no explanation why the normative goal for those living in populations with worse health should be lower than for those living in healthier populations. Using different norms also destroys any possibility of using health gaps to compare population health status of different communities or the same community overtime. Murray et al. [10] have shown that health gaps based on local life expectancy have the perverse property that as mortality declines, the health gap actually increases. Using stylised survivorship functions, Figure 2 shows for two populations, with life expectancies at birth of 25 and 37.5, respectively, the mortality component of the health gap defined using local life expectancy (figure A and B) and the mortality component of DALYs defined using a standard life expectancy (figure C and D).[36] By inspection, it is clear that the population norm defined by using local life expectancy appears to have no obvious interpretation and shifts substantially as life expectancy increases. It is because the norm shifts when using local life expectancy that the gap increases as mortality declines. Without any justification, Williams claims that a health gap where the norm is explicitly debated and established such as DALYs is a 'fiction' but the norm defined by local life expectancy is a 'fact'. Clearly both are normative but in one case the norm has a convincing justification, and in the other it does not.

With the growing interest in summary measures of population health, there is an urgent need for a reasoned discourse on the desirable properties of summary

measures for different uses. WHO is orchestrating a global dialog on the uses, design and estimation of summary measures of population health. The original work on the GBD represents a useful start but we must recognise that there is considerable scope to improve future versions of summary measures that are to be used in the next rounds of the Global Burden of Disease enterprise.

Relating health gains from interventions to changes in summary measures of population health

It is essential to distinguish clearly efforts to quantify the health gains from interventions for cost-effectiveness analyses from efforts to apply summary measures of population health in cross-national comparisons or other uses. Imagining that the most desirable summary measure of population health has been identified, one could argue logically that the benefits of a health intervention should be measured as the expected difference in this summary measure for a population with and without the intervention, *ceteris parabis*. If the summary measure used is a health gap (**HG**), then the benefits of an intervention must be formally defined as:

$$InterventionBeneftis = \sum_{t=0}^{t=\infty} \left[HG_t^s - HG_t^c \right] \delta^t$$

where HG_t^c is the health gap with the intervention at time t and HG_t^s is the health gap without the intervetion at time t, and δ is the discount factor.

Alternatively, one could argue that summary measures are primarily meant for comparative purposes and not for the evaluation of health interventions. Thus, there should in general be consistency between the approach used to develop summary measures and that used to estimate the benefits of interventions, without formally defining the benefits as the change in the summary measure. For example, health state preferences might be the same for both uses, but the benefits of interventions could be measured as the increase in healthy years of life lived. With consistency, it is hard to imagine a case where the benefits from an intervention would not equal the change in a health expectancy. However, with changes in survival at very old ages, there are a number of situations where changes in a health gap may not equal the benefits of an intervention evaluated in terms of the extra number of healthy years lived. The counter-argument to this view is that if the health benefits of an intervention are evaluated in a way that is inconsistent with the change in the summary measure, then the summary measure must not adequately reflect how society values health outcomes.

In many cases, analysts who have in principle argued that they are evaluating health interventions in terms of the change in a summary measure have made logical errors. For example, it is a mistake to equate the benefits of an intervention that averts 1000 deaths at age 5 to the total health gap implied by the deaths of 1000 5 year-olds in some population. Clearly, preventing 1000 5-year olds from dying today will increase the number of deaths in the future at older ages, and thus increase the health gap in future years. This must be taken into

account when estimating the benefits of an intervention. Preston [37] and Murray [38] have shown that the change in the present value of future health gaps can be approximated with local cohort life expectancy; it is critical to recognise that this is only an approximation and the only accurate method to estimate benefits is to model future health gaps and health expectancies directly. Below, where the fair innings argument of Williams is examined in detail, this distinction will be important.

Describing and valuing health states

As illustrated in Figure 1, a key step in the construction of a health expectancy or a health gap is comparing time lived in a health state worse than full health with time lived in full health (in health expectancies) and with time lost due to premature mortality, compared to some normative goal (in a health gap). Two sets of issues are common to both health expectancies and health gaps: the conceptual framework and measurement strategy to describe health states and the conceptual framework and measurement strategy to value time spent in health states. The literature on both description and valuation of health states is vast and rapidly expanding [39]. There is no possibility of analysing the subtleties of this literature and their manifold implications for summary measures of population health in this article; Murray (1996) provides a more detailed discussion with regards to the original GBD approach. Rather let us highlight some key points of near universal consensus and some major areas of controversy. This will, we hope, illuminate a range of misunderstandings in Williams' essay.

Health states need to be described in multiple dimensions such as mobility, self-care, pain, cognition, affect, etc. A wide range of instruments have been developed in various languages to use individual responses to measure various dimensions or domains of health states [39]. Some instruments sacrifice the capacity to discriminate between health states by restricting the number of questions or items in the survey and restricting the number of response categories in order to increase measured reliability – for example, this is the strategy used in Euroqol EQ5D, which includes five domains with three level categories on each [40]. Other instruments such as SF-36 have many more items and more response categories per item. Increased discriminatory power often comes at the price of increased complexity, which may have important implications for valuation to time spent in a health state. A fundamental problem with current self-reported instruments is a lack of cross-cultural comparability (including comparisons of the same community over periods long enough that cultural health norms may have changed). This is not simply a question of language and the interpretation of the meaning of different categorical responses in different languages. The endpoints of scales for a given domain such as best or worst mobility may also have very different meanings across different cultures or across socio-economic groups within a society. A classic example comes from Australia, where the aboriginal population with much higher mortality than the rest of the Australian population reports better health status on surveys. In response to the question how do you rate your overall health status, 2 per cent report their health as poor

and 10 percent report their health as fair, as compared to 4.5 percent in the rest of the Australian population reporting poor health and 16 percent reporting fair health [41].

For the foreseeable future, this means that summary measures of population health for comparative purposes must make use of survey results on self-reported health status instruments with great care, and then only if supported by many other condition-specific epidemiological datasets. In order to decompose summary measures into component causes (diseases, injuries or risk factors), many other condition-specific data sets will needed. For causal attribution it will be critical to link diseases, injuries and risk factors to one or several average health states. In the original GBD work, we have emphasised the need to map between aetiologies (diseases, injuries and risk factors) to relatively homogeneous health states that can be described on average in various domains and valued. A thorny issue in causal attribution is comorbidity. If on average we know that individuals with only condition x are in a health state characterised by performance in several domains, and individuals with only condition y are in another health state characterised by different levels of performance in the same domains, what is the likely performance in these domains of individuals with conditions x and y? What is the valuation of time spent in the health states related to condition x alone, y alone or x and y together? In the GBD 1990, an extremely simplistic additive valuation model was used to deal with comorbidity. If on average time spent in the health state of individuals with condition x was valued v1 and on average time spent in the health state of individuals with condition y was valued v2 then time spent by individuals with both conditions was valued v1+v2. Such an additive model can easily be rejected as being implausible. Substantial effort will be required to improve on the estimation of the prevalence of non-independent comorbidity for future iterations of the GBD.

As there are many conceptual and measurement issues in developing adequate descriptions of health states, there are numerous measurement issues in eliciting an individual's valuations of time spent in health states worse than full health. One of the main objectives for the ongoing work on the GBD 2000 effort led by the World Health Organization is to facilitate reliable and valid measurements of valuations of time spent in health states in populations across the world. We are sure that nearly all analysts in this area, including Williams, share similar goals. Through the efforts of many researchers including efforts organised through various national burden of disease studies, some key lessons for improving health state valuations have emerged. First, reflecting a long-standing finding in psychometrics, more valid and reliable measurements are obtained if individuals are asked to value a range of health states from very mild to very severe [42]. Second, combining multiple methods to elicit valuations such as visual analogue, time trade-offs, ordinal rankings, person trade-offs, probably provides more valid results. Inconsistencies in the results from various methods for a set of health states can be fed back to individuals to prompt them to deliberate over their responses. Third, more cognitively complex valuations techniques such as the standard gamble, person trade-off and even the time trade-off become increasingly difficult to use with less educated individuals. If large scale

empirical assessment in many different countries to inform health state valuations for the GBD are to be achieved, instruments that are reliable and valid for populations with widely varying educational attainment need to be developed. Difficulties in valuation surveys are illustrated by a time trade-off survey in the UK that reported standard deviations for valuations of more than 0.6 [43].

Regardless of the resources available, it would not be feasible to measure health state valuations of the population for every possible health state. It is therefore important to develop predictive models that allow an analyst to impute health state valuations from information about the levels on various domains of health status associated with a particular state. To date, there have been at least four published attempts to develop systems that can be used to map from the levels on a set of domains of health status to valuations of health states described along these domains, including the Quality of Well-Being (QWB) scale in the United States [44], the Disability and Distress Scale in the United Kingdom [45], the EuroQol system [43] and the Health Utilities Index (HUI) in Canada [46]. Taken together, the various efforts at linking health state valuations to the domains of health status have suffered from a combination of measurement problems, limited population-based datasets, analytical strategies that have not taken into account measurement error, and attempts to fit implausible models. More conceptual, methodological and empirical work is needed to develop robust models for this purpose.

At the time that the GBD 1990 was underway, and even today, there is no body of empirical measurement of health state descriptions and valuations that can be used (a) to describe the average health state in multiple domains associated with different diseases, injuries and risk factors and (b) to value these average health states. As an effort to provide a practical interim solution to these major data deficiencies, we used a multiple methods (ordinal rankings, person trade-off, time trade-off and visual analogue) approach with small groups of public health professionals to measure values for approximately 20 health states ranging from mild to severe.

Since the development of the original protocol for health state value measurement in the GBD, a series of convenience samples of international public health practitioners has been organized, and a number of modifications and refinements of the original protocol have been examined. In ten different groups, valuations for 15 to 22 states – with a set of 14 states common to all exercises – have been measured using a multi-method approach with internal consistency checks and group discussions. The study locations have included the United States, Mexico, Brazil, the Maghreb countries (Morocco, Algeria and Tunisia), Japan, the Netherlands, and four multi-national groups of health care practitioners. The GBD approach to health valuation has also served as the foundation for a number of experiments in Europe, starting with the Disability Weights Project for Diseases in the Netherlands and continuing with ongoing research by the European Disability Weights Project. Also underway is a multi-country, multi-informant validation study of the GBD disability weights, with results from 14 different countries available thus far [47]. Table 1 summarizes

the Pearson's product-moment correlation coefficients comparing the median health state valuations for each state across the various groups. Overall, the intraclass correlation coefficient for the ten studies is 0.954.

These results indicate that this measurement approach yields similar values in groups from very different communities. The work completed by Ustun and colleagues in 14 countries [47], which measures rank correlations for a set of 17 health conditions, provides further evidence that valuations of health states appear to be quite stable across diverse settings. This stability of valuations, in spite of the heterogeneity of the respondents, suggests that there may be some predictable relationship between a given health state, as described along domains of health status, and valuations for that state obtained through preference measurement methods. Clearly, much empirical work remains to substantiate this hypothesis. We suspect as large scale health state valuation exercises are undertaken in many countries, important variation in average health state valuations will be found, particularly with respect to the contribution of selected domains such as sexual function or pain. Nevertheless, the magnitude of this variation, we expect, will not have major implications for summary measures of population health. At the end of the day, everybody agrees that the health state associated with quadriplegia (paralysis from the neck down) is worse than the health state associated with vitiligo (patches of whitened skin).

Equity, interventions and summary measures

Williams raises an original argument on equity in choosing life saving interventions across age-groups. In brief, he argues that society may be willing to accept a reduction in total population life expectancy in order to obtain greater equality of life expectancy across different groups. He argues that if this willingness to trade total life expectancy for the distribution of life expectancy were measured and operationalized, it would lead to equity weights used to adjust the estimated health benefits of interventions based on the age of the beneficiary. As presented by Williams [11], key steps in the argument are missing. For example, Williams moves from group indifference curves to age-specific 'equity weights' without any derivation or even definitions. In fact, moving from the concept of an indifference curve on health outcomes across individuals to a unique set of 'equity weights' as a function of the age of the beneficiary is far more complex than portrayed, and requires many assumptions that go unstated.

When Williams applies his fair innings argument to a summary measure of population health such as DALYs, he unfortunately commits a fundamental error. He calculates the health gap represented by a death at each age and divides by local life expectancy. He proceeds to call this an age-specific 'equity weight'. We are not told why this is an equity weight nor how this weight would be used in any decision-making context. In this case, he seems to be using local life expectancy as a measure of health benefits from a hypothetical intervention that prevents death in an individual with the average risk profile of the population at a

given age x. No reason is given why this measure of benefit should be compared to the normative statement that a death at age x represents a gap of y years.

Rather, we can derive meaningful implied equity weights if one evaluates the benefits of interventions in terms of a change in a summary measure of population health as opposed to the incremental number of healthy life years lived. More formally, an equity weight implied by using a summary measure of population health to evaluate the benefits of an intervention would equal:

Equity Weight = (change in summary measure of health due to the application of an intervention)/ (change in years of healthy life lived due to the application of an intervention)

This depends, as noted, on the choice of summary measure and on the magnitude of the years of life gained through the application of an intervention at each age. If we examine a set of interventions that prevent death in fully healthy individuals at each age with the average level of risk in the population, benefits of the intervention in life years can be approximated with cohort life expectancy of that population [37]. Again simplifying, if we assume that mortality rates are constant such that period life expectancy and cohort life expectancy are equal, we can calculate the equity weights by age for this set of interventions using period life tables.

Murray [38] has already estimated these implied equity weights for DALYs [48] and the results are reproduced in Figure 3. Evaluating the benefits of an intervention in terms of DALYs[0,0] averted will tend weight equal gains in terms of healthy life years as less important at older ages than at younger ages. Because the assessment is based on DALYs[0,0] this result is not because of age-weighting or discounting. In fact, the health gap method of summarizing population health in general incorporates a type of fair innings concept. Years of extra healthy life added through interventions that prevent death at older ages that approach the fair innings ceiling captured in the population norm for survival in full health are accorded less weight than years of life added to individuals who have not been so fortunate to survive to this age. If interventions are funded to minimise the health gap such as DALYs, this has built in equity considerations such that reductions in social group health disparities will be favoured.

Goodness and fairness

In the literature on health expectancies and health gaps, other values in addition to health state preferences have been incorporated such as age weights, time preference and various distributional concerns. For example, Cutler and Richardson [28] proposed a form of health expectancy which includes an individual's time preference; some forms of DALYs include age weights and time preference; and there are also proposals for equity-weighted DALYs [35]. Such discussions about which values should or should not be included in a summary measure of population health raise several fundamental questions: (1) Is a value such as discounting for future health widely held in the population? (2) Are there other reasons to exclude widely held values based on other first

principles (such as excluding gender or race discriminatory preferences)? and (3) Even if preferences are widely held, and there are no other reasons to reject these values, would it be better to include the value in the summary measure or keep it distinct? The last of these three has critical implications for the design of summary measures. Some values such as concern for the distribution of health outcomes may figure prominently in public decision-making, but it is still debatable whether such fairness concerns should be integrated directly into the summary measure of population health or rather measured independently. The advantages of including fairness concerns in a summary measure is that it places these issues firmly in the health agenda. The disadvantage is that including fairness considerations directly in a summary measure of population health can complicate the summary measure profoundly and does not allow for different trade-offs between goodness and fairness.

Perhaps more importantly, it is clear that there may be fairness concerns that are central to the choice of interventions but are not as relevant to the comparative use of summary health measures. Nord [49] has drawn attention to a preference for giving the same health benefit to the sick as to the healthy, a form of distributional concern. One cannot argue that such a priority to the sick would be relevant to measuring population health even if it is critical to the debate on resource allocation across interventions and beneficiaries. Keeping fairness and goodness considerations distinct in the construction of summary measures of population health allows us to keep track of these different uses and needs. In general, a much sharper distinction needs to be made in the debate on the construction of summary measures of population health and the ethical dimensions of intervention choice. Some values such as discounting or age-weighting may be considered types of fairness considerations (discounting is related to intergenerational fairness and age-weighting to fairness across age groups) Or they may be seen as components of goodness. If the latter, there is a much stronger case that they be incorporated into a summary measure of population health.

Conclusions

Williams' critique has raised a number of issues concerning the uses and importance of measuring, summarizing and interpreting population health status. With the burgeoning interest in summary measures of population health, his observations are particularly timely. Yet they offer only a partial, and at times, misinformed view of the range of considerations that surround the design, estimation and use of summary measures of population health, both health expectancies and health gaps. We urge Williams and other commentators to make the very fundamental distinction between assembling the vast body of empirical epidemiological estimates of diseases, injuries and risk factors, which constitutes one of the main functions of the Global Burden of Disease Study, and the methodological, ethical and conceptual issues that pertain to the development of summary measures of population health. Williams' claim that summary measures of population health are irrelevant to policy formulation is unconvincing and is contradicted by the interest shown by many countries and

international agencies such as WHO in estimating summary measures of population health.

We believe that the GBD and subsequent work on burden of disease analysis at the national and sub-national level has stimulated broader interest in the design, estimation and use of summary measures of population health. The body of work on the burden of disease has also demonstrated the feasibility of estimating health gaps and their use in quantifying the contribution of diseases, injuries and risk factors to population health. And practical methods for enhancing the internal consistency of epidemiological estimates essential for the calculation of summary measures have been developed and disseminated through this work. Notwithstanding this contribution, there is great scope for improvement in all aspects of the GBD endeavour. The World Health Organization through its commitment to assessing the GBD for the year 2000 and to leading an international dialog on the development of summary measures of population health will, we hope, advance this agenda.

An important area that will require further reflection and discourse is the incorporation of distributional concerns into summary measures of population health and the estimation of the benefits of health interventions. Health systems have many goals but nearly everyone would agree that improving average levels of population health and reducing health inequalities are two of the most important. It remains an open debate whether distributional values should be incorporated into the design of summary measures of population health or whether separate measures of the distribution of health across individuals should be routinely assessed. Likewise, it is also a separate debate whether distributional values should be directly incorporated into assessing the benefits of health interventions or kept as a separate component of the evaluation of health interventions. Thanks to Williams raising this issue, it now emerges more clearly that evaluating interventions in terms of cost compared to the change in a health gap often includes a built in set of equity weights favouring interventions that benefit individuals that have not yet had the benefit of surviving to an older age. The population effect, therefore, of allocating health resources to minimize a health gap such as DALYs will tend *ceteris paribus* to reduce social group differences in health.

1 Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* 1999; 8:1-8.

2 Many national burden of disease studies have been completed or are underway. Bobadilla JL. *Searching for essential health services in low- and middle-income countries: A Review of Recent Studies on Health Priorities*. Human Development Department, World Bank, 1996. For selected published work see Lozano R. Frenk J. Gonzalez MA. El peso de la enfermedad en adultos mayores, Mexico 1994. *Salud Publica de Mexico*1994; 38: 419-429.; Lozano R. Murray CJL. Frenk J. Bobadilla J. Burden of Disease Assessment and Health System Reform: Results of a study in Mexico. *Journal for International Development* 1995.; 7(3): 555-564.; Fundación Mexicana Para La Salud. *Health and the Economy: Proposals for Progress in the Mexican Health System*. Mexico, Salud, 1995.; República de Colombia Ministerio de Salud. *La Carga de la Enfermedad en Columbia*. Bogota, Ministerio de Salud, 1994.; Ruwaard D. Kramers PGN. *Public health status and forecasts*. The Hague: National Institute of Public Health and Environmental Protection, 1998.; Bowie C. Beck S. Bevan G. Raftery J. Silvertion F. and Stevens J. Estimating the burden of disease in an English region. *Journal of Public Health Medicine* 1997; 19:87-92.; Concha M., Aguilera, X. Albala C. et al. Estudio ecarga de enfermedad informe final. Estudio Prioridades de Inversion en salud Minsterio de Salud, 1996.; Murray CJL. Michaud CM. McKenna MT. and Marks JS. *U. S. Patterns of Mortality by County and Race: 1965-1994*. Cambridge: Harvard Center for Population and Development Studies and Centers for Disease Control, 1998.; Murray CJL. Mahapatra P. Ashley R. Michaud C. George A. Horbon P. Akhavan D. et al. *The Health Sector in Mauritius: Resource Use, Intervention Cost and Options for Efficiency Enhancement*. Cambridge, Harvard Center for Population and Development Studies, 1997.

3 Work has commenced at WHO on a new Global Burden of Disease Study for the year 2000. The study will incorporate several changes to the diseases, injuries and risk factors suggested by national burden of disease efforts as well as empirical databases and research information which have come to light since the 1990 Study.

4 In order to provide a forum for the exchange of information and promotion of national burden of diseases studies, an International Burden of Disease Network was launched in Atlanta in March 1997. The network is coordinated by the Centre for Health Care Development, Liverpool, UK. A description of the objectives and activities of the network is available in the report of the Atlanta meeting or from the coordinator (Howard Seymour: email: howard@ched.org).

5 World Health Organization. . *World Health Report 1999. Making a Difference*. Geneva: World Health Organization; 1999.

6 Brundtland GH. *The global burden of disease*. Speech delivered on December 15, 1998, Geneva.

7 World Bank. *World Development Report 1993. Investing in health*. New York: Oxford University Press for the World Bank, 1993.

8 For selected articles see Anand S. Hanson K. Disability-Adjusted Life Years: A Critical Review. *Journal of Health Economics* December 1997; 16(6): 685-702.; Anand S. Hanson K. DALYs: Efficiency versus Equity. *World-Development* February 1998; 26(2): 307-10.; Andrews G. Mathers C. and Sanderson K. The burden of disease. *Medical Journal of Australia* 1998; 169: 156-8.; Anil G. Burden of disease and cost of ill health in India : Setting priorities for health interventions during the Ninth Plan. National Council of Applied Economic Research 1997; 29(2): 133-72.; Appleby L. Shaw J. Amos T. and Dennehy J. Global burden of disease [letter; comment]. *Lancet* 1997; 350: 143.; Barendregt J J. Bonneux L. and Van der Maas P. J. DALYs: the age-weights on balance [see comments]. *Bulletin of the World Health Organization* 1996; 74: 439-43.; Berman S. Otitis media in developing countries. *Pediatrics* 1995; 96: 126-31.; Bradley DJ. Tropical diseases: the burden and its implications. *Schweizerische Medizinische Wochenschrift. Journal Suisse de Medecine* 1997; 127: 1592-7.; Bradshaw D. and Schneider M. Priority setting in health care: burden of disease cannot be jettisoned [letter; comment]. *Australian & New Zealand Journal of Public Health* 1998*;* 22: 517.; Bremberg S. Health promotion in school age children. *Scandinavian Journal of Social Medicine* 1998; 26: 81-4.; Christie S. and Tobias M. The burden of infectious disease in New Zealand. *Australian & New Zealand Journal of Public Health* 1998; 22: 257-60.; Connors GL. and Hilling L. Prevention, not just treatment. *Respiratory Care Clinics of North America* 1998; 4: 1-12.; Curlin P. and Tinker A. Women's health. *Infectious Disease Clinics of North America* 1995; 9: 335-51.; Eisenberg L. Global burden of disease [letter; comment]. *Lancet* 1997; 350: 143.; Fernandez Martin J. Pereira Candel J. and Torres Cantero A. Una agenda a debate: el informe del Banco Mundial "Invertir en salud". *Revista Espanola De Salud Publica* 1995; 69: 385-91.; Finlay JF. Law MM. Gelmon LJ. and de Savigny D. A new Canadian health care initiative in Tanzania [see comments]. *CMAJ* 1995; 153: 1081-5.; Foster S. and Phillips M. Economics and its contribution to the fight against malaria. *Annals of Tropical Medicine & Parasitology* 1998; 92: 391-8.; Guralnik JM. Fried LP. and Salive ME. Disability as a public health outcome in the aging population. *Annual Review of Public Health* 1996; 17: 25-46.; Gwatkin DR. Global burden of disease [letter; comment]. *Lancet* 1997; 350(141): discussion 144-5.; Hinman AR. Quantitative policy analysis and public health policy: a macro and micro view. *American Journal of Preventive Medicine*1997; 13(1):6-11.; Hyder AA. Rotllant G. and Morrow RH. Measuring the burden of disease: healthy life-years. *American Journal of Public Health* 1998; 88: 196-202.; Jamison DT. Saxenian H. and Bergevin Y. Investing in health wisely. The role of needs-based technology assessment. *International Journal of Technology Assessment in Health Care* 1995; 11: 673-84.; Kane M. Global programme for control of hepatitis B infection. *Vaccine* 1995; 13 (Suppl 1): S47-9.; Karim MS.

Disease pattern, health services utilization and cost of treatment in Pakistan. *JPMA - Journal of the Pakistan Medical Association* 1993; 43: 159-64.; Kochi A. Nunn P. Dye C. and Tayler E. Global burden of disease [letter; comment]. *Lancet* 1997; 350: 142.; Kramers PG. and van der Maas PJ. 'Volksgezondheid toekomst verkenning' 1997. IV. Gezondheid en levensverwachting gewogen. *Nederlands Tijdschrift voor Geneeskunde* 1998; 142: 1277-80.; Lancet. Global burden of disease. *Lancet* 1997 Jul 12; 350(9071): 141-144.; Lozano-Ascencio R. Frenk-Mora J. and Gonzalez-Block M. A. El peso de la enfermedad en adultos mayores, Mexico 1994. *Salud Publica de Mexico* 1996; 38: 419-29.; Maetzel A. Costs of illness and the burden of disease [editorial; comment]. *Journal of Rheumatology* 1997; 24: 3-5.; Martens WJ. Niessen LW. Rotmans J. Jetten TH. and McMichael AJ. Potential impact of global climate change on malaria risk [see comments]. *Environmental Health Perspectives* 1995; 103: 458-64.; Matsubayashi K. Okumiya K. Nakamura T. Fujisawa M. and Osaki Y. Global burden of disease [letter; comment]. *Lancet* 1997; 350: 144.; Mbizvo M T. Reproductive and sexual health: a research and developmental challenge. *Central African Journal of Medicine* 1996; 42: 80-5.; McKee M. and Britton A. The positive relationship between alcohol and heart disease in eastern Europe: potential physiological mechanisms. *Journal of the Royal Society of Medicine* 1998; 91: 402-7.; Meerding WJ. Demographic and epidemiological determinants of healthcare costs in the Netherlands: cost of illness study. *BMJ* 1998 Jul 11; 317 (7151):111-5.; Meltzer MI. Rigau-Perez JG. Clark GG. Reiter P. and Gubler DJ. Using disability-adjusted life years to assess the economic impact of dengue in Puerto Rico: 1984-1994. *American Journal of Tropical Medicine & Hygiene* 1998; 59: 265-71.; Mooney G. Irwig L. and Leeder S. Priority setting in health care: unburdening from the burden of disease [editorial] [see comments]. *Australian & New Zealand Journal of Public Health* 1997; 21: 680-1.; Mosley WH. Population change, health planning and human resource development in the health sector. *World Health Statistics Quarterly - Rapport Trimestriel de Statistiques Sanitaires Mondiales* 1994; 47: 26-30.; Oortwijn WJ. Vondeling H. and Bouter L. The use of societal criteria in priority setting for health technology assessment in The Netherlands. Initial experiences and future challenges. *International Journal of Technology Assessment in Health Care* 1998; 14: 226-36.; Politi C. Carrin G. Evans D. Kuzoe FA. and Cattand PD. Cost-effectiveness analysis of alternative treatments of African gambiense trypanosomiasis in Uganda. *Health Economics* 1995; 4: 273-87.; Prabhakar R. Tuberculosis--the continuing scourge of India. *Indian Journal of Medical Research* 1996; 103: 19-25.; Ramaiah K. D. Kumar KN. Ramu K. Pani SP. and Das PK. Functional impairment caused by lymphatic filariasis in rural areas of south India. *Tropical Medicine & International Health* 1997; 2: 832-8.; RIVM. *Health, prevention and health care in the Netherlands until 2015.* Bilthoven, The Netherlands Elsevier 1997.; Roberts I. Global burden of disease [letter; comment]. *Lancet* 1997; 350, 144.; Robine JM. Measuring the burden of disease. *Lancet* 1998; 352: 757-8.; Sayers BM. and Fliedner TM. The critique of DALYs: a counter-reply. *Bulletin of the World Health Organization* 1997; 75: 383-4.; Schwartlander B. Global burden of disease [letter; comment]. *Lancet* 1997; 350: 141-2; discussion 144-5.; Seim AR. Godal T. and Lie SO. En globaL utfordring i prioriteringsdebatten.

Tidsskrift for Den Norske Laegeforening 1997; 117: 38-42.; Victora CG. and Van Haecke P. Vitamin K prophylaxis in less developed countries: policy issues and relevance to breastfeeding promotion. *American Journal of Public Health* 1998; 88: 203-9.; Yu MY. and Sarri R. Women's health status and gender inequality in China. *Social Science & Medicine* 1997; 45, 1885-98.; Zuber PL. McKenna MT. Binkin NJ. Onorato IM. and Castro KG. Long-term risk of tuberculosis among foreign-born persons in the United States. *JAMA* 1997; 278, 304-7.

9 Work on the preparation of the GBD 2000 Study is being carried out at WHO and includes major efforts to revise disease and injury burden estimates for over 100 conditions and for several risk factors including hazards such as indoor air pollution and BMI which were not quantified in the GBD 1990 Study, revised projection methods and scenarios, new conceptual and empirical work on health status measurement and valuation, measurement of health inequalities, development of international classification systems and the operation of an information dissemination service. A major component of the GBD 2000 effort will be a conference on summary measures of population health scheduled to take place later in 1999.

10 This section makes liberal use of a more in-depth analysis by Murray CJL. Salomon JA. and Mathers C. On summary measures of population health. *Bulletin of WHO,* 1999 (in submission). In this paper, the authors have proposed criteria for choosing among various summary measures that have been proposed or are in use.

11 Williams A. Inter-generational equity: an explorationo f the 'fair innings' argument. *Health Economics* 1997; 6:117-132.

12 Murray CJL. Acharya AK. Understanding DALYs. *Journal of Health Economics* 1997; 16: 703-730; Murray CJL. and Lopez A. Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet* 1997; 349: 1269-1276; Murray CJL. and Lopez A. Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global Burden of Disease Study. *Lancet* 1997; 349: 1347-1352; Murray CJL. and Lopez A. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997; 349: 1436-1442; Murray CJL. and Lopez A. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997; 349: 1498-1504; Murray CJL. and Lopez A. eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996.; Murray CJL. and Lopez A. *Global Health Statistics*. Cambridge, Harvard University Press, 1996.; Murray CJL. and Lopez A. Evidence-based health policy—lessons from the Global Burden of Disease Study. *Science* 1996; 274:740-743.

13 There are many different types of DALYs. The standard notation is DALYs[r,K] where r is the discount rate and K is the age-weighting parameter. The more common forms of DALYs that have been widely published are DALYs[0.03,1] which is the form published in the original World Development Report 1993 where the discount rate is 3% and there is full age-weighting and DALYs[0,0] where the discount rate is zero and age-weights are uniform for all ages.

14 Field MJ. Gold GM. eds. *Summarizing Population Health: Directions for the Development and Application of Population Metrics*. Institute of Medicine, Washington, D.C. National Academy Press, 1998.

15 For some examples of this extensive literature see Katz S. Akpom CA, Papsidero JA, Weiss ST. Measuring the health status of populations. In: Berg RL, ed. *Health Status indices*. Chicago, Hospital Research and Educational Trust, 1973, 39-52.; Sanders BS. Measuring community health levels. *American Journal of Public Health,* 1964; 54:1063-70.; Chiang CL. *An index of health: mathematical models. Public Health Services Publications 1000 Series2. No.5.* Washington DC, National Center for Health Statistics, 1965.; Fanshel S. Bush JW. A health-status index and its application to health services outcomes. *Operations research,* 1970; 18(6):1021-1066; Sullivan DF. *Conceptual problems in developing an index of health. US public Health Service Publication Series No. 1000. Vital and Health Statistics Series 2. No. 17.* National Center for Health Statistics, 1966.; Piot M. Sundaresan TK. *A linear programme decision model for tuberculosis control. Progress report on the first test-runs.* Geneva, World health Organization, 1967, WHO/TB/Tech.Information/67.55.; Preston SH. Health indices as a guide to health sector planning: a demographic critique. In Gribble JN. Preston SH. eds. *The epidemiological transition. Policy and planning implications for developing countries*. Washington D.C., National Academic Press, 1993.; Ghana Health Assessment Project Team. A quantitative method of assessing the health impact of different diseases in less developed countries. *International journal of epidemiology,* 1981;10(1):72-80.

16 Bone MR. International efforts to measure health expectancy. *J. Epidemiol. Community. Health* 1992; 46, 555-558.; Bronnum HH. Trends in health expectancy in Denmark, 1987-1994. *Dan. Med. Bull.* 1998;45, 217-221.; Mutafova M. van-de-Water HP. Perenboom RJ. Boshuizen HC. and Maleshkov C. Health expectancy calculations: a novel approach to studying population health in Bulgaria. *Buletin of the .World Health Organanization,* 1997; 75; 147-153.; Sihvonen AP. Kunst AE. Lahelma E. Valkonen T. and Mackenbach JP. Socioeconomic inequalities in health expectancy in Finland and Norway in the late 1980s. *Soc.Sci.Med.* 1998; 47, 303-315.; Valkonen T. Sihvonen AP. and Lahelma E. Health expectancy by level of education in Finland. *Soc.Sci.Med.* 1997; 44, 801-808.

17 Robine JM. et al. eds. *Calculation of health expectancies: harminization, consensus achieved and future perspectives*. London, John Libbey Eurotex,

1993.; Mathers C. McCallum J. Robine J-M. eds. *Advances in health expectancies: proceedings of the 7th meeting of the International Network on Health Expectancy (REVES), Canberra*. Canberra, Australian Institute of Welfare: AGPS, 1994.

18 Institute of Medicine. *New vaccine development. Establishing priorities volume II. Diseases of importance in developing countries*. Washington DC: National Academy Press, 1986.

19 World Health Organization. *Investing in health research and development. Report of the Ad Hoc Committee on Health Research Relating to Future Intervention Options*. Geneva, World Health Organization, 1996.

20 More complicated approaches for R&D prioritisation can be derived on the basis of health maximization when there is an earmarked research and development budget and whether or not the health sector budget is fixed or influenced by the set of technologies available at a given time. This discussion should be taken as simply illustriative of the problems of estimating cost, effectiveness and the probability of success.

21 Institute of Medicine. *Improving priority setting and public input at the National Institutes of Health*. Washington DC, National Academy Press, 1998.

22 Hartley M. Funding priorities at the National Institutes of Health hot topic at Hastings Center Conference. Haley, S. *Washington Fax: An Information Service* (http://www.washingtonfax.com/p1/1999/19990210.html.), Bradie Metheny, 10 February 1999.

23 Mooney G. Irwig L. Leeder S. Priority setting in health care: unburdening from the burden of disease. *Aust NZ J Public Health* 1997; 21(7):680-681.

24 Jamison D. Global Burden of Disease and Injury Series Foreword. In: Murray CJL. and Lopez A. eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996.

25 Commission of the European Communities, International Monetary Fund, Organisation for Economic Cooperation and Development, United Nations. World Bank. *System of National Accounts 1993*. Brussels, Luxembourg, New York, Paris, Washington,D.C., 1993.

26 Figure 1 graphically illustrates the magnitude of both health expectancies and health gaps only when a population has a stable distribution with a zero population growth rate. In practice, health expectancies are not sensitive to differences in the age structure of different populations. Health gaps are usually reported in absolute terms so that health gaps are sensitive to variations in the age distribution of different populations although age independent forms of health gaps can be formulated.

27 One health expectancy is disability-free life expectancy in which time spent in any health state categorized as disabled is assigned arbitrarily a weight of zero, and time spent in any state categorized as not disabled is assigned a weight of one (i.e., equivalent to full health). The operational definition of disabled varies by data source and questions used in various surveys and censuses. Not surprisingly, disability-free life expectancy cannot be compared in a meaningful way across communities or even in the same community over time.

28 There is an extensive literature on health expectancies, for examples see Cutler DM. Richardson E. The value of health. *American Economic Review*, 1998; 88(2): 97-100.; Cutler DM. Richardson E. Measuring the health of the U.S. population. *Brookings Paper: Microeconomics,*1997; 217-271.; Katz S. Branch LG. Branson MH. Papsidero JA. Beck JC. Greer DS. Active life expectancy. *The New England Journal of Medicine,* 1983; 309(20):1218-24.; Robine JM. et al. eds. . *Calculation of health expectancies: harminization, consensus achieved and future perspectives*. London, John Libbey Eurotex, 1993.; Mathers C. McCallum J. Robine J-M. eds. *Advances in health expectancies: proceedings of the 7th meeting of the International Network on Health Expectancy (REVES), Canberra*. Canberra, Australian Institute of Welfare: AGPS, 1994.; Murray CJL. and Lopez AD Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global Burden of Disease Study. *Lancet* 1997; 349: 1347-1352; Erickson P. Wilson R. Shannon E. *Years of healthy life*. Hyattsville, Maryland: US National Center for Health Statistics, 1995. Health capital proposed by Cutler and Richardson was proposed as an age-specific subjective cohort health expectation although the concept can be generalized to a summary measure of population health.

29 Sullivan DF. *Conceptual problems in developing an index of health. US public Health Service Publication Series No. 1000. Vital and Health Statistics Series 2. No. 17*. National Center for Health Statistics, 1966.

30 Katz S. Branch LG. Branson MH. Papsidero JA. Beck JC. Greer DS. Active life expectancy. *The New England Journal of Medicine,*1983; 309(20):1218-24.

31 Branch LG. Guralnik JM. Foley DJ. Kohout FJ. Wetle TT. Ostfeld A. Katz S. Active life expectancy for 10,000 Caucasian men and women in three communities. *Journal of Gerontology* 1991; 46(4): M145-50.; Rogers RG. Rogers A. Belanger A. Longer life but worse health? Measurement and dynamics. *Gerontologist,* 1990; 30:640-9.; Mathers CD. Robine J-M. How good is Sullivan's method for monitoring changes in population health expectancies. *Journal of Epidemiology and Community Health,* 1997; 51: 80-86.

32 The original concept of a mortality gap was proposed by Dempsey M. Decline in tuberculosis: the death rate fails to tell the entire story. *American review of tuberculosis* 1947; 56: 157-164. Some of the evolution of mortality gaps is discussed in Murray CJL. Rethinking DALYs. In: Murray CJL. and Lopez A.

eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996, chapter 1.; and, Romeder JM. and McWhinnie JR. Potential years of life lost between ages 1 and 70: an indicator of premature mortality for health planning. *International Journal of Epidemiology,* 1977; 6:143-151.

33 Murray CJL. Rethinking DALYs. In: Murray CJL. and Lopez A. eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996, chapter1.; World Bank. *World Development Report 1993: investing in health.* New York: Oxford University Press for the World Bank, 1993.; Hyder AA. Rotllanat G. Morrow RH. Measuring the Burden of Disease: Healthy Life-Years. *American Journal of Public Health,*1998; 88(2):196-202.; Ghana Health Assessment Project Team. A quantitative method of assessing the health impact of different diseases in less developed countries. *International journal of epidemiology,* 1981; 10(1):72-80.

34 Murray CJL. Rethinking DALYs. In: Murray CJL. and Lopez A. eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996, chapter1.; Barnum H. Evaluating healthy days of life gained from health projects. Social Science Medicine 1987; 24(10):833-841.

35 Finn Diderichsen. National DALY-study in Sweden (www.fhinst.se)

36 This figure is easily constructed. If S(x) is the survivorship function, then the population health norm, G(x) is defined such that G(x+L(x))=S(x) where L(x) is the loss of potential life (local life expectancy or the standard life expectancy) at each age.

37 Preston S. Health indices as a guide to health sector planning: a demographic critique. In: Gribble JN. Preston SH eds. *The epidemiological transition. Policy and planning implications for developing countries*. Washingtong DC: National Academy Press, 1993.

38 Murray CJL. Rethinking DALYs. In: Murray CJL. and Lopez A. eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996, chapter1.

39 See for example McDowell I. and Newell C. *Measuring health. A guide to rating scales and questionnaires. Second Edition*. Oxford: Oxford University Press; 1996.; Krabbe PFM. Essink-Bot M. and Bonsel GJ.; The comparability and reliability of five health-state valuation methods. *Social Science and Medicine* 1996; 45:1641-1652.; Brazier J. Usherwood T. Harper R. Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *Journal of Clinical Epidemiology* 1998; 51(11): 1115-1128; Nord E. Methods for quality adjustment of life years, *Social Science and Medicine* 1992; 34: 559-569.

40 Van Agt HME. Essinck-Bot ML. Krabbe PFM. et al. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. *Social Science and Medicine* 1994;39:1537-1544.

41 Australian Institute of Health and Welfare. *Australia's health 1996: the fifth biennial report of the Australian Institute of Health and Welfare*. Canberra, AGPS, 1996.

42 Teghtsoonian R. Range Effects in Psychophysical Scaling and a Revision of Stevens' Law. *American Journal of Psychology*1973; 86(1): 3-27.; Poulton EC. The New Psychophysics: Six Models for Magnitude Estimation. *Psychological Bulletin* 1968; 69(1): 1-19.

43 Dolan P. Gudex C. Kind P. and Williams A. The Time Trade-off Method: Results from a General Population Study. *Health Economics* 1996; 5: 141-154.

44 Kaplan RM. and Bush JW. Health-related quality of life measurement of evaluation research and policy analysis. *Health Psychology* 1982;1:61-80.

45 Rosser R. Kind P. A scale of valuations of states of illness: is there a social consensus? *International Journal of Epidemiology*, 1978; 7: 347-358.

46 Feeny DH. Torrance GW. and Furlong W.J. Health Utilities Index. in *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia, Lippincott-Raven, 1996, 239-251.; Torrance GW. Feeny D. Furlong WJ. Barr RD. Zhang Q. Wang Q. Multiattribute utility function for a comprehensive health status classification system: Health Utility Index Mark 2. *Medical Care* 1996; 34: 702-722.

47 Ustun TB. Rehm J. Chatterji, S. Trotter R. Room R. Bickenbach J. Are disability valuations universal? Multiple-informant ranking of the disabling effects of different health conditions in 14 countries, *Lancet* 1999, (in submission).

48 In fact the analysis was for years of life lost (YLLs), the premature mortality component of DALYs which is the same as DALYs if there everyone alive is in a state of full health.

49 Nord E. Pinto JL. Richardson J. Menzel P. Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* 1999; 8:25-39.
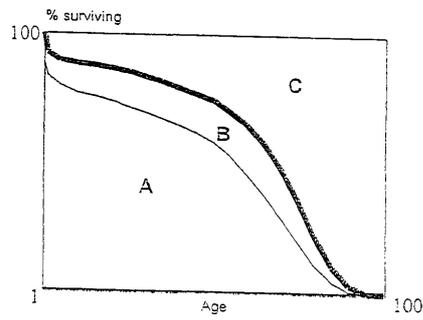
Christopher J.L. Murray
Alan D. Lopez

Table 1. Pearson's Correlation Coefficients for Median Disability Weights for Ten Exercises, Based on 14 Conditions Common to All Exercises

	Int'l I	Nether.	Maghreb	Japan	GBD	Int'l II	CDC	Brazil	Mexico	Int'l III
International I	1.00									
Netherlands	0.98	1.00								
Maghreb	0.96	0.95	1.00							
Japan	0.94	0.88	0.91	1.00						
GBD	0.97	0.95	0.95	0.92	1.00					
International II	1.00	0.98	0.95	0.93	0.97	1.00				
CDC	0.99	0.99	0.95	0.90	0.95	0.96	1.00			
Brazil	0.95	0.91	0.90	0.83	0.91	0.94	0.94	1.00		
Mexico	0.96	0.92	0.92	0.96	0.92	0.96	0.93	0.97	1.00	
International III	0.96	0.95	0.92	0.93	0.93	0.97	0.96	0.97	0.96	1.00

Christopher J.L. Murray
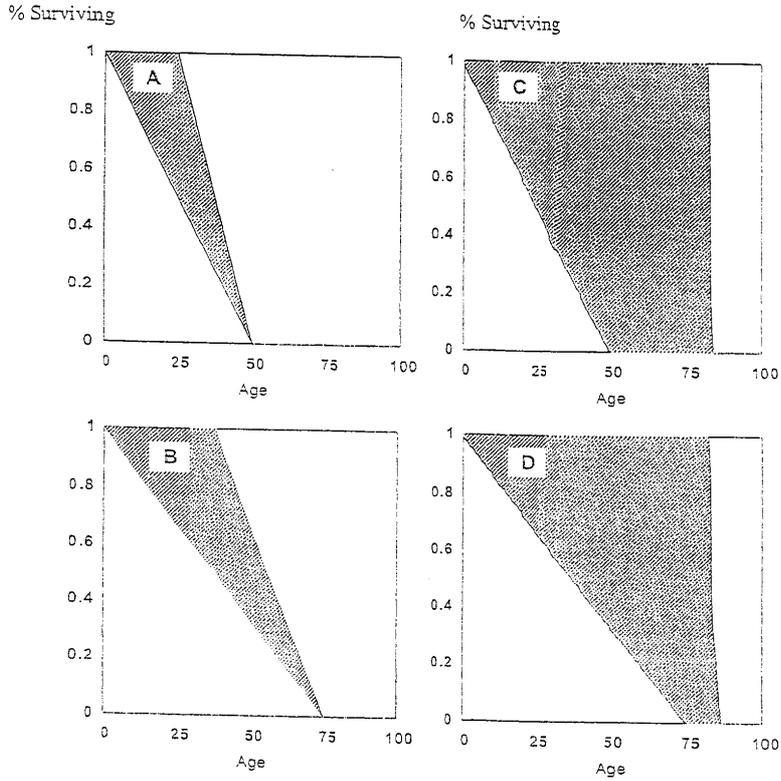Alan A Lopez

Figure 1.

The survivorship function for a hypothetical population.

Christopher J.L. Murray
Alan A Lopez

Figure 2.

Survivorship curves and health gaps in two hypothetical populations

The two graphs at the top represent a population with life expectancy at birth of 25 years, wile the two graphs at the bottom indicate a population with life expectancy at birth of 37.5 years. In each graph, the bold diagonal represents a survivorship norms based on local life expectancy, while in graphs C and D, the health gap is calculated using standardized survivorship norms based on a model life table with e(0)=82.5.

Christopher J.L. Murray
Alan D. Lopez

Figure 3.

Ratio of Change in DALYs to change in the number of life—years
lived in a population due to the application of an intervention
that prevents a single death at a given age.