DISABILITY ADJUSTED LIFE YEARS (DALYS) FOR DECISION-MAKING?
An overview of the literature

Julia A Fox-Rushby

OHE

DISABILITY ADJUSTED LIFE YEARS (DALYS) FOR DECISION-MAKING?
An overview of the literature

Julia A Fox-Rushby

Office of Health Economics
12 Whitehall London SW1A 2DY

About the author
Dr Julia Fox-Rushby is Senior Lecturer at the Health Policy Unit of the London School of Hygiene and Tropical Medicine. In addition to DALYs, her current research interests focus on the measurement and valuation of health related quality of life in low- and middle-income countries, on applications to cost-effectiveness analyses in those settings, and on considering the generalisability of economic evaluations across settings.

OFFICE OF HEALTH ECONOMICS

Terms of Reference

The Office of Health Economics (OHE) was founded in 1962. Its terms of reference are to:

● commission and undertake research on the economics of health and health care;

● collect and analyse health and health care data from the UK and other countries;

● disseminate the results of this work and stimulate discussion of them and their policy implications.

The OHE is supported by an annual grant from the Association of the British Pharmaceutical Industry and by sales of its publications, and welcomes financial support from other bodies interested in its work.

Independence

The research and editorial independence of the OHE is ensured by its Policy Board:

Chairman:
Professor Tony Culyer – *University of York*
Members:
Professor Michael Arnold – *University of Tübingen*
Professor Patricia Danzon – *The Wharton School of the University of Pennsylvania*
Professor Naoki Ikegami – *Keio University*
Dr Trevor Jones – *Director General of the Association of the British Pharmaceutical Industry*
Ms Chrissie Kimmons – *Glaxo SmithKline plc*
Professor David Mant – *University of Oxford*
Dr Nancy Mattison – *The Mattison Group Inc*
Dr John Patterson – *AstraZeneca plc and President of the Association of the British Pharmaceutical Industry*
Professor Sir Michael Peckham – *University College, University of London*

Peer Review

All OHE publications have been reviewed by members of its Editorial Board and, where appropriate, other clinical or technical experts independent of the authors. The current membership of the Editorial Board is as follows:

Professor Christopher Bulpitt – *Royal Postgraduate Medical School, Hammersmith Hospital*
Professor Martin Buxton – *Health Economics Research Group, Brunel University*
Professor Tony Culyer – *Department of Economics and Related Studies, University of York*
Dr Jennifer Dixon – *The King's Fund*
Professor Hugh Gravelle – *Centre for Health Economics, University of York*
Mr Geoffrey Hulme – *Director, Public Finance Foundation*
Professor Carol Propper – *Department of Economics, University of Bristol*
Professor Bonnie Sibbald – *National Primary Care R&D Centre, University of Manchester*
Mr Nicholas Wells – *Head of European Outcomes Research, Pfizer Ltd*
Professor Peter Zweifel – *Socioeconomic Institute, University of Zurich*

CONTENTS

GLOSSARY OF ABBREVIATIONS

DALE	Disability adjusted life expectancy
DALY(s)	Disability adjusted life year(s)
DfID	(UK) Department for International Development
EQ5D	EuroQol five dimensional questionnaire on health-related quality of life
HIV	Human Immunodeficiency Virus
ICD9	9th revision of the International Classification of Disease
ICIDH	International Classification of Impairments, Disabilities and Handicaps
NIH	(US) National Institutes of Health
PTO	Person trade-off technique for valuing health states
QALY(s)	Quality adjusted life year(s)
WHO	World Health Organisation
WHOQOL	WHO quality of life questionnaire
YLD(s)	Year(s) of life lost from disability
YLL(s)	Year(s) of life lost from premature death

8

EXECUTIVE SUMMARY

What are DALYs?
Disability adjusted life years (DALYs) are used to calculate life years
lost from a wide range of diseases and injuries, adjusted for assump-
tions about the severity of mental or physical disability as well as age
and discounted over time. Thus DALYs are one way of measuring the
loss from living a shorter life with disease rather than living a longer
life without disease. At present DALYs are a tool of estimation using
secondary data, rather than one developed for primary data collection.

Why were DALYs developed?
DALYs were developed to help shape the of health policies of interna-
tional institutions such as the World Bank and World Health
Organisation. Their introduction was designed to broaden the usual
focus of these institutions from measuring disease in terms of mortal-
ity and morbidity to including an estimate of the impact of morbidi-
ty. As the focus of the institutions was international, the DALY was
intended to enable many forms of comparison: across diseases, coun-
tries, curative/preventive care as well as different time periods.

Once calculated, the two prime purposes of DALYs were: as an
input to the calculation of the global burden of disease and as an out-
come measure for use in cost-effectiveness analysis. Both were intend-
ed to influence the:

- prioritisation of health care spending within and across countries
  for curative and preventive care;
- flow of funds within health research and development;
- identification of disadvantaged groups for targeting health inter-
  ventions;
- composition of training for clinical and health practitioners;
- methods for assessing performance in health projects and health
  systems.

How are DALYs constructed?
DALYs have evolved over time and are still under development. The
Mark 1 version was presented in the World Bank's 1993 'World
Development Report' and the Mark 2 (superseding Mark 1) was out-
lined in the WHO's and World Bank's 'Global Burden of Disease'

series. Both versions are calculated as weighted combinations of the following four components.

Life expectancy: Where DALYS are used to measure the burden of a disease, this is based on the 'standard expected years of life lost' approach. It imposes an ideal length of life expected on each population and measures the burden of a disease in terms of life lost from that point. The idealised standard chosen was the highest national life expectancy at birth observed globally. This was in Japan, with females having a life expectancy at birth of 82.5 years and males 80 years, after accounting for the riskier behaviour of men. However, in cost-effectiveness analysis we need to measure the impact of an intervention, i.e. the different qualities and quantities of life experienced by patients with and without treatment. In those cases it is recommended that DALYs are calculated using the relevant national life expectancy, if mortality is stable and the intervention lasts no more than one year. If these conditions do not hold, then either a cohort approach or a dynamic population model will be needed.

The value of life at different ages: Each year of life lived is weighted differently according to the age of a person. The highest values are given to 20-40 year olds and the lowest values to those younger than five or older than 90. The shape of the age weighting function was chosen because it reflected two views:

- people aged 20-40 are more likely to be looking after young children or elderly relatives and therefore poor health during these years would be particularly detrimental to others;
- because people's productivity (reflected in the value of wages) tends to be higher during this age range.

The value of future time: Each year of future life is discounted at an annual rate of 3 percent so that each additional year is, from the perspective of the present, worth a little less than the preceding year. Therefore, at the age of 40, an additional year of life is worth 0.3 compared with 0.5 at the age of 20 and 0.1 at the age of 60. This pattern of values is argued to reflect individuals' preferences for benefits soon-

er rather than later, as well as the small risks of death in any particular year and the diminishing marginal utility of additional life years.

The value of avoiding disability: The impact of disease on disability is weighted between zero (representing healthiness) and 1 (representing death), which is an inversion of the scale used by quality adjusted life years (QALYs). This part of the DALY has evolved most over time. The weights for the Mark 2 version were developed using technical descriptions of 22 conditions (e.g. blindness, active psychosis). A small group of international health professionals were asked to rate and discuss the value of each condition in two ways, by considering the value of extending the lives of healthy people compared with:
- extending the life of people with the specified condition;
- curing the disability of people with the specified condition.

The median point from this group was the weight adopted for the condition. An arbitrary cutting of the 0-1 scale into seven unequal sizes was used to group the 22 conditions and then used to place several hundred other diseases into each category (for which each was assigned the mid-point value of the category). To move from these severity weights to a disability weight for a disease involved multiplying the severity weight by the proportion of incident cases of a disease expected to suffer disability at each level of severity.

How have DALYs been used?
The table below shows the four main ways in which DALYs have been used to date by the original developers and others. The main differences are that the original developers apply DALYs globally rather than locally, to a wider range of diseases, and that they are more consistent in the assumptions they use for weighting DALYs.

The most frequent use for DALYs by both groups has been as an indicator of the burden of disease. There are a few cost-effectiveness analyses and none for interventions amongst people from high-income countries. The sectoral analyses have combined national burden of disease estimates with cost-effectiveness analyses for a limited number of interventions. These have been used to plan the provision of a minimum package of essential health services in over 20 low/middle

12   **How DALYs have been used**

Type of use	By DALY developers	By others
Burden of disease	World Health Organisation/ World Bank global burden of disease exercise	National (e.g. Australia) and sub-national (e.g. regions in the UK) burdens of disease; individual diseases
Cost-effectiveness analysis	The World Bank health sector priorities review for low/middle income countries	Cost-effectiveness analysis of alternative treatments in specific low-income countries
Sectoral analysis	Mexico (country) Tanzania (district)	Turkey (country)
Planning health R&D	WHO (1996) report on investing in health R&D	NIH post-hoc justification of R&D spending

income countries and the results have often formed the basis of World Bank loans for health sector development.

What are the principal criticisms of DALYs?
Disagreement with the presence and contribution of each weighted component:
- The use of an idealised life expectancy in calculating the burden of a disease is not relevant to national planning and fosters, rather than reduces, inequity between countries. In effect, older people in richer countries are accorded more weight than younger people in poorer countries.
- The value of living a healthy life at different ages during a person's lifetime may have become confused with the value of time and the greater likelihood that older people are disabled. Therefore the impact of disability at different ages may have been double counted.
- As the developers of DALYs rejected human capital theory for valuing age, it is inconsistent of them to draw on the same theory for valuing future time. The arbitrary choice of a 3 percent annual discount rate has little empirical evidence to justify it and is unlikely to reflect globally held preferences.

- The two ways used to lead health professionals to choose weights for conditions are based on unacceptable ethical principals. Also, forcing a consistency in answers between the two approaches means that even the preferences of the few professionals selected were sometimes ignored.

  In practice one of the main problems in using DALYs is the way that analysts report results. Few report the assumptions used or test the sensitivity of results to the assumptions. This hampers comparability between studies and lends support to calls for the time and age weights to be taken out of the baseline disability weighting formula.

Disagreement about whose values are, and should be, represented:
- The limited focus of DALYs on disease and some aspects of disability, means that many benefits of interventions that improve health and welfare are not measured. It also means that the only way that co-morbidity within a person is included is by adding up weights for each disease separately. Unfortunately this means that the benefits expected from an intervention may not be realised because it is not linked to the reality of disease and illness within people.
- There has been a strong reaction against the view that people experiencing disease are unable to give a truthful account of their condition and that the biomedical models of disease and preferences of English speaking health professionals should be used exclusively to determine global resource allocation.
- The DALY approach ignores decades of research showing that the understanding and experience of disease and sickness depends fundamentally on context. Therefore disability weights are unlikely to be universal and assuming that they are they could lead to inefficient interventions being provided.
- The DALY relies on data that are known to be gender biased and much gynaecological morbidity is missing from the estimates of burden of disease.
- DALYs are not counted as lost prior to a live birth and therefore the unborn have no claim on resources (unless the health of the mother is affected). There are also concerns that DALYs may

affect the rights of a person to be born if they are known to have a higher probability of being physically or mentally challenged.

● Those with less ability to return to full health have a lower claim on resources, a group which is more likely to include the poor.

Concerns about the quality of data: The quantity of guesstimates used in calculating DALYs requires more open reporting of the sources of data and their reliability, and at least some sensitivity analysis based on the levels of uncertainty.

Are DALYs more widely applicable than EQ5D quality adjusted life years (QALYs)[1]?

DALYs have been applied more widely and more consistently across more countries and diseases than EQ5D QALYs because the original developers of the DALY have made the estimates. However, when others use DALYs there are many differences in the way they calculate them. For example, age and future time may or may not be included, different disability weights may be used, and even the standard expected years of life lost may not be used – and yet they are still called DALYs! The EQ5D QALYs have been used more widely in assessing the effectiveness and efficiency of health interventions because this tool has been operationalised for primary data collection, e.g. in clinical trials, whereas DALYs have not.

Neither of these instruments evolved from studying conceptions of health in the countries targeted for their use but instead are based on the values of the researchers who developed the instruments, with the EQ5D developers having a wider range of disciplines and countries represented. Therefore neither is likely to represent a universally held view of health, and the 'centres' of the instruments are more likely to represent views held by well-educated, wealthy, white, middle-class professionals from Europe and North America.

1 QALYs can be calculated using many methods. I chose the EQ5D approach for comparison because it one of the few generic measures based on population derived preferences for health states in relation to death and because it is the only health index developed from the beginning for use in more than one country.

Are DALYs better than EQ5D QALYs as an outcome measure in economic evaluation?

Evidence on the validity, reliability and sensitivity of the EQ5D is considerable and shows that the instrument reaches acceptable standards. There is almost no published evidence for DALYs which, coupled with a lack of openness in their design, means that it is difficult for researchers to conclude that this is an acceptable measure. EQ5D QALYs are more inclusive than DALYs because they can:

- include the impact of side-effects;
- account for the impact of co-morbidities;
- involve the patient in measuring their own health state; and
- reflect the values of the general public in particular countries rather than of a small group of 'experts'.

These criticisms could change only: if DALYs were operationalised and tested for primary data collection; if they were redesigned to measure health rather than disease; and if the reporting of methods and results were improved.

In countries where the EQ5D has not been tested, or is not considered relevant, the DALY may be the only option available and therefore used. However, the reliability and validity of the outcome measure must still be questioned and the results subjected to more rigorous sensitivity analysis than is happening currently, particularly as the disability weights have not been evaluated for any intervention and are unable to distinguish alternative interventions.

As DALYs and QALYs share the same basic idea of combining the impact of mortality and morbidity into one measure designed for resource allocation decisions, they also share (and are unlikely to shed) some of the same criticisms. For example, discrimination against the aged and those less likely to return to full health such as the permanently disabled and the poor, and that 'health' rather a broader notion of human welfare is used to allocate resources.

DALYs or QALYs for decision-making?

As DALYs only partially reflect the impact of disease on people's lives, they offer at best a limited picture of the size of the problem. They are useful if the intention is to eradicate a disease or to ascertain the scale

of resources required for palliative care. However, estimates of the burden of disease are not a sufficient basis for resource allocation, as they say nothing about how interventions are likely to reduce the problem, or about the opportunity cost of treating one problem rather than another. Therefore increasing spending in this area of research is unlikely to be an efficient use of resources for decision-makers.

Cost-effectiveness analysis can be used by decision-makers at different levels, each of which leads to different conclusions about the current relevance of using QALYs relative to DALYs.

International decision-making: DALYs offer a quick, broad-brush approach to estimating the impact of interventions for different diseases in different regions of the world based on expert views. DALYs are therefore of more interest to international decision-makers such as the WHO and World Bank. QALYs have never been used in this context and so provide no data, however rough.

National decision-making: In high-income countries, DALYs offer no advantages over the existing QALY measures. In middle-income countries where there is some development in QALY research, either they or DALYs might be relevant. In low-income countries, where the World Bank ties loans to the calculation of DALYs and where little research funding exists for the development of nationally applicable measures, it is less likely that locally relevant QALYs will be developed for decision-making.

Recommendations
1. Everyone using DALYs should evaluate the impact of assumptions on their results routinely.
2. Move DALYs away from disease to a health or welfare based measure and to a people-centred measurement and valuation process.
3. The current assumptions about the impacts of interventions on DALYs need to be explained, justified and also tested.
4. Testing the assumptions about the impacts of interventions on DALYs requires that a measurement tool is operationalised and tested for validity, reliability and equivalence across countries.

This process needs to be culturally fair for results to be meaningful to local populations and in low-income countries it is likely that external agencies will need to help fund this process.

5. In high-income countries researchers should not adopt DALYs to assess the effectiveness of health technologies.

6. For national decision-making in middle-income countries it is time to invest in the development of locally appropriate measures of health and welfare for use in economic evaluation of health technologies.

7. Measures used globally should not claim universality if the development of the measure did not access a range of local views first. If the idea of universality is considered relevant, research should begin by investigating conceptions of health to more accurately identify the 'centre' of universality prior to developing and testing tools for international measurement.

8. Stop funding more of the same burden of disease studies in low- and middle-income countries and provide more decisional studies based on the cost-effectiveness of interventions.

1 INTRODUCTION

Disability-adjusted life years (DALYs) were introduced in 1993 amid the pages and policy directives of the World Development Report 'Investing in Health' (World Bank, 1993). DALYs are a measure of life years lost from disease, adjusted for assumptions about disability as well as the impact of age and future time. They were launched to widen the measurement of disease from the presence of morbidity and mortality, usually cited by the World Bank, to include the impact on disability in a commensurable way with mortality. The World Development Report was ground breaking for two reasons:

- it was the first attempt to assess the global burden of disease by region of the world in this way; and
- it was the first attempt to bring together results from various cost-effectiveness analyses of health interventions by disease using one outcome measure (DALYs) to recommend global health policies.

This monograph outlines the construction, uses and criticisms of DALYs and hence summarises the current state of thinking about this relatively new measure. The monograph highlights the range of values underlying the development and use of the measure. I reflect on the impact that DALYs have had on the types of agendas set and decisions made about which interventions ameliorate disease in high-, middle- and low-income countries.

The integration of quantity of life with aspects of its quality in one measure is not new. Indeed DALYs have been argued to perform a similar function to quality adjusted life years (QALYs) (Bevan et al., 1999) and have even been described as a 'variant' (Williams, 1999) or 'clone' of the QALY (Bobadilla, 1996). A simplified schematic comparison of DALYs and QALYs is given in Figure 1, which shows that DALYs lost and QALYs lived could be considered the obverse of each other. Where a change in health is measured for one disease following an intervention, DALYs averted would equal QALYs gained, if quality of life were measured and valued in the same specific way by each measure.

In making a comparison between QALYs and DALYs, it is difficult to write about QALYs as if they are one measurement method. Rather, they are an approach. Examples of techniques to calculate the 'Q' part in QALYs include:

- generic pre-scored questionnaires, such as the EQ5D[2] (Brooks, 1996) or Health Utilities Index (Torrance et al., 1996), which have valuation matrices for attaching weights to different health states; or

- scenario approaches to valuing descriptions of specific disease states, e.g. using time trade-off to place a value on having a particular treatment for breast cancer (Ashby et al., 1994).

However, because of the difficulty in running time trade-off exercises, the valuation matrix approach appears to be increasingly popular for assessing the impact of medical technologies[3].

This monograph concentrates on DALYs and includes the findings of a thorough review of the literature[4]. Section 2 explains in detail

Figure 1 **Relationship between QALYs and DALYs**
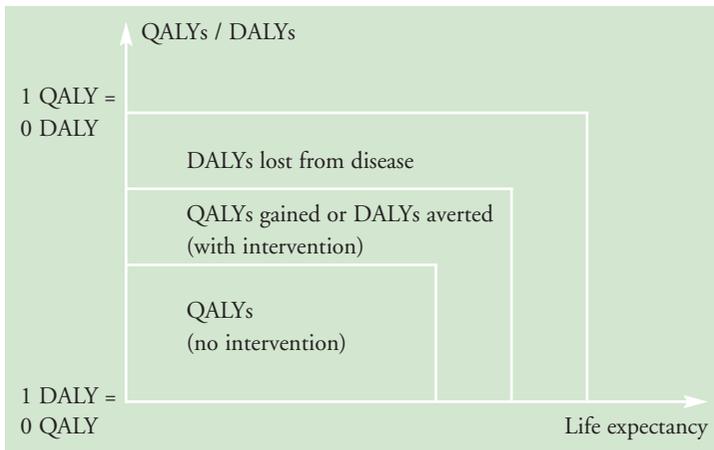
2 It is called the EQ5D because it was developed by the EuroQol group and this version had five dimensions. The nomenclature was developed to avoid confusion with earlier versions of the instrument that had six dimensions.

3 However, readers should be aware that the valuation matrix for pre-scored questionnaires for calculating the 'Q' of QALYs is based on the results of a valuation approach such as the TTO. The difference when compared with the 'scenario' approach is that: the values are not attached to any particular disease; and the values do not have to be ascertained each time a disease/health state is measured.

4 Appendix 1 outlines the search and selection criteria used for this literature review.

what DALYs are as well as how and why they are constructed. It shows that DALYs incorporate assumptions about life expectancy, time and age-group preferences within their calculation, and that the focus is on disease rather than health. It also shows the crucial role that health professionals have played in their development.

Section 3 moves on to show how DALYs have been used in practice, distinguishing between use by their original developers and by others. In a relatively short time the DALY appears to have become one of the most talked about approaches to measuring health across the globe. Indeed the 1993 World Development Report and its accompanying volume (Jamison et al., 1993) have been described as 'probably the most important document in international health that has appeared in at least a decade' (Hinman, 1997 p8). The DALY's impact on decision-making in international organisations is undisputed (Laurell and Arellano, 1996; Bobadilla and Cowley, 1995). Within a year of their introduction, DALYs had been used as the outcome measure in cost-effectiveness analyses for sector-wide decision-making in the health sectors of eight countries including Mexico, Mauritius, Turkey and Indonesia (Bobadilla and Cowley, 1995) and applied in many more countries as a measure of the burden of disease. The speed of uptake reflected:

- the need to take decisions, even when information was not available;
- the power of the agencies developing, and requiring use of, the tools;
- the paucity of experience in using or developing alternative, QALY-based instruments in low- and middle-income countries.

Early reactions to the 1993 World Development Report (World Bank, 1993) heralded important changes in thinking about international health at the World Bank (Foege, 1994) whilst at the same time reflecting views that the methods and assumptions used in DALYs were controversial (Lancet, 1993). Publication of some of the methods in 1994 (Murray and Lopez, 1994a) facilitated greater consideration in the academic press of the construction and meaning of DALYs and their use in analyses of burden of disease and cost-effectiveness by people other than the developers. In 1996 a revision to DALYs was

produced (Murray and Lopez, 1996a), which took account of some of the published discourse, and there has since been an increasing literature using and criticising DALYs as an outcome measure.

Section 4 of this monograph summarises the range of criticisms of DALYs to date. It focuses first on the four main methods of weighting designed by Murray, and progresses to outline concerns about the value systems embedded within the design of DALYs, as well as the quality and cost of existing data. The most persistent criticisms are that:

- DALYs discriminate against people with disabilities;
- the minimisation of DALYs would increase inequalities of health across populations;
- the DALY is too reductive a view of health;
- the DALY does not account for the context in which disease is experienced;
- DALYs do not reflect either individual or societal preferences in any one country;
- DALYs will not help identify the most efficient, welfare maximising interventions.

The last part of Section 4 outlines the debates raised concerning uses of DALYs in decision-making. In particular it is argued that the tool has been most useful to international organisations like the World Bank, but that using DALYs to measure the burden of disease is not helpful in helping decide what to *do* about disease. Indeed, as Williams (1999) argues, spending vast amounts on estimating the burden of disease is taking valuable research resources away from cost-effectiveness analyses that could help decision-makers decide what to do.

Like all methods for measuring and valuing the impact of treating disease, understanding of the measurement process changes over time. Section 5 briefly describes the range of on-going work and potential approaches for the future. It shows that, over time, there appears to be greater convergence between the QALY and DALY methods as DALYs adopt some of the characteristics of existing questionnaire based approaches for QALYs, e.g. from the EQ5D. Secondly there is some consideration of separating out the measurement of perceived

changes from the values attributed to those changes. However, it remains to be seen which ideas will be adopted.

The discussion in Section 6 reflects on the added value of DALYs over other approaches for calculating QALYs, specifically EQ5D-constructed QALYs (readers who would like an introduction to the EQ5D are directed to Appendix 2). The answer will depend on who is using the measure, where, for what purpose and from which viewpoint. In considering whether DALYs are more widely applicable than QALYs, I argue that both have been applied in varied ways, although the DALY has been applied in more countries. However, as DALYs have moved away from a strictly biomedical orientation, it is difficult to argue that either of them has accessed universally held views of 'health'. DALYs have only sought the views of health professionals, whereas QALYs constructed with the EQ5D have been used to measure, rather than estimate, change in health and have used random samples of the general population to provide the values of such changes to an individual.

As the DALY has not been operationalised as a primary data collection tool, it has not been used directly to measure change in 'health' and is a tool for estimation based on secondary data and a range of assumptions. Thus as an outcome measure for use in economic evaluation, DALYs are argued to provide a quick route to estimating the partial impact of a disease or health intervention. However, at present there is no way to distinguish impacts of different interventions for the same disease on the 'D' of DALYs because there are only two possible weights by disease: with or without an (unnamed) intervention. Therefore, relative to questionnaire based approaches to measuring QALYs, the DALY not only adds little but favours interventions that have more side-effects or that people find it harder to adapt to, as well as diseases that have more co-morbidities, as the benefits of such interventions are systematically overestimated. Much of this problem arises because the focus of DALYs is on disease rather than health or wellbeing.

In this monograph I consider the advantages of DALYs relative to QALYs in the context of high-, middle- and low-income countries, as well as at the international, national and sub-national levels of decision

making. I argue that decision-makers in low- and middle-income countries and at the international level are more likely to use DALYs rather than QALYs because:

- DALYs are required to be used in order to agree health sector loans with the World Bank;
- QALY approaches have hardly been used in these settings so there are few alternative measures to DALYs; and because
- there has been little investment in research to develop locally-specific, health-related quality of life measures in these settings.

In high-income countries, especially those in which health-related quality of life measures such as QALYs have been developed, there is much less need for DALYs, although there is still an onus to develop understanding of QALY gains for a wider range of interventions across a wider range of disease areas to help sectoral decision-making.

The conclusions presented in Section 7 focus on the achievements of DALYs to date with respect to their four principal uses as:

- an indicator of the burden of disease;
- a director of research and development;
- a measure of effectiveness; and
- a unit of benefit in cost-utility analysis for micro and sectoral evaluations.

I conclude that the use of DALYs in assessing the global burden of disease has helped people think about the distribution of disease based research funding across the globe. The main concerns raised are about the quality of data input, the method of dealing with co-morbidity, that the context of disease experience is ignored and the implied equity weights emanating from the assumptions. With respect to the use of DALYs as an indicator of burden and measure of effectiveness or benefit, I suggest that it would help clarify matters to separate actual change in quantities from the values attached to the changing quantities. This would clarify the impact of weighting for age and time preference; the impact of assuming 'ideal' life spans; and the desired distributions of benefits to 'healthy' people and people who are disabled (i.e. suffering from any disease). It would also allow alternative value functions to be considered, and the impact of judgements to be understood.

24

I also question the usefulness of DALYs that focus on disease rather than people, because it is people who experience disease, sickness and interventions that change disease. Useful decisional information needs to incorporate both perspectives. However, to move to maximising QALYs, rather than minimising DALYs, may still lead to a sub-optimal allocation of resources, even from the viewpoint of efficiency, as the desired objective function may be broader than or different from what QALYs are measuring. The likelihood of this may well increase with the number of countries included, as views about the constituents of health and well-being have not yet been investigated in the process of developing internationally used measures of 'health'. Existing measures will therefore increasingly impose values on, rather than reflect values of, new populations.

Finally, I also highlight the disparity between the demand for and supply of decisional information for health sector prioritisation by comparing the large and increasing quantity of data produced on burden of disease globally with the fact that only 107 economic evaluations were published between 1984-97 that had anything to do with low- or middle-income countries. An average of one publication per country is hardly a sufficient basis for global policy making. Thus, if cost-effectiveness analysis is to be used, far more resources should be targeted to the use of DALYs in economic evaluation rather than to burden of disease exercises.

2 THE CONSTRUCTION OF DALYS

2.1 What is the DALY?

25

DALYs are the present value of future years of lifetime lost through premature mortality, plus the present value of the adjustment to years of future lifetime to allow for the average severity (frequency and intensity) of any mental or physical disability caused by a disease or injury. This shows that DALYs are a measure of something lost rather than something gained, so that DALYs lost are not desired themselves, but rather their reduction. 'Popular' explanations have even described DALYs lost as 'the enemy' (McNamee citing Feacham, 1993, p105).

2.2 The founding principles of DALYs

In 1994, Murray outlined four principles on which he argued DALYs were based (Murray, 1994, p4-5):

- to the extent possible, any health outcome that represents a loss of welfare should be included;
- the characteristics of the individual affected by a health outcome that should be considered in calculating the burden of disease should be restricted to age and sex, being the only two characteristics which mean the same in all cultures;
- the measure treats like health outcomes as like;
- time is the unit of measure for the burden of disease.

However, whilst not counted among the founding principles, it is clear that four other views have also dominated thinking:

- a desire to 'decouple epidemiological assessment from advocacy so that estimates of the mortality or disability from a condition are developed as objectively as possible' (Murray, 1996, pp1-2). This resulted from research showing that claims concerning adult and child mortality by disease when summed exceeded total deaths in each age group by 200-300 percent (Murray and Lopez, 1994a);
- that 'it is better to make an informed estimate of disability flowing from a particular condition than to have no estimate at all' (Murray and Lopez, 1994b, p55), to ensure that under- or unresearched diseases do not drop off research or policy agendas;
- that 'health research and action agendas are global not national'

and that 'the health profiles and the agendas for health research and health policy for the developing world and the developed world are converging' (Murray, 1998). Both beliefs underpin views that facilitate the construction of global outcome measures and global policy agendas;

- that international agencies have a critical role in influencing national health policies in developing countries, in terms of: aiding the expensive and time intensive process of priority setting; allaying 'the problems of inertia'; and in facilitating long-term commitments to research (Murray, 1990, p310).

2.3 The architecture of DALYs

DALYs have evolved over time and there are now two markedly different versions. Whilst what I shall call the DALY Mark 2 version has superseded the Mark 1, both are referenced and used in the literature and a Mark 3 version being developed appears to draw on some of the ideas of the original version. Therefore, I will draw out comparisons of the Mark 1 and Mark 2 versions in four ways, through which the construction of DALYs can be viewed. This section will show in detail how the DALY is constructed for a single individual of given age and sex. Individual DALYs then have to be summed up for the whole population to calculate total DALY burdens or treatment impacts.

The four pillars of DALYs involve different methods of weighting for:
- life expectancy;
- age;
- future time;
- disability.

These are discussed in turn in the remainder of this section of the monograph.

2.3.1 Life expectancy
The method for calculating life expectancy in burden of disease estimates is based on the 'standard expected years of life lost' method. This imposes an 'ideal' length of life expected on each population and measures the burden of a disease in terms of life lost relative to that.

The idealised standard chosen was the highest national life expectancy observed globally. This was in Japan, where females were shown to have a life expectancy at birth of 82.5 years and males 80 years[5]. The difference was estimated as the biological difference in gender, after accounting for men behaving badly[6].

The main reason for adopting a standardised life expectancy across the world was to avoid the problem of valuing deaths differently across the world. If each country's life expectancies for current birth cohorts were used as the base for calculations of disease burden, a 40 year old's death in Nairobi, Kenya would be worth less (in terms of DALYs lost) than a 40 year old's death in London, because life expectancy in England is higher (Murray, 1994, p7). The implication for basing allocation of World Bank resources on burden of disease calculations in this simple example would be to give money to the English and not the Kenyan person, which was not desired. The 'standard expected years of life lost' approach was preferred[7] for measuring the burden of disease because it combined the cohort expectation approach in treatment of deaths at older ages with the egalitarian nature of potential years of life[8]. The method of calculating life expectancy remained the same for Mark 1 and

5 The standard life expectancies at each age are given in Murray (1996, p17). These life expectancies were chosen because of the need to make international comparisons to aid the World Bank's decisions about where to lend money for reforming health sectors.
6 Including, for example, men smoking and drinking more than women as well as engaging in more risky behaviours and taking more hazardous employment (Murray, 1996, p17-18).
7 Compared with three alternative approaches: potential years of life lost, period expected years of life lost and cohort expected years of life lost.
8 This allows life expectancies for each age cohort to be calculated and is particularly useful where age specific mortality rates are changing (mostly mortality is falling over time, although more recently mortality in some sub-Saharan Africa countries has increased, attributed to HIV/AIDS). However mortality has to be estimated as we do not know the future, and a range of projection models exist to do this. The problem with period expected life lost is that current mortality rates are applied as the 'ideal' or highest possible age of death and where mortality changes over time this can lead to odd results as reference standards change (Murray, 1996, p13). The difference between period expected and cohort expected years of life lost on the one hand and potential years of life lost on the other is that in the former more weight is given to deaths at older ages because expectation of remaining life does not fall to an arbitrary zero at a particular age (using standard expected years of life lost it falls to zero at 82.5 for women). A full discussion of the arguments is given in Murray (1996, pp10-16).

Mark 2 DALYs. The tables of life expectancy at each age used to calculate DALYs are reproduced in Murray (1994, p9; and 1996, p17).

2.3.2 Weighting age

Age weights were introduced into DALYs to reflect beliefs that years of life vary in value depending on an individual's age. Two justifications have been given (Murray, 1994 and 1996). First, it was argued that human capital theory allows the value of individuals to be calculated in the same way as the value of machines and that measures of productivity, including wages, can be a good approximation of this. Such indicators show an inverted U-shaped relationship of value across time. Secondly, age weights attempt to capture the belief that social roles, and in particular the level of 'welfare inter-dependence', vary with age. For example, people between the ages of 20 and 50 are more likely to have young children or elderly relatives dependent on them and therefore it can be judged that it is more important to be alive and healthy during these times than before or after. Latterly Murray (1996, p56) cited evidence from Bussbach et al. (1993) that students and elderly people prefer to be healthy at particular ages.

A continuous age weighting function was chosen for both Mark 1 and Mark 2 DALYs, to ease the computational burden that would have accompanied a discrete weighting system, of the form:

$$C \, xe^{-\beta x}$$

where:

ß and C = constants

x = age in years

e = base for natural logarithm (approx 2.718)

The shape of the age function is given in Figure 2, when C = 0.1658 and ß = 0.04. It can be seen that the relative value rises until around the age of 25 and then falls[9]. This suggests that adding 10 years of life to a 10 year old is worth more than adding 10 years of life to a 50 year old.

9 Note that Barendregt et al. (1996, p441) showed that the combined effect of discounting and age weighting extended the 'hump' of the graph in Figure 2 and therefore added further value to life years in age groups 27-38.

Figure 2 **Relative value of a year of life lived at different ages incorporated into DALYs**
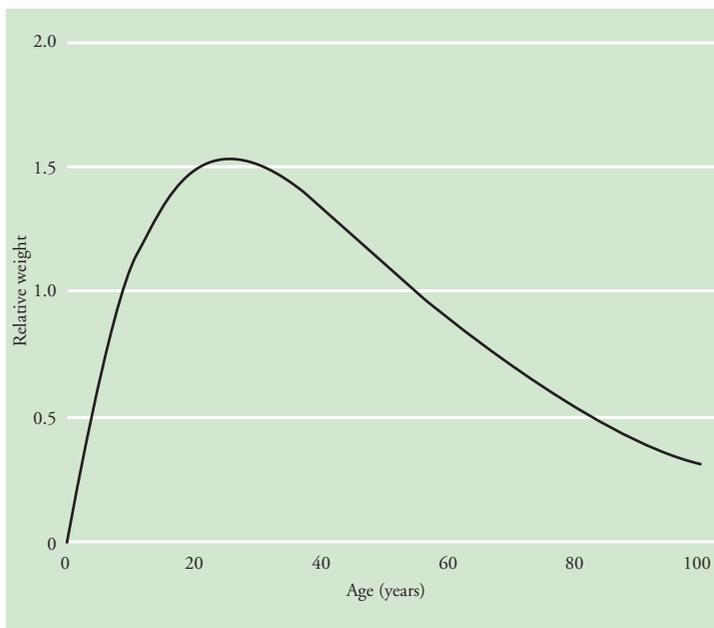
Source: Murray (1996, p60).

The only difference between Mark 1 and Mark 2 DALYs is the size of the constant *C*, which equals 0.16523 in Mark 1 DALYs and 0.1658 in Mark 2 DALYs. The purpose of the constant is to ensure that the total number of DALYs lost without age weighting equals the total number of DALYs lost after accounting for age weights.[10] The

10 As Murray (1994, p10) states, the size of the constant would change if either the age weighting function changed (e.g. by changing ß) or if the age and sex pattern of the global burden of disease in real populations changed. Therefore if the burden of disease was calculated at a different point in time or for a specific country, the size of the constant should change. This quickly becomes a complex exercise and WHO has adopted the pragmatic approach of using 0.1658 in all country and regional estimates of the burden of disease and for cost-effectiveness analysis (Evans D, personal communication). This means that total age-weighted DALYs will not equal total non-age weighted DALYs in such cases.

choice of value for ß (which controls the shape of the age weighting function) as 0.04 was made on the basis of:

- 'a modified Delphi method'[11] with a group of public health 'experts';
- the judgement that 'only a narrow range of ß provides reasonable age patterns between 0.03 and 0.04';
- the conclusion that 'results are largely insensitive to the specific ß chosen but are sensitive to the difference between equal and unequal weights';
- 'an informal polling of the advisory board' (all quotes are from Murray, 1996, p10).

2.3.3 Weighting future time

The concept of pure time preference is drawn on to justify reducing the weight of successive increments of future time; i.e. an individual would prefer a benefit today rather than in one year's time (assuming there is no risk or uncertainty[12]). However, in addition to this, Murray (1996) also draws on two further reasons why individuals are likely to have a positive rate of time preference: a 1 percent annual risk of personal death; and a diminishing marginal utility to additional life years. The latter argument is linked to additional life years increasing the capacity of an individual to access utility through consumption.

The arguments concerning moving from individual to social rates of time preference are premised on similar bases. Pure time preference is discussed with reference to the possibility that one generation may be more concerned with the welfare of future generations than an individual is in their own private day to day transactions; and that groups of individuals today may be more willing to save if they join together

11 The Delphi technique is an approach used for establishing and developing consensus. It is usually based on experts being sent a self-completion questionnaire, analysing the responses, feeding back a summary of the group's views and asking respondents to re-evaluate their own views given the results. If a substantial amount of disagreement remains (e.g. rankings are still disparate) a further round of feedback may ensue (Bowling, 1997, pp362-3).

12 When this assumption is relaxed, the effect is to decrease or increase the rate of time preference depending on whether future income becomes more or less valuable relative to current income, which depends on whether marginal utility is convex or concave.

rather than if they remain on their own. Risks of death, rather than being attributed to an individual, are considered in relation to societies as a whole with respect to extinction of an individual society (through war or natural disaster) and of the human race (through a cataclysmic event) – both of which are considered to be much smaller than 1 percent per year but still with a positive risk. The same argument concerning expectations of future economic growth is given. However, Murray outlines an additional argument called the 'disease eradication paradox'. This is based on the case where a disease is conceived as having a stable and infinite incidence and the possibility of an intervention to eradicate it. He argues 'society would necessarily redirect all its resources to disease eradication programmes and reduce all other investments in all curative and preventive programmes to zero' and that, therefore, a discount rate high enough to counteract such excessive sacrifice by current generations is needed.

In choosing the rate at which to discount, Murray writes about his struggle between, on the one hand, using a positive (albeit low) discount rate to capture the uncertainty that increases with time and the need to reduce problems of excessive sacrifice and, on the other hand, using a zero discount rate that ignores the cumulative effect of small future uncertainties (with an option of counteracting the excessive sacrifices through arbitrary delineations of budgets for current and future generations). A positive discount rate of 3 percent per year is used in both Mark 1 and Mark 2 DALYs, a rate which Murray felt was likely to represent 'the lower limit of acceptability for those economists who are persuaded by opportunity cost arguments… and the upper limit for public health practitioners who are willing to accept a positive rate of discount' (Murray, 1996, p54)[13]. The discount rate was introduced into the DALY formula for both Mark 1 and Mark 2 as a continuous discounting function of the form:

13 In a final section, Murray (1996, p54) also raises the possibility of discounting the expected benefits of preventive interventions differently for communicable and non-communicable diseases, suggesting that the discount rate in the latter case should perhaps be lower. The main argument propounded is that the annual decline of communicable disease has been much faster and therefore the future benefits from prevention are likely to be smaller and more uncertain.

e^{-rt}

where:

r = discount rate expressed as a decimal (e.g. 0.03)
e = base for natural logarithm
t = time

In estimating the burden of disease, DALYs Mark 2 are presented with a discount rate of 0 percent as well as 3 percent.

2.3.4 Weighting disability

The broad approach adopted was to weight the impact of disease between zero (representing healthiness) and 1 (representing death)[14]. It is this area that has been subject to greatest change between the Mark 1 and Mark 2 DALYs.

Whilst both Mark 1 and Mark 2 DALYs used the same name, '*disability* adjusted life years', and the conception of disability is in both cases argued to arise largely from the work of the International Classification of Impairments, Disabilities and Handicaps (ICIDH)[15], what is covered has changed over time. Thus, in 1994, Murray wrote 'we have chosen to measure disability, not handicap[16]' (p11), whilst for the Mark 2 DALYs he stated that the underlying con-

---

14 Note that this is an inversion of the quality adjusted life year (QALY) type scales for which zero represents death and 1 healthiness.

15 Where impairment is defined as 'any loss or abnormality of psychological, physiological, or anatomical structure or function'. 'Disability has to do with compound or integrated activities expected of the person or body as a whole, such as are represented by tasks, skills and behaviour. It is defined as any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being. Handicap relates to the social consequences of deficiencies in organs and activity performance. It is defined as a disadvantage for a given individual, resulting from an impairment or a disability that limits or prevents the fulfilment of a role that is normal (depending on age, sex and social and cultural factors) for that individual' (Reynolds Whyte and Ingstad, 1995, pp6-7).

16 Disability in the ICIDH (WHO, 1980) concerns the loss of functional capacity emanating from an impaired organ (mental, sensory, internal or externally visible body), whereas handicap concerns the social and cultural consequences of a disability (or impairment to an organ).

Table 1 **Classifications of disability used in Mark 1 DALYs**

Class 1	Limited* ability to perform at least one activity in one of the following areas: recreation, education, procreation or occupation
Class 2	Limited* ability to perform most activities in one of the following areas: recreation, education, procreation or occupation
Class 3	Limited* ability to perform activities in two or more of the following areas: recreation, education, procreation or occupation
Class 4	Limited* ability to perform most activities in all of the following areas: recreation, education, procreation or occupation
Class 5	Needs assistance with instrumental activities of daily living such as meal preparation, shopping or housework
Class 6	Needs assistance with activities of daily living such as eating, personal hygiene or toilet use

Note: *Limited ability was arbitrarily defined as 50 percent of full activity.

Source: Murray (1994, p12).

struct 'lies somewhere between disability and handicap' and 'is best described as an average level of handicap'[17] (Murray, 1996, p34).

One of the most significant changes between DALY Marks 1 and 2 is the way in which the impact of disease on disability is described. Table 1 shows the classifications of disability used originally. It can be seen that the top four classes are measured in terms of ability to function for work, leisure, procreation and education. The bottom two classes measured quite different abilities involving managing aspects of a home environment and basic activities of daily living. Mark 2 DALYs moved away from this to specific definitions of individual disease states and conditions. Boxes 1 and 2 show the range and operationalisation of disease states and conditions considered. The main

17 This is an interesting claim, as Murray previously adopted the position that disability and not handicap allowed 'like outcomes to be treated as like'. Using handicap, it was argued, would exacerbate inequalities as different expectations are held in different social contexts (Murray, 1994, p12).

Box 1 **The 'indicator conditions' chosen to reflect different dimensions of non-fatal outcomes in Mark 2 DALYs***

- Largely physical manifestations
  - e.g. blindness, deafness, below-the knee amputation

- Neuro-psychiatric conditions
  - e.g. unipolar major depression, active psychosis, Down's syndrome without cardiac malformation, mild mental retardation

- A condition with social/group interaction consequences
  - e.g. vitiligo on face

- Pain
  - e.g. severe migraine, angina, severe sore throat

- Sexual/reproductive functioning
  - e.g. erectile dysfunction, infertility, recto-vaginal fistula

Note: *The remaining conditions, not categorised in the above manner, include paraplegia, quadriplegia, severe anaemia, fracture of radius in a stiff cast, dementia, watery diarrhoea, two standard deviations below weight-for-height, rheumatoid arthritis.

Source: Murray (1996).

reason given for the change in descriptions was that the original, Mark 1, classification was not appropriate for children as they would, by definition, be considered disabled because of their dependency on others. Other than considering the literature, no description of the processes used to arrive at the descriptions of disease or health states for either version of DALYs is given, except that definitions of the indicator conditions used in constructing Mark 2 DALYs were revised following testing by 'students and fellows from a number of countries at the Harvard School of Public Health' (Murray, 1996, p36).

The methods for determining weights for DALYs Mark 1 and Mark 2 are also very different, although the idea that the severity weight is a combination of the perceived intensity of disability and the

Box 2 **Examples of the descriptions given for 'indicator conditions'**

> **Blindness**
> Maximum visual acuity with the best possible correction is less that 3/60. (A person is unable to distinguish the fingers of a hand at 3 metres, or has less than 5 per cent of remaining vision as compared to a normally sighted individual).
>
> **Dementia**
> An individual with multiple cognitive deficits that include memory impairment and aphasia (difficulty producing the names of individuals and objects) and apraxia (impaired ability to execute motor activities despite intact motor abilities, sensory function, and comprehension of the required task).

Source: These definitions and all others used are given in Murray (1996, pp94-96).

time spent in the disabled condition remains the same in both versions. Little explanation was given about the processes involved for the Mark 1 DALY. We are told that weights were chosen 'based on the word definitions and the set of disabling sequelae in each class. *De facto*, they used a magnitude estimation method[18] to choose a number between 0 and 1 for each of the six classes. Their votes were averaged to generate the final class weights'. The values were given by a group of 'independent 'experts" (public health practitioners) 'who had not been involved in the estimation of the incidence, duration or mortality of any disease' (Murray, 1994, p13).

A protocol was developed for the process of eliciting values for the indicator conditions of Mark 2 DALYs (see Murray, 1996, pp90-96). The person trade-off (PTO) approach was used in a deliberative group process, based on two thought experiments, which are described in Box 3. Each individual within the group was asked:

18 Later Murray states this was a rating scale (Murray, 1996, p34).

- from the viewpoint of a decision-maker purchasing interventions and for each indicator condition[19], to make two individual choices: to trade quantity of life for healthy or disabled individuals (PTO1) and to trade off quality of life for improvements in 'healthy' versus disabled individuals (PTO2);

- to reflect on any differences between their own values (including a rank ordering exercise) and to arrive at one internally consistent weight to reconcile any differences in the disability weight implied by the two PTO questions. To do this, each participant was given a 'PTO1-PTO2 equivalence table' (Murray, 1996, pp97-8). This allowed the experts to make their decisions based on Murray's expectations of the relationship between PTO1 and PTO2;

- to participate in a group discussion which focused on who had the highest and lowest values in the group for each condition for PTO1 and PTO2, and why;

- to reconsider their own values in private and make changes if they wished to.

The final value selected for the group composite disability weight was the median PTO value for each condition[20]. Murray and Lopez (1997a, p379) reported that the correlation coefficients between various group exercises exceeded 0.9, indicating a high degree of agreement between them.

The group of people providing the disability weights for the DALY Mark 2 had been invited to an 'international consensus meeting on disability weights' which was convened at the WHO in Geneva in August 1995. The group included individuals from each region of the world, of whom 60 percent were male and 40 percent female[21]. They were also described as health providers from 'diverse cultural backgrounds' (Murray, 1996, pp38-39) with knowledge of, but not suffering from, any of the conditions (Murray, 1996, p93). The critical role

19 In considering each indicator condition, raters were asked to 'evaluate the average individual with the condition described taking into account the average social response or milieu' (Murray, 1996, p38).

20 Unfortunately no data have been published on the mean, standard deviation or any other information to describe the distribution of values.

21 We are not told the total number of people, although we are told that the process was designed for a group of 8-12 health providers (Murray, 1996, p36).

Box 3 **Two types of thought experiment given to raters of the**
'indicator conditions' used to elicit disability weights for Mark 2
DALYs

THOUGHT EXPERIMENT 1 (PTO1)

You are a decision maker who has only enough money to buy one of two
mutually exclusive health interventions.

Either
    Choose intervention A which will:
    Extend the life of 1,000 healthy individuals for exactly one year, at
    which point they will all die.

    If you do not purchase this, they will all die today.

Or
    Choose intervention B which will:
    Extend the life of n ($\geq 1,000$) *blind* individuals for one year.

    If you do not purchase this, they will all die today.

THOUGHT EXPERIMENT 2 (PTO2)

You are a decision maker who has only enough money to buy one of two
mutually exclusive health interventions.

Either
    Choose intervention A which will:
    Extend the life of 1,000 healthy individuals for exactly one year, at
    which point they will all die.

    If you do not purchase this, they will all die today.

Or
    Choose intervention B which will:
    Cure the disability of n *blind* individuals, who will live exactly one year
    with or without the intervention.

    With the intervention they will live in perfect health, without the
    intervention they will continue to live for one year with the given dis-
    abling condition.

Source: Adapted from Murray (1996, p91).

of the facilitator at the consensus meeting was highlighted, in terms of needing to encourage participants constantly to take the exercise seriously and carefully, as well as to challenge individuals about their chosen values, in an attempt to elicit reliable and valid assessments.

Following the derivation of weights for each indicator condition by the above group, the spectrum of disease states or conditions between death and no disability or disease was divided arbitrarily into seven disability classes (Murray, 1996, p40). The weights were divided, as shown in Table 2, and disability weights for each disability class were fixed at the midpoint. Next, the individual disability classes and the 22 'benchmark' indicator conditions were used in the mapping of disease states through their disabling sequelae for several hundred

Table 2 **Revised disability classes and weights (DALYs Mark 2)**

Disability class	Severity weights	Mid-point*	Indicator conditions
1	0.00-0.02	0.01	Vitiligo on face, weight-for-height less than two standard deviations
2	0.02-0.12	0.07	Watery diarrhoea, severe sore throat, severe anaemia
3	0.12-0.24	0.18	Radius fracture in a stiff cast, infertility, erectile dysfunction, rheumatoid arthritis, angina
4	0.24-0.36	0.30	Below-the-knee amputation, deafness
5	0.36-0.50	0.43	Recto-vaginal fistula, mild mental retardation, Down's Syndrome
6	0.50-0.70	0.60	Uni-polar major depression, blindness, paraplegia
7	0.70-1.00	0.85	Active psychosis, dementia, severe migraine, quadriplegia

Note: *Weight applied by class to distribution of disability in the global burden of disease exercise, version 5.

Source: Adapted from Murray (1996, pp40-41).

other 'conditions', which did not have individual descriptions.
    The same group that provided the weights for the indicator con-
ditions was also asked to use a rating scale to decide the distribution
of people with each 'condition' in its treated and untreated form by
broad disability class and, once again, views were amended following
a group discussion. In this exercise, many pieces of information were
considered: 'where treatment was judged to change the distribution of
severity by class, and not simply the incidence, duration or case-fatal-
ity rate of a condition, the group developed a separate distribution for
the treated form of a condition' (Murray, 1996, p41). A single dis-
ability weight for each condition was constructed by multiplying the
per cent in each class by the mid-point severity weight for the range
defining that class. An illustrative example is given in Table 3. The

Table 3 **Estimated proportion of incident cases developing a
disability and the distribution of these disabilities by severity
class: an illustrative example in Latin America and the Caribbean**

Age group (years)	Percentage of incident cases developing a disability	Percent distribution by severity class[a] of those developing a disability						
		1 0.01[b]	2 0.07[b]	3 0.18[b]	4 0.30[b]	5 0.43[b]	6 0.60[b]	7 0.85[b]
0-4	100	35	15	15	15	10	5	5
5-14	100	35	15	15	15	10	5	5
15-44	100	35	15	15	15	10	5	5
45-59	100	35	15	15	15	10	5	5
60+	100	35	15	15	15	10	5	5

Notes: [a] Distributions across the seven disability classes sum to 100 percent.
Data on these distributions may come from a variety of sources in a specific
country, and will vary by location and condition. If not available, they
would have been estimated by the expert group.
[b] Weight for time spent in each disability class (the mid-point of the disabili-
ty class).

Source: Based on an example in Murray and Lopez (1994b, p59).

Box 4 **Summary of steps used to weight disease states in Mark 2 DALYs**

1. Descriptions of 22 indicator conditions devised.

2. Weighting of named indicator conditions using two person trade-off (PTO) questions with a small group (n<13) of international health providers.

3. Each provider required to make values from PTO exercises consistent with expectations.

4. Values of individual experts shown to group, with discussion of reasons for highest and lowest values.

5. Chance for experts to re-assign their individual values of indicator conditions.

6. Median value of group assigned to each indicator condition, and all indicator conditions ordered on a scale between 0 and 1.

7. Arbitrary slicing of the 0-1 disability scale into seven categories (each category containing between 2-5 of the 22 indicator conditions).

8. Seven category rating scale was used, by the same small group of international health providers, to agree (following group discussion) the distribution across the seven categories of each condition in treated and untreated groups of people for:
   - the indicator conditions;
   - 'several hundred other conditions'[a].

9. All conditions should have had separate information on the duration of time spent in each disability class for each condition that was used to calculate the final weight[b].

10. Distributions of a condition were converted to a single disability weight by multiplying the percentage in each part of the category scale by the mid-point severity weight for that classification for the relevant distribution range decided for that condition (treated and untreated).

Notes: [a] Values for a total of 237 sequelae are given in Annex 3 of Murray and Lopez (1996a, pp412-416). Murray and Lopez (1996, p202) also point out that the disabling sequelae were selected in consultation with disease 'experts' who collaborated on the global burden of disease exercise.
[b] At times steps 8 and 9 were done together in a composite step by the group of raters when information was not available.

implied disability weight in this example, assuming equal time was spent in each disability class, would be 0.2015[22]. However, given the gaps in knowledge of the distribution of some conditions by severity of disability, the group often had to make a combined judgement concerning the distribution and time spent with each form of the condition (Murray, 1996, p41). Box 4 summarises the steps used to weight disease states in Mark 2 DALYs.

2.4 How to calculate Mark 2 DALYs

DALYs for a specific condition are the sum of years of life lost (YLL) plus years of life lived with disability (YLD):

DALYs [r, K, ß] = YLLs + YLDs

where:

r = discount rate expressed as a decimal

K = age weighting modulation factor

ß = parameter from the age weighting function

To reflect the base case recommended and used by Murray and Lopez (1996a,b) this would mean that $r = 0.03$, $K = 1$ and ß $= 0.04$. The notation DALYs [0.03, 1, 0.04] provides a quick way for evaluators to highlight any key changes in the assumptions of the DALYs they calculate.

The method used to calculate DALYs differs according to the purpose of analysis. The following two sections exemplify these differences by first showing the calculation of DALYs when estimating the burden of disease and then showing the different DALY calculation method that is used in cost-effectiveness analysis of a health intervention. Both examples focus on calculating DALYs for one individual.

2.4.1 Estimating the burden of one case of disease using DALYs

The relevant formulas were outlined by Murray (1996, pp65-6) and are shown in Appendix 3. Homedes (1995) usefully exemplified the

---

22 Note therefore that, with the same assumptions about duration, the maximum disability weight is 1.0x0.85 and the lowest is 1.0x0.01.

calculation of Mark 1 DALYs for an individual, using four illustrative case studies. Using a similar approach but a different example, I show how Mark 2 DALYs can be calculated for an individual.

Imagine that a woman becomes sick with bipolar depression at age 35, has it for 10 years and dies as a result. How many DALYs [0.03, 1, 0.04] have been lost? To calculate this we need to estimate the numbers of DALYs lost due to both her premature mortality, YLLs, and the disability she experiences while suffering from the condition, YLDs, and then sum them. The expected time horizon for this woman is shown below, by age:

Time horizon without treatment:

Depression	Death
Age 35	45

The calculation of YLDs focuses on the period during which she is living with depression, from 35-45 years. By substituting the following values into the general equation shown in Appendix 3, the YLD based DALYs can be calculated:

$K = 1$
$C = 0.1658$
$r = 0.03$
$a = 35$
$ß = 0.04$
$L = 10$
$D = 0.6$ (from Murray and Lopez, 1996a, p415)

The number of DALYs lost due to disability[23] from age 35 is then 6.95.

The calculation of the YLLs requires two steps, in this example. First to calculate the life lost from age 45 onwards, and secondly to

---

23 The formula for this specific example in Excel for YLDs is
=0.6*(1*0.1658*EXP(0.03*35)/(0.03+0.04)^2)*(EXP(-1*(0.03+0.04)*(10+35))*(-(0.03+0.04)*(10+35)-1)-EXP(-1*(0.03+0.04)*35)*(-(0.03+0.04)*35-1))+((1-1)/0.03)*(1-EXP(-1*0.03*10)).

discount this value to age 35 so that all DALYs can be added up from age 35 onwards (if the woman had died immediately at age 35, this latter step would not be necessary). The first step requires a straight-forward substitution of values into the general YLL equation shown in Appendix 3 using the following values:

$K = 1$
$C = 0.1658$
$r = 0.03$
$a = 45$
$ß = 0.04$
$L = 38.72$ (the life expectancy at age 45, using the tables provided in Murray, 1996)

The DALYs lost from the age of 45 onwards, due to lost years of life[24] are then 20.66. The conversion of this to DALYs calculated at age 35 uses the following formula:

DALYs at age x = (DALYs at age $x+s$)e^{-rs}
where:
s = number of years to be discounted
r = discount rate expressed as a decimal

Therefore, substituting in the appropriate values:

DALYs at age 35 = 20.66 x $e^{(-0.03 \times (45-35))}$ = 15.31

Therefore, from the time of onset of disease at age 35, the total number of YLLs lost due to premature death equals 15.31. Adding this to the years of life lost with disability (YLDs = 6.95), gives the total number of DALYs lost of 22.26.

---

24 The formula in Excel for YLLs in this example is
=(1*0.1658*EXP(0.03*45)/(0.03+0.04)^2)*(EXP(-1*(0.03+0.04)*(38.72+45))*(-(0.03+0.04)*(38.72+45)-1)-EXP(-1*(0.03+0.04)*45)*(-(0.03+0.04)*45-1))+((1-1)/0.03)*(1-EXP(-1*0.03*38.72)).

To move from this calculation for an individual to the total number of DALYs lost in a community due to bipolar depression would require calculating the DALYs lost to each individual with bipolar depression, using the approach shown above, and then adding them all up. With full knowledge, this would give the most detailed figures. However, in reality the calculations tend to work from a population level down, given assumptions e.g. about proportions of the population with the disease.

2.4.2 Estimating the DALYs averted for one case for use in cost-effectiveness analysis

It is important to note that, for calculating DALYs averted as part of a cost-effectiveness analysis, the 'standard expected years of life lost' approach is not used. What is needed is the difference between years lived with and without the intervention being studied. Standard expected years of life lost measures neither. Therefore, in cost-effectiveness studies the relevant national life expectancy is recommended as a good approximation of life expectancy, provided that mortality is stable and interventions have a duration of less than one year (Murray, 1996, p20, citing Preston, 1993).

If, however, underlying mortality is changing over time, each new birth cohort will experience a different life expectancy and the current local life expectancy becomes a less accurate representation of future life for interventions that impact on particular age groups. Therefore a cohort life expectancy is recommended as the basis for estimating change with and without an intervention (Murray, 1996, p20). However, the situation quickly becomes more complex if the evaluation considers an intervention lasting several years that changes age-specific mortality rates. Not only is a great deal more information required (age specific mortality rates and the impact of the interventions on mortality at each age) but a move to a population model is also required to capture the dynamic nature of the impact of the intervention.

For ease I will use the same example as above. The first step is to use local life tables rather than the standard expected life tables to calculate YLLs without treatment (in this example the estimation of YLDs will be the same as above because the 10 years alive with dis-

ability is a given). The YLLs will therefore differ by country/life table used. In this example I use Chile, where the remaining life expectancy for women at age 35 is 44.13 years and at 45 is 34.73 years[25]. I also use WHO's approach and assume the Global Burden of Disease value of the constant (see footnote 10). The YLLs at age 45 are 19.97, which when discounted to age 35 values becomes 14.80 YLLs. Therefore the total DALYs lost as at age 35 without treatment are 6.95 (YLDs) plus 14.80 (YLLs), which equals a total of 21.75 DALYs[26].

Now imagine if this woman received treatment for her depression at age 35, and that she does not die at age 45 but lives for her expected life span, in the treated state. Her new time line is shown below.

Time horizon with treatment:

Depression treated   Death

Age 35   79.13
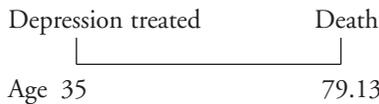
How many DALYs would be averted by the intervention? A DALY weight for the treated form of the disease is given in Murray and Lopez (1996b, p415)[27] as 0.302 (a fall of 0.298 from the untreated DALY weight of 0.6, as used in the example in Section 2.4.1). Using the YLD formula from Appendix 3 with the values: $K = 1$; $C = 0.1658$; $r = 0.03$; $a = 35$; $ß = 0.04$; $L = 44.13$; $D = 0.302$; the DALYs now associated with her condition[28] equal 7.94. This means the total number of DALYs [0.03, 1, 0.04] averted following treatment of the woman is 21.75 – 7.94 = 13.81. As Fox-Rushby and Hanson (2001)

25 These are the female life expectancies at ages 35 and 45 respectively using the Chilean life tables from the United Nations, 1982 Model Life Tables.
26 See Fox-Rushby and Hanson (2001) for full calculations.
27 Note that whilst we are given this value we do not know what the treatment is. Also, the value is meant to represent the 'average' response of people to treatment of bipolar depression. In an average, some people will be better, and may have a 100 percent recovery with no need for further treatment and no remission. Others may not recover at all following treatment.
28 Using the Excel formula: =0.302*(1*0.1658*EXP(0.03*35)/(0.03+0.04)^2)*(EXP(-1*(0.03+0.04)*(44.13+35)))*(-(0.03+0.04)*(44.13+35)-1)-EXP(-1*(0.03+0.04)*35)*(-(0.03+0.04)*35-1))+((1-1)/0.03)*(1-EXP(-1*0.03*44.13)).

show, changing the age and discount weights would alter the DALYs averted from treatment at age 35 as follows:

DALYs averted [0.03, 1 0.04] = 13.81
DALYs averted [0.03, 0, 0] = 13.76
DALYs averted [0, 1, 0.04] = 25.07
DALYs averted [0, 0, 0] = 27.40

Therefore, in this example, discounting (at 3 percent rather than 0 percent) has a much bigger impact on the DALY calculation than does the choice of age weights. Once the DALYs averted from an intervention have been calculated for an individual, the total for the community can be summed and then compared with the difference in costs between providing the treatment and not providing it, to yield the incremental effects for an incremental cost-effectiveness ratio[29].

29 Because this compares treatment versus no treatment, the incremental cost-effectiveness ratio will equal the average cost-effectiveness ratio. It equates, in this instance, with the approach forwarded by Murray et al. (2000). However, note that this approach takes no account of the impact of any other existing treatments because the comparison is with no treatment.

3 THE USE OF DALYS

The range of proposed uses of DALYs outlined by Murray and colleagues (Murray, 1994; Jamison in Murray and Lopez, 1996a; Murray, Salomon and Mathers, undated) include:

- setting health service priorities for curative and preventive care;
- setting priorities for health research and development;
- identifying disadvantaged groups and targeting health interventions;
- providing a comparable measure of output for intervention, programme and sector evaluation and planning;
- generating a forum for informed debate of values and priorities;
- allocating training time for clinical and health practitioners;
- assessing the performance of health systems[30];
- comparing geographical and secular variation in health[31].

One of the frustrations that developers of health-related quality of life measures have had is that their measures are not always used for the purposes for which they were intended (Rosser, 1990). Criticisms of their use in inappropriate ways cannot be a criticism of the measure *per se*, but of its use. Therefore, this section distinguishes uses by the

30 Most recently, the World Health Report (WHO, 2000) has used DALYs and disability adjusted life expectancy as part of an assessment of the performance of countries' health care systems. Disability adjusted life expectancy was combined in an index along with measures of: the distribution of health in the population; the overall responsiveness of the country's health care system; the distribution of responsiveness; the distribution of financial contribution (WHO, 2000, p27); and were then compared with health expenditure per capita (WHO, 2000, p43). The results were widely reported in the popular press, and government officials were charged with finding out why their country ranked where it did.

31 A series of related papers by Jamison and colleagues have outlined the role of the findings from the World Development Report (World Bank, 1993) in the broader context of health sector reform (Aiyer et al., 1995; Jamison et al., 1995; Jamison, 1997). They have consistently advocated that the main areas for improvements include: strengthening the household's capacity to improve health; improving national spending on health; and promoting diversity and competition in the financing, organisation and provision of health services (particularly with reference to extending the coverage of health insurance). Also see Musgrove (1995) for an outline of how the emphasis on cost per DALY averted is related to choices for alternative funding mechanisms.

original developers[32] and by others in each of the four principal areas in which DALYs have been used:

- estimates of the burden of disease;
- individual studies of cost-effectiveness;
- 'sectoral' planning;
- directing research and development effort.

3.1 Estimates of the burden of disease

There have been a number of iterations in calculating the global burden of disease. The third attempt was the basis of the World Development Report (World Bank, 1993) that covered 109 categories of ICD9[33], disaggregated by: eight regions of the world, five age

32 Separating out the developers of the DALY from 'others' is quite difficult given the breadth of multi-authored papers. For the purposes of this review I have included three authors; Christopher JL Murray, Alan D Lopez and Dean T Jamison. Murray was sole author of the two main papers which outline the conceptual basis of the DALY but has jointly edited the Global Burden of Disease series with Lopez, within which is the most comprehensive presentation of data on DALYs. Jamison was selected because he was the project manager of the World Development Report in 1993 (within which was the inaugural presentation of DALYs) and the lead author evaluating the cost-effectiveness analysis in the Health Sector Priorities Review (Jamison et al., 1993) which was presented in the World Development Report (World Bank, 1993). Any papers authored, or co-authored by these three were included for review within this section. I also include any papers in the collected volumes edited by Murray and Lopez (1996a-c). Christopher Murray is currently the Executive Director of the Evidence and Information for Policy Cluster of the WHO and was Professor of International Health Economics and Director of the Burden of Disease Unit at Harvard University during the original launch of DALYs. His academic background encompasses medical training at Harvard and a DPhil in Economics from Oxford University. Alan Lopez's academic background includes mathematics followed by postgraduate degrees in statistics and medical demography. He has worked at the WHO since 1980 and is currently co-ordinator of the Epidemiology and Burden of Disease Unit there. Dean Jamison read philosophy and later engineering sciences at Stanford University, followed by a PhD in economics at Harvard. He is currently a Professor of Social Research Methodology and Director of the Program on Global Health and Education at the University of California, Los Angeles. He has also spent many years as a senior economist in the research department at the World Bank.

33 This is the 9th revision of the International Classification of Disease, which is used to code and classify mortality data from death certificates. The 9th revision covers the period 1979-1998.

Table 4 **Top 10 causes of DALYs [0.03, 1, 0.04] (percent of total)**
in descending order, 1990

Developed regions of the world			Developing regions of the world		
Disease or injury	**DALYs ('000s)**	**% of total**	**Disease or injury**	**DALYs ('000s)**	**% of total**
Ischaemic heart disease	15,950	9.9	Lower respiratory infections	110,506	9.1
Unipolar major depression	9,780	6.1	Diarrhoeal diseases	99,168	8.1
Cerebrovascular disease	9,425	5.9	Conditions arising during the perinatal period	89,193	7.3
Road traffic accidents	7,064	4.4	Unipolar major depression	41,031	3.4
Alcohol use	6,446	4.0	Tuberculosis	37,930	3.1
Osteoarthritis	4,681	2.9	Measles	36,498	3.0
Trachea, bronchus and lung cancers	4,587	2.9	Malaria	31,705	2.6
Dementia and other degenerative and hereditary central nervous system disorders	3,816	2.4	Ischaemic heart disease	30,749	2.5
Self-inflicted injuries	3,768	2.3	Congenital abnormalities	29,441	2.4
Congenital abnormalities	3,480	2.2	Cerebrovascular disease	29,099	2.4

Source: Murray and Lopez (1996a, p262).

groups, gender and cause. The fourth iteration, presented in Murray
and Lopez (1994a) added the disease categories 'other neoplasm' and

'skin disease'. The fifth iteration[34] was published in 1996 (Murray and Lopez, 1996a-d, 1997b-e; Lancet, 1997).

The results show, for example, that in 1990 the top cause of DALYs lost in developed regions of the world was ischaemic heart disease, whereas for the developing regions it was lower respiratory infections. Table 4 shows the much larger proportion of DALYs lost due to infectious disease in developing compared with developed regions. One other finding was the larger than expected contribution of road traffic accidents to total DALYs lost for developed countries.

The results have also shown the percentage contribution of DALYs lost by age group. Figure 3 shows that the distribution of DALYs lost by age differs dramatically across regions. In Sub-Saharan Africa 52 percent of the burden is due to mortality and morbidity in the 0-4 year age group, whereas in the established market economies more than 50 percent of the burden is in the 45+ age group. 41 percent of all DALYs lost globally occurred in the 15-44 year age group, which has resulted in a call for 'global health action to promote health and to prevent disease and injury among young adults' (Murray and Lopez, 1996a, p156).

Table 5 gives an idea of the greatest level of detail attained in the global burden of disease exercise, showing the spread of cardiovascular disease by age and sex for the established market economies. It can be seen that roughly 50 percent of the burden of cardiovascular disease is due to ischaemic heart disease, and that men lose 30 percent more DALYs to cardiovascular diseases than women. 65 percent of the burden of cardiovascular disease occurs in the over 60s, particularly in women.

The loss of DALYs has also been attributed to risk factors. In developing countries malnutrition is the main risk factor causing DALY loss, followed by poor water supplies, sanitation and hygiene.

34 Of the ten volumes, three have been published to date. The categories of this disease burden exercise: dropped pelvic inflammatory disease, detailed categories of diarrhoeal disease, and eclampsia; and added dengue, Japanese encephalitis, and detail of two conditions arising during the peri-natal period including low birth weight and birth asphyxia/birth trauma; separated out schizophrenia, obsessive-compulsive disorders and panic disorder; and also separated out ten categories of congenital abnormalities. It also included several risk factors such as indoor air pollution.

Figure 3 **DALYs lost by age and region, 1990**
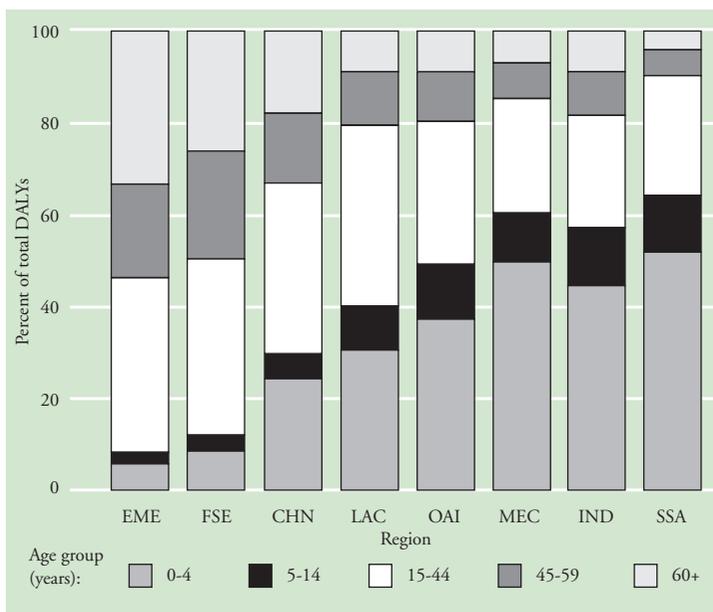
Notes: EME = established market economies; FSE = former socialist economies of Europe; CHN = China; LAC = Latin America and Caribbean; OAI = Other Asia and islands; MEC = Middle Eastern Crescent; IND = India; SSA = Sub-Saharan Africa.

Source: Murray and Lopez (1996a, p256).

In the developed regions, tobacco and alcohol were the main risk factors. Murray and Lopez called for intensification of the expanded programme of immunisation in developing countries and for reducing the DALYs caused by addictive substances in developed regions (Murray and Lopez, 1997e, p1441). Murray and Lopez have also provided a series of projections of DALYs, by cause, to 2020. An aging population, the spread of HIV, the increase in DALYs due to tobacco and the decline in death from communicable diseases were judged to be the main causes of the expected change in patterns of disease (Murray and Lopez, 1997d).

Table 5 **DALYs [0.03, 1, 0.04] by age and sex for cardiovascular diseases (thousands) in established market economies[a], 1990**

	Males						Females					
	0–4	5–14	15–44	45–59	60+	Total (%)	0–4	5–14	15–44	45–59	60+	Total (%)
Cardiovascular diseases: total	62	28	1,430	2,714	6,176	10,410 (100%)	50	22	723	1,170	5,965	7,930 (100%)
Rheumatic heart disease	0	0	14	17	26	57 (0.5%)	0	0	15	28	61	104 (1.3%)
Ischaemic heart disease	0	0	490	1,589	3,339	5,418 (52%)	0	0	133	484	2,841	3,458 (44%)
Cerebrovascular disease	9	6	395	552	1,460	2,422 (23%)	7	5	305	427	1,810	2,554 (32%)
Inflammatory heart disease	12	6	169	111	114	412 (4%)	14	5	84	52	116	271 (3.4%)

Notes: [a] Andorra, Australia, Austria, Belgium, Bermuda, Canada, Channel Islands, Denmark, Faeroe Islands, Finland, France, Germany, Gibraltar, Greece, Greenland, Holy See, Iceland, Ireland, Isle of Man, Italy, Japan, Liechtenstein, Luxembourg, Monaco, Netherlands, New Zealand, Norway, Portugal, San Marino, Spain, St Pierre and Miquelon, Sweden, Switzerland, United Kingdom and United States.

Source: Murray and Lopez (1996a, p543).

Perhaps not surprisingly, the most frequent use of DALYs by people other than the original developers of the measure is the citation of results from the different global burden of disease exercises. There has been a profusion of organisations using or discussing DALYs, including national agencies such as Rehabilitation International[35], the UK Public Health Laboratory Service[36], the World Resources Institute[37] and the US National Institutes of Health (NIH). Sub-national organisations such as departments of health in cities as well as directors of public health have also used DALYs in their reporting[38]. Other DALY-using groups include aid wings of governments[39], charities[40], commercial companies (Friedman, 2000) and teachers[41]. The citations in the DALY literature fall into two principal categories: those that accept the results and use the figures to call for greater funding of specific disease based interventions, research and training[42]; and those that question the estimates of DALYs lost for specific diseases.

Calls for increased expenditure on interventions and research have

35 See www.rehab-international.org/measuringquality.htm
36 See www.phls.co.uk/whoweare/corpplan.htm
37 See www.igc.org/wri/wr-98-99/dalys.htm
38 For example, see
WWW.STAFFHEALTH.CO.UK/HADocs/REPORTS/PUBLIC%20HEALTH/HEALTH/NEXT-PAGE.HTM   and
http://PHPH.DHS.CO.LA.CA.US/PH/REPORTS/PHLAUCLA/PHLAUCLA.HTM
39 See, for example, the UK Department for International Development's strategy paper 'Better Health for Poor People' (1999), which uses DALYs and the global burden of disease estimates for targeting development (p11), even though the DALY is not yet considered suitable as a performance measure.
40 E.g. United Cerebral Palsy of Prince George's and Montgomery Counties, http://www.ucppgmc.com
41 Taylor (1997, p156) cited the work as a 'fascinating and plausible reference book', that he would draw on for senior undergraduate and graduate level health geography courses.
42 See for example Gupta et al. (1994), who believe that the global burden of disease does not under-report cancer deaths in India (contrary to Murray and Lopez's view), and Meerding et al. (1998) who considered the main causes of disability by age in The Netherlands compared with the global burden of disease results.

54   been particularly noticeable in the areas of mental health[43], infectious diseases such as HIV/sexually transmitted diseases (Van Dam, 1995), tuberculosis (Prabhakar, 1996; Squire, 1997), otitis media (Berman, 1995,) and malaria (Foster and Phillips, 1998). Fèvre and Barnish (1999) have pointed to the role DALYs and other economic indicators of the burden of disease may play in attracting resources for combating disease, in their case malaria, even though others have argued that the advocacy is now decoupled from DALYs and the epidemiology (Andrews et al., 1998a). A more unusual use has been by Jonsson (1998), who attempted to link estimates of DALY loss due to diabetes to costs of illness. He concluded that any increase in incidence, and therefore prevalence, would have a considerable economic impact. Following a summary of injuries and non-communicable diseases among children in developing countries, Deen et al. (1999) called for more attention to be paid to the area, and for country estimates of the burden of disease and the cost-effectiveness of intervention strategies to be made.

Estimates of DALYs lost have been questioned in two ways. One approach has been to provide alternative measures of impact alongside DALY estimates. For example:

- Ramaiah et al. (1997) developed a disease specific questionnaire to quantify the effect of lymphatic filariasis on domestic and occupa-

---

43 See for example: Eisenberg (1997a) 'the nations of the world must make a major commitment to upgrade the quality of mental health services'; Blue and Harpham (1994 p12) 'sufficient evidence … now exists to warrant a greater emphasis being placed on common mental disorders'; Shore (1999 p336) who concludes that 'the next task for research is to prove what clinicians experience everyday in their work with individual patients – that modern treatment of psychiatric disorders is worth the money and is even, in many cases, a bargain'. Andrews (2000) argued that using burden of disease figures alone, mental health services would warrant a three-fold increase in spending relative to other health services in Australia. However, he also argues that spending should be related to the cost-effectiveness of interventions but notes the absence of data. Therefore he outlines four alternative ways of splitting the budget for mental health care, assuming that half of all care is emergency rather than a discretionary service. Bremberg (1998) also used the global burden of disease results to focus on the mental health of children 'since this is the leading cause of DALYs lost in this age group'. Finally, see a published series of supportive letters by Appelby et al. (1997) and Eisenberg (1997b).

tional activities and physical movement. They found that 66 percent of all patients reported limitations in occupation and 56 percent of women reported limitation in travel and domestic activities, and that the most severe restriction was amongst those patients with acute episodes of adenolymphangitis. This, they suggested, provided evidence that DALY estimates were likely to be too low because previously there had been no data on functional impairment and disability and they had shown that it was high;

- Guerrant et al. (1999) studied the physical fitness, activity and cognitive development of a longitudinal cohort of 26 children aged 6.5-9 years in Goncalves, a shantytown near Fortaleza, Brazil. They argued that potentially there was a substantial impact from early childhood diarrhoea and, if confirmed, there would need to be significant revisions to estimates of DALYs lost from diarrhoea and DALYs averted by interventions in cost-effectiveness analyses[44];

- Andrews et al. (1998b) have been among the few to suggest that DALYs lost as a result of mental disorders may have been overestimated. They showed that mental disorder had particularly high levels of co-morbidity and that therefore the measured global burden of disease was likely to overestimate any potential gains for treating this burden of disease, and especially for affective disorders and substance abuse.

Such criticisms of DALY estimates were foreseen. Indeed Mansourian (1996, p334) suggested that one of the advantages of using DALYs is that it would help highlight the data most urgently needed.

The second approach to criticising the burden of disease estimates has been to question the categorisation by cause. For example:

44 Kale (1998), rather than providing new estimates of burden, drew on Remme's work (Remme, 1998) to claim that the global burden of diseases underestimates DALYs lost from onchocerciasis. The two main reasons given were the reductive measure of benefits (which does not account for the social costs of disfigurement from river blindness) and the definition of blindness adopted (which assumes only those with 3/60 sight are blind, whereas Kale argues that due to the environment even those with more sight than that can be treated as practically blind).

- the burden of blindness has been considered to be underestimated because it 'is limited by the categorisation of the major causes of blindness within children within other groups of disorders such as malnutrition, communicable disease, perinatal disorders and congenital abnormalities' (Rahi et al., 1999, p387);

- Hyder and Morrow (1999) show the impact of assuming that DALYs lost begin at the age of death rather than at the age of disease onset. They argue that this causes the DALY to underestimate the burden of HIV and any other disease with a strong secular trend.

Appendix 4 outlines eight studies that provide new estimates of burdens of disease using DALYs[45]. Four of these studies included communicable diseases, two non-communicable diseases, one covered all types of diseases and one focussed only on providing a new set of disability weights for The Netherlands. Four of the studies were funded, at least partly, by national governments and two by research organisations, with the other two not stating the source of funding. Only three papers focussed on a specific country (The Netherlands, Australia and UK), as the majority considered globally distributed populations such as the world's poor, or endemic malarial areas.

The main reasons for using DALYs in these studies were the desire to use a single outcome measure (Martens et al., 1995) and to make comparisons with other diseases (Chan, 1997) but also because the impact of the global burden of disease exercise led researchers to believe that DALYs were more likely to be used in the future to influence policy (Stouthard et al., 1997; Hollinghurst et al., 2000). Hollinghurst et al. (2000) took a more unusual slant as they calculated 'avoidable' DALYs based on considering whether effective treatment was available. The eight papers were split between using Mark 1 and Mark 2 DALYs. It is noticeable that, despite desires for comparison, the assumptions and sources of data used in the construction of these DALYs differ significantly between studies. Even for those presenting base case DALYs [0.03, 1, 0.04], the sources of data either differed or were unclear.

---

45 These papers were found using the search strategy set out in Appendix 1.

3.2 Cost-effectiveness analysis

One of the most extensive uses of DALYs in cost-effectiveness analysis was provided by Jamison et al. (1993) and reported in the World Development Report (World Bank, 1993). Using the DALY Mark 1, a comparative analysis of 50 specific health interventions in 24 disease areas was presented for developing countries. The results were an amalgam of conclusions from the ensuing chapters of the book even though it was recognised that the methods of assessing cost and effectiveness of intervention varied by chapter, thus complicating comparisons, but they were accepted as 'reasonable first approximations'. Nevertheless Jamison stated that 'our concern is with generalization – with addressing trends and findings that are important for a sufficiently large number of countries that they assume significance for the developing world as a whole'[46] (Jamison, 1993). The broad policy conclusions drawn were that:

- there is no strong tendency for primary prevention or public health interventions to have a superior cost-effectiveness;
- virtually no cost-effective interventions require more specialised facilities than those available at district hospitals;
- a broad array of interventions are likely to make economic sense, whatever the income level of the country;
- governments should plan to expand those services with costs below $75 per DALY averted, for which a list was provided.

Appendix 5 reviews the 16 papers found[47] that used DALYs as an outcome measure in cost-effectiveness analyses. It can be seen that 11 of these were published after 1997. The majority focus on interventions for communicable disease, and there is a preponderance of papers about Sub-Saharan Africa. These studies tended to be funded by single agencies and most often by an international organisation or a bilateral aid agency. Nearly all papers compare the results of their study with the 'cut-off points' used by the World Bank to signify good

46 One reason for confidence in the results is that the World Development Report claimed 'differences in cost-effectiveness between one intervention and another are often much larger than either the variation from one locale to another or the uncertainty of the estimates' (World Bank, 1993, p62).

47 These papers were found using the search strategy set out in Appendix 1.

or poor value for money. They also compare their results with the cost-effectiveness ratios for treating other diseases in different countries around the world. Compared with the papers using DALYs for burden of disease studies, summarised in Appendix 4, a greater proportion of the cost-effectiveness analyses, in Appendix 5, focus on 'hypothetical' populations, but there is a similar split between those using Mark 1 and Mark 2 DALYs.

With respect to the construction of DALYs, it is interesting to note that, when the assumptions were made explicit, all used a discount rate of 3 percent. However, four papers chose to exclude the age weighting, and in another six it was not clear what had happened. When using age weighting, 0.04 was always used. With respect to disability weights, the picture is different: two papers used only YLLs to calculate DALYs and therefore the issue was not relevant; in eight cases it was unclear whether weights had been applied or what rate was used if they had; in two papers the weights applied were those developed in Murray and Lopez (1996a); and in four cases the weights cited had been drawn from other literature or assumptions of the authors. There is as much variation in the approaches taken to life expectancy: four papers did not state the life expectancy used or the source of data; five papers appeared to use the standard life expectancies presented in Murray and Lopez (1996a), where life expectancy at birth is 80 years or more; and the remaining seven used various estimates for local cohorts. It is important to note that those using the standard life tables reproduced in Murray (1996) will overestimate the relative efficiency of the interventions.

Of the 16 papers, eight did not present any sensitivity analysis concerning assumptions involved in estimating DALYs. Of those that did, the variables most frequently examined were the discount rate, life expectancy and case fatality rate. The disability and age weights were only examined once each.

3.3 Sectoral prioritisation exercises

The developers of the DALY combined burden of disease and cost-effectiveness analyses to provide 'sectoral analyses'. The underlying

belief was that priorities for health sector spending 'should go to those health problems that cause a large disease burden *and* for which cost-effective interventions are available' (World Bank, 1993, p63). This had a significant impact on the proposals for governments to consider adopting the 'minimum' and 'essential' packages of care recommended[48]. The justification for the design, content and financing of an essential package for national health services was outlined by Bobadilla et al. (1994) for low- and middle-income countries[49], based on the findings of Jamison et al. (1993) and the global burden of disease exercise. Whilst the mechanisms for undertaking the sectoral analysis are a little vague in Bobadilla et al. (1994), Bobadilla and Cowley (1995) elucidate the World Bank's determination of an essential health services package in practice. Table 6 shows the steps planned. It can be seen that, following an assessment of the burden of disease, broad judgements of the cost-efficacy[50] of interventions are made as a first stage process in selecting interventions and prior to any estimation of local costs.

There are several references testifying to the use of DALYs in prioritisation exercises. For example, Andrews et al. (1998a, p157) wrote that 'developing countries such as Mexico have successfully applied the DALY based methodology to define priorities and essential packages of care, and over 20 other countries are trialing this approach. The United States, Japan, the Netherlands, Sweden and most recently Australia have begun or are planning national burden of disease studies'. However, whilst the approach outlined in Table 6 was followed in Mexico, Bobadilla and Cowley (1995) also show that countries with higher mortality rates and lower incomes do not manage in practice to

48 An alternative, illustrative, model of resource allocation using DALYs for health sector decision-making is given in Murray (1995a,b) and Murray et al. (1994).

49 Considering packages of interventions, rather than single interventions in isolation, was considered particularly advantageous because of: the ability to consider jointness in production as well as co-morbidities; potential synergies between treatment and prevention; and the possibility of improving the use of specialised resources through screening and referral.

50 Efficacy is distinguished from effectiveness, although without specific definition in Bobadilla and Cowley (1995). It would appear that the distinction is made between the potential (efficacy) and actual effectiveness in practice, given a country's different health care scenarios.

Table 6 **Basic steps in designing a package of essential health services in middle-income countries**

Step	Activity
1	Estimate the burden of disease due to premature mortality and disability for approximately 100 diseases and injuries by age, sex and geo-political region
2	Estimate the burden of disease for a selected list of proximate risk factors, for example, inadequate water supply and sanitation, protein energy malnutrition, tobacco consumption and alcohol abuse
3	Select interventions that fulfil one of the five criteria: favourable cost-efficacy; ability to address diseases, injuries or risk factors with a high burden (about 1 percent or more of the total); high supply or demand; high recurrent expenditure; ability to address diseases of public concern. Health interventions should include three types: public health, community outreach, and clinical (outpatient and hospital)
4	Estimate costs for the selected interventions and clusters of interventions that are justifiably provided together. This includes making normative assumptions about technical efficiency and desired quality of care
5	Assemble information on efficacy of interventions and estimate the expected effectiveness for the country's different health scenarios
6	Define time frame and population coverage (current and targeted) to estimate incremental cost per intervention per year (including overheads and cost of limited diagnostic and health services)
7	Congruent with the current model of care and constrained by the financial resources available, construct a package of cost-effective interventions that can be delivered in the same facilities and through common mechanisms

Source: Bobadilla and Cowley (1995, pp553-554).
© John Wiley & Sons Limited. Reproduced with permission.

work through all these steps. For example, countries in East Africa considered mortality only and a Ugandan study considered neither burden of disease nor cost-effectiveness analysis specific to Uganda.

One of the most useful documents detailing this work is by Bobadilla (1996), who reviewed the scope and methods of setting priorities in World Bank lending projects. He showed that all the large studies took place in middle-income countries[51] (mostly in Latin America and the Caribbean) and that 'none of the 19 countries undertook any analysis of the types of disabilities and only two did any serious analysis of risk factors'[52]. He highlighted that the most ambitious study took place in Turkey at an estimated cost of over US$3 million, and that negotiations had shown concern not to allow control to be in the hands of foreign universities and consultants. In discussing the ownership of the process by individual countries Bobadilla also commented that World Bank staff or foreign consultants had written 17 out of 24 reports.

In the application of the DALY, Bobadilla stated (1996, p24) that 'half of the countries made significant modifications, and one (Russian Federation) used a different indicator. Five countries did not discount future health losses and all the countries that discounted used a three percent rate. Five countries did not apply any age weights' and three used a different disability scale. He concluded that 'only seven of the 19 study results are comparable in terms of the indicator and scope of the study, five belong to the Latin American and Caribbean region'. However, for nationally run studies, it is not surprising that there are differences because, as Bobadilla (1996, p25) pointed out, 'international comparability of burden of disease studies should not be a primary concern of countries'. He also showed that five countries used no cost data, a further four had cost data on fewer than 20 interventions, and only three focussed specifically on cost-effectiveness (for a maximum of 35 interventions).

51 Bobadilla (1996, p15) suggested that this was probably linked to the intensity of World Bank lending rather than any other characteristics.
52 He also developed this point to suggest that the consequences of this were that no country proposed any rehabilitation services (when analysis has shown this to be very cost-effective), and that therefore resource allocation was significantly biased. The lack of consideration of risk factors was also argued to reduce the emphasis given to public health and preventive activities (Bobadilla, 1996).

Few examples of sectoral prioritisation exercises are published[53]. The findings of a four year project based in two districts in Tanzania that set out to 'test the hypothesis that burden of disease and cost-effectiveness analysis should provide the basis for health services planning in low income countries' (Finlay et al., 1995, p1083) are on the point of being reported (The Economist, 2002). This article suggested that using a combination of data on the burden of disease and on the cost-effectiveness of interventions to plan the provision of an additional 80 cents spending per person resulted in major health improvements: infant mortality fell by 28 percent and mortality of children under five by 14 percent between 1999 and 2000. It will be interesting to read the full reports soon.

Parallel to the on-going debates about, and construction of, essential national health services is the development of the World Bank initiated 'sector-wide approach'. Cassels and Janovsky (1998) explain that this approach is intended to move from the individual project-based funding of most donor agencies to their contributing instead to the funding of the entire health sector[54]. Feacham justified the sector-wide approach as a way for national governments to co-ordinate donors, and ensure that there is a single set of disbursement, procurement, evaluation and reporting procedures (Abbasi, 1999a, pp1207-8). To counterbalance donor agencies losing their power to completely determine their own priorities within countries, they were invited to participate jointly (without a controlling interest) in the planning of health sector activities as a whole. Cassels and Janovsky (1998) drew on around four years' experience from four countries (Zambia, Ghana, Pakistan and Bangladesh) to suggest that whilst the results looked

53 Carter et al. (2000) reported the results of a priority setting exercise for cancer control using programme budgeting and marginal analysis. They used DALYs as the outcome measure and the global burden of disease results from the study by Mathers et al. (2000) that was commissioned for the Australian Ministry of Health. At the same conference, Oswald (2000) reported the results of a prioritisation exercise in two states in the US based on Tengs et al. (1995) and a model of the population's health and demographic characteristics. I also understand that, sometime in the future, WHO are planning to calculate and report the cost-effectiveness of 100 interventions in 17 regions of the world.

54 All private and public institutions that have a stake in health, including the household (Cassels and Janovsky, 1998).

promising, the sector-wide approach is quite varied in its implementation and therefore not a universally applicable model[55]. However, it appears that this is a framework within which the use of DALYs in calculating the burden of disease and expected cost-effectiveness of health interventions might be given substantial support.

3.4 Indicating the direction for health research and development

The developers of DALYs have also been involved in the report 'Investing in Health Research and Development' (Ad Hoc Committee on Health Research Relating to Future Intervention Options, 1996), which used DALYs to consider the justifications for future research and development. Table 7 shows the broad approach adopted, with three selected examples. The Ad Hoc Committee report states that only a few selected examples could be evaluated rigorously and that many evaluations relied heavily on expert judgement and qualitative analyses (p9). The 17 recommendations in the report were aimed at governments, industry and other investors, concerning the allocations of funds to, and within, health research and development budgets. The recommendations covered a wide range of specifications such as:

- reallocations of money (e.g. from HIV to TB);
- increasing the size of resources for developing and evaluating packages of care (e.g. for the 'Mother-Baby package');
- the development and evaluation of policy instruments (e.g. for tobacco control);
- studying the impact of research on policy;
- establishing a forum to bring together governments, donors and researchers[56].

55 But note that some Pakistani officials have been critical of the 'problem of leadership among donors' because just having what might be a good idea does not mean it can be implemented suddenly. They complained that: systems in Pakistan were said not to be geared up for such a change; the changes suggested were not based on the reality of the situation on the ground; and that system-wide changes may take a long time to trickle down to an impatient and poverty stricken populace (Abbasi, 1999b, p1133).

56 The Global Forum for Health Research was set up in January 1998 and reports of its activities, in particular its re-direction of research, can be found in Global Forum for Health Research (2000) and on the website www.globalforumhealth.org

Table 7 **Steps to inform resource allocation for research and development in health: selected examples**

Condition or risk factor	Need	Opportunity: promise of R&D effort					Investment requirement
	Disease burden (rank of 96 causes)	Primary reasons for persistence of burden	Current knowledge base/R&D capacity	Desired intervention/estimated likely cost-effectiveness (in US$)	Probability of success	Current effort, additional cost (in US$)/time frame	Conclusion
Pneumonia	High (1)	Failure to use existing tools efficiently	Good	Package for integrating management of the sick child Very high (<US$ 50 per DALY averted)	High	Current investment relatively small; required further costs and time frame modest (US$ 15 million over 3 years)	High priority for investment
Malaria	High (11)	Lack of tools	Moderate to good	Malaria vaccine Very high (<US$ 30 per DALY averted and <US$ 15 in some circumstances)	High	Current investment relatively small; required further costs and time frame moderate (US$ 50 million over 10 years)	High priority for investment
Leprosy	Low (95)	Failure to use existing tools efficiently	Good	Leprosy vaccine Low (US$ 2,543 per DALY averted compared with <US$ 50 for multi-drug therapy based on passive case finding and treatment, or US$ 42 – US$ 7,200 based on active case finding)	Moderate	Not assessed	Low priority for investment

Source: Ad Hoc Committee on Health Research Relating to Future Intervention Options (1996, p9).

Two other organisations have also published work concerning the use of DALYs in the prioritisation of research and development. Oortwijn et al. (1998, p226) reported initial experiences with a new model of priority setting for the US$21 million annual budget of the Investigative Medicine Programme of the Dutch Health Insurance Council[57]. They concluded that four dimensions were important for considering 'relevance'. The first was the burden of disease, for which they cited the World Development Report as 'playing a role', despite noting its 'controversial' method of measuring burden[58].

The other, possibly better-known, example is the retort published concerning the NIH, whose credibility had been questioned by advocacy groups with respect to their choice of disease areas for research funding (Varmus, 1997, 1999; Gross et al., 1999). Following pressure, an analysis 'to determine whether the amount of funding the NIH allocates to research on particular diseases is associated with commonly available measures of the burden of disease' was undertaken. The DALY was among the outcome measures chosen by the NIH because 'it incorporates the social values placed on various aspects of the physical, mental and social function' (Gross et al., 1999, p1883). Gross et al. found that the relation between NIH funding and DALYs lost in 'established market economies' was statistically significant at the 99 percent level (Spearman's rho = 0.62, p < 0.001), and that such a relationship could not be found when NIH funding was related to incidence or prevalence, although there was a positive association with mortality (Spearmans' rho = 0.4, p = 0.03). Turning the tables on the advocacy groups, it was pointed out that such groups can mislead policy makers if they only provide information using a single indicator to measure the burden of disease, as the results are extremely sensitive to the measure chosen.

57 This was described as the most important programme for health technology assessment in the Netherlands, and is accessed predominantly by universities and university hospitals.
58 The other dimensions included: searching Cochrane reviews; the HARMET and Eur-ASSESS European projects which aim at developing consensus in methods for economic evaluation; and implementation of results.

4 CRITIQUES OF DALYS

There is a growing critical literature focussing on the construction and use of DALYs. The aim of this section is to outline the range of criticisms raised in the literature to date. Many criticisms have focussed on the specific assumptions concerning approaches to weighting DALYs and I begin with these. This is followed by various concerns raised about: whose values are, and should be, represented; the existence and quality of data; and how helpful DALYs are for decision-making in the health sector. The range of criticism is summarised in Box 5. It is particularly noticeable that the majority of critiques centre on notions of value. Perhaps this is because, as Sundby (1999 p280) stated:

> 'the initial presentation of the concepts as a tool in a World Bank policymaking document, already demonstrates that it is difficult to separate the critique of the methodology itself from its application in a politicised debate on health reform'.

Box 5 **The range of critiques of DALYs**

1. The appropriateness of weights in DALYs for:
   - life expectancy
   - age
   - future time
   - disability

2. Whose values are, and should be, represented?
   - conceptions of 'health', 'disease' and 'disability'
   - the role of 'experts'
   - the universality of disability weights
   - gender bias
   - human rights
   - integrating equity

3. Quality of data

4. DALYs' usefulness for decision-making:
   - costly and time-consuming to construct
   - international versus national priorities
   - burden of disease versus cost-effectiveness analysis

4.1 Alternative views about the weights in DALYs

4.1.1 Life expectancy

Two aspects of life expectancy have been questioned: the choice of an 'ideal' standard for life expectancy; and the size of the difference between male and female life expectancy. In choosing to use DALYs to calculate the burden of disease, Bowie et al. (1998, p41) question the relevance of using Japanese life expectancies for national or regional priority setting in the UK. They argue that using Japanese life tables would give more weight to the elderly relative to the young than if British life tables were used. This highlights the potential differences in the needs or desires of national and international decision-makers and also reflects the concern that any one intervention is unlikely to have a major impact on overall DALYs because of all the other contributors to mortality and disability. However, Musgrove (2000, p111) suggests that national decision-makers should use the 'standard expectation of life' approach because 'if all the diseases contributing to the burden were controlled to the extent that they are in Japan, life expectancy would rise to 80 or more years, just as in that country'[59].

Williams (1999 p4) develops Bowie et al.'s criticism and considers the implication of assuming a hypothetical standard for life expectancy for cross-country comparisons. Whilst acknowledging the appealing moral position, he criticises the 'strong political judgement' on two grounds. First he queried whether such a 'fiction' should be introduced as part of the composite measure of DALYs. He suggested that it would be clearer if two sets of calculations were made, so that DALYs using actual and hypothetical life expectancy could be compared. This would allow people to see the impact of the political judgement made for international comparisons, and would presumably be more acceptable for national level data analysis. Williams' second concern focussed on the implications of the current decision to base DALYs on a life expectancy of 80/82.5 (male/female) in all countries. He wrote that it

---

59 Presumably this would also entail 'interventions' to include the same diet, level of exercise and occupation – i.e. almost 'becoming' Japanese.

'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 were 40, then each actual year carries an 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 global burden of disease calculations, this process generates some very peculiar results concerning intergenerational equity, both within and between countries.' (Williams, 1999, p5)

Using Figure 4, Williams shows these implied equity weights for four different regional model life tables, when life expectancy at birth is 50 years. It shows that that those who survive into older ages are accorded more weight to the remaining years of their lives than are younger people in every region, and particularly in the 'East' and 'South' regions[60]. Figure 5 takes this data one stage further, dividing the 'East' life tables into 'poor', 'middle-income' and 'rich' countries on the basis of life expectancy[61]. This shows that old people in richer countries are given more weight than young adults in poor countries, which he imagined was 'an unintended and unwanted side-effect' in a measure purportedly trying to avoid the impact of inequalities in life expectancies.

Anand and Hanson (1997, p690) criticise the choice of 2.5 years difference in life expectancy between males and females as arbitrary. First they highlight the significant disagreement regarding the relative contributions of biological, environmental and social factors leading to biological-genetic differences. Secondly, they criticise the generalisation of Murray's method, based on projecting an observation of narrowing gender gaps among higher income quintiles in urban Canada, across time and space. Most specifically they argue that the actual gap

60 Note that these are the equity weights implied as a direct result of using a hypothetical standard for life expectancy. As such they are different from the age weights imposed additionally which were outlined in Section 2.3.2.
61 As Williams argues that life expectancy is closely correlated to real income per head (Williams, 1999, p5).

Figure 4 **Implied age weights when female life expectancy at birth is 50, for different regional patterns**
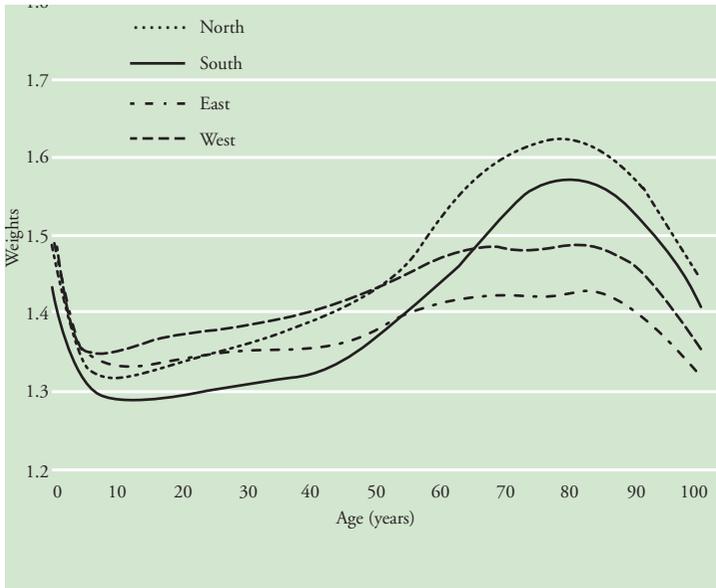
Source: Williams (1999, p5).
© John Wiley & Sons Limited. Reproduced with permission.

is likely to be wider, and that therefore the reduction effects a movement of resources away from women if burden of disease estimates are used for decision-making. However, although Murray has argued that the biological difference in mortality lies somewhere between 1.9 and 3.2 years (Murray, 1996, p18), he has also questioned the inequity of the difference in length of life between men and women, as well as suggesting that the biological difference could just be added as one more contributor to the burden of premature death (Murray and Acharya, 1997, p712).

4.1.2 Age weights
Debate concerning the age weights in DALYs has centred on: questioning the conceptual basis of weighting by age; potential double

Figure 5 **Implied weights for 'East' only with differing life expectancies at birth**
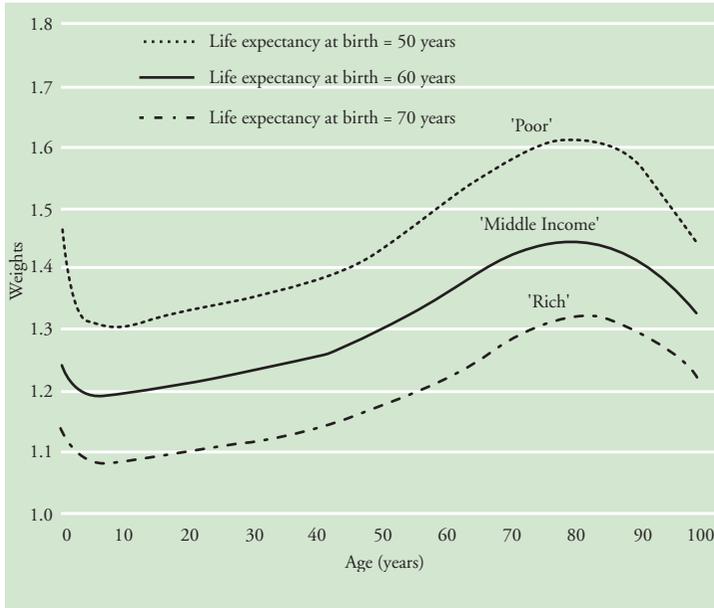
Source: Williams (1999, p5).
© John Wiley & Sons Limited. Reproduced with permission.

counting introduced by age-weighting; and the evidence for an 'appropriate' rate.

Four arguments have been put forward that question the conceptual basis behind weighting life years by age:

- a principal of 'universalism of life' has been invoked to argue that the value per life year should be common to all people regardless of their age (Anand and Hanson, 1997, p694)[62];

- using notions of dependency discriminates against those with fewer dependencies and social ties. However, as much of health

62 This assumes no diminishing marginal utility of additional life years (Donaldson et al., 2002)

care is given to an individual rather than on the basis of other peoples' dependence on them, such principles are irrelevant as an ethical base for allocating health care resources (Brock, 1998, p78);

- principles are used inconsistently. For example, to weight by age because it captures different social roles raises the question: why stop there, why not also weight according to, for example, a person's occupation or tax bracket or productivity (Anand and Hanson, 1997, pp691-2; Paalman et al., 1998, p17)?

- the one reason for stopping at age weighting is because an individual can expect to move through all the life stages whereas this would not be the case if ethnic, income or gender groups were used for weighting (Brock, 1998, p78). However Murray did not draw out this reasoning.

A number of authors are concerned about the possibility of double-counting the impact of age or are at least worried that the issue has not been worked through clearly from concept to method, results and interpretation. Bowie et al. (1998, p42) raise the question 'is the weight sought a function by age or individual's social responsibilities?' and worry also about the interaction between age weighting and discounting. Anand and Hanson (1997) ask for the exact question that was posed to the TB programme managers[63] and cited as part of the evidence for the values given to preventing a death at different ages. They wonder whether information had been elicited in a sufficiently careful way to be sure that the values reflect an intrinsic value of time lived at different ages and asked 'how do we not know that it is not reflecting their view of income levels and productivity through the life cycle?' (p692). If this were the case, then selecting people with a different expected pattern of lifetime earnings would probably alter the findings. Like Tsuchiya (1999, p268), they also question whether the value of living a healthy life at different ages was considered separately from the likelihood of disability. If, for example, the programme managers were properly informed about other adjustments in the DALY formula (for example reductions in human function) the values

---

63 This included 116 people from developing countries who had attended a tuberculosis training course over a period of five years Murray (1996).

given in age-weighting would also reflect expectations of future disability and hence double count the impact of disability.

Questions concerning the evidence basis for age weights can be considered in the light of Tsuchiya's recent literature review (Tsuchiya, 1999). She found some limited empirical evidence showing that people do value health benefits differently depending on the age of the recipient of health care. However, she also concluded that the evidence to suggest a particular rate for efficiency-based age weights is extremely limited. Given these contentions, and the suggestion that age weighting is not used consistently in cost-effectiveness analysis even by one of the developers of DALYs (Paalman et al., 1998, p17), it is not surprising that there have been calls for age weighting to be dropped, at least from the base case DALYs (Barendregt et al., 1996). Williams (1999) suggests that age weights could be treated as an optional extra.

4.1.3   Weighting future time

The questioning of weighting future time lived has focussed on why discount and, if so, at what rate and with what functional form?[64] Some critics have focussed on the nature of the good, for example stating that health is not commensurate with money as health cannot be re-invested elsewhere and so it should not be considered as a financial flow and therefore should not be discounted (DfID, undated). Some have focussed on the inconsistency of arguments used. For example, Anand and Hanson (1997, p89) have claimed that Murray's rejection of the human capital approach as a basis for age-weighting also implies a rejection of the view that health is commensurate with money and that therefore any of the arguments presented concerning future consumption are irrelevant in any justification of discounting.

---

64 The only views reviewed here are those raised in the context of evaluating DALYs. Other relevant arguments made in this area, but not yet directed specifically at DALYs, are the possibility of adopting alternative functional forms for discounting future life years saved (Cairns and van der Pol, 1997) and disentangling the effect of time preference from the effect of the duration of health state (Dolan and Gudex, 1995).

With respect to the social rate of time preference and inter-generational transfers, Anand and Hanson (1998, p309) disagree with the use of a positive rate of discount because it can lead to forms of environmental degradation as future lives are valued less than current lives. Weighting the same life or disease differently by time period does not accord with their 'intuition or common language'. They also question why Murray feels uncomfortable with the implications of a zero discount rate in recommending eradication of a disease, and wonder whether feeling uncomfortable is a sufficient basis for a decision or an excuse to fix results[65].

As expected by Murray, the magnitude of the discount rate has been questioned. Sauerborn et al. (1998) criticise the choice of 3 percent as arbitrary. They draw on the mixed empirical evidence to question the size of the discount rates, arguing that much more evidence for 3 percent is needed given the profound impact it can have on the ordering of burden of disease. In particular they suggest that community-based studies in countries across the world should be used to establish a rate rather than imposing a 3 percent discount rate across all societies. It would be interesting to establish whether a higher private rate of time preference for health does occur in places where individuals face a higher annual probability of death and secondly whether this is related to a higher social rate of time preference.

The functional form of the discounting element of DALYs has been questioned in two ways. First, Sauerborn et al. (1998), drawing on existing literature (e.g. Gafni, 1995; Johanesson, 1992), question whether the exponential function fully characterises an individual's time preference and suggest that it may depend on age but that more research is required to determine this essentially empirical question. Secondly, Elbasha (2000) has questioned the representation of time in the DALY discounting formula as a continuous rather than as a discrete variable. Three reasons are given to justify a discrete approach: the use of discrete time intervals is more common, especially in eco-

65 However, contrast this with their own arguments based on 'intuition'. It would appear that they feel uncomfortable with age weighting (Anand and Hanson, 1997, p698).

nomic evaluations; it corresponds to the way that health and economic data are usually reported (i.e. specified time intervals such as an hour, a day, a week, a month, a year); and discrete intervals are considered more plausible, because age standardisation formulas use discrete weighting functions. Elbasha reformulated both the age weighting and the discounting formula in a discrete-time version, and compared the estimates of DALYs lost due to cataracts in Africa in 1990. The impact of the reformulation is greatest in diseases of short duration and for older people. Thus DALYs as currently formulated are argued to underestimate the burden of disease in these cases, although it is not clear how the total burden for all diseases would change.

4.1.4 Weighting disability
The conceptualisation, measurement and valuation of disability is the area that has been subject to greatest debate. This section focuses on questions raised about whether disability should or should not be weighted, use of the PTO approach, the effect of the presence of comorbidity and whether 'adaptive' weights would be more appropriate. Other issues concerning the definition of health and disability and methods used to elicit values are discussed in the next section.

Several people have questioned the commensurability of death and disability. Anand and Hanson (1997, p689) take the softer critical stance, noting the convenience of being able to combine imperfect health and death in a single indicator, but stating that treating death purely as another health state results in a significant loss of information. They suggested that DALYs should always be presented separately as YLLs and YLDs (even if trade-offs between them could be demonstrated). This would have the advantage of allowing those who believe that the two events were incommensurable (e.g. Field and Gold, 1998, p17, drawing on Kamm, 1993) and that a life, whatever its quality, is always preferred to death, to adopt different decisions and to understand the impact of moving from one indicator to another.

Arnesen and Nord (1999a, p1425) develop a stronger critique and invoke the Declaration of Human Rights to argue that rights to life extension should be equal for all people and not dependent on the

level of disability[66] – the core positioning of the disability rights movement that policy analyses should not make distinctions among lives that are still worth living. In an emotive illustration of this view, Groce et al. (2000, p2) decry the alignment of people with disabilities as lying somewhere between life and premature death.

The particular problem with the Mark 2 DALY's approach to the PTO, of forcing experts into adopting a particular view of the value of disabled versus healthy lives, was set out in Arnesen and Nord (1999a). For example, the first PTO question asks about the relative value of extending the life of 1,000 sighted and 1,000 blind people (see Box 3 above). An expert may reasonably conclude that these are of equal value if they do not want to discriminate between the lives of disabled and able-bodied people, such that the resulting disability weight (1,000/1,000) = 1. The second PTO question is different (see Box 3) and does not require that a final weight of less than 1 is based on a supposition that the lives of disabled people are worth less than those of the able-bodied. Say an expert considers that relieving 5,000 people of blindness is as valuable as prolonging the lives of 1,000 healthy people. In this case the health state 'blindness' would be given a disability weight of 0.2. Therefore, quite logically, the PTO exercise results in two different disability weightings for the same health state: 1 or 0.2 according to which question is being asked. However, the current methodology does not allow this positioning and the expert is forced to make their valuations 'consistent'. Under the pressure of such forced consistency, an expert might end up selecting PTO1 = 1,100 and PTO2 = 11,000, which together yield a disability weighting of 0.09. As Arnesen and Nord (1994, 1999a) state, 'this weighting, however, does not correspond to any actual preference of the respondent' and they conclude that the disability weightings are merely 'artefacts' threatening the validity of the entire global burden of dis-

66 This was in response to receiving training in a workshop of European researchers working with DALYs, as the trainees were told that they were expected to give a higher value to saving healthy people than disabled people because if they were equal 'it implied that being disabled is as good as being non-disabled' (Arnesen and Nord, 1999a, p1424).

ease exercise. They suggest that the first PTO question should simply be dropped from the valuation exercise.

As part of any academic piece of research, it would have been interesting to be able to access information on the extent to which these issues were demonstrated. Unfortunately none of this data has been published. Perhaps this criticism is keenly felt, given that a respondent from one of the valuation exercises felt compelled to write about the experience (Nygaard, 2000). Not only did she indicate that people had raised such issues during the meeting she attended but had 'resisted' weighing the saving of lives of some against improving the health of others. She concludes that 'the validity of valuations made in the face of serious doubts and outright resistance and refusal to participate are worth questioning'. Secondly, she raises a related concern about whether the use of average scores are a good representation of any one individual with the disease as she hints that there may well have been bimodal distributions of values. These issues are compounded given the arbitrariness of the classifying the seven categories of disability. However, once again with no data presented on such issues, serious questions are raised about the meaning of the disability weights and what they represent.

Daniels' (1998, p66) suggests that what is needed is a different approach. He suggests that 'more complex methods that lead subjects through a series of questions that import arguments and reasons' should be the basis for making these trade-offs between quality and quantity of life because complex philosophical exploration is more likely to uncover evidence about why people weight as they do and not simply uncover unconsidered tastes. He states that:

> 'if a good moral argument persuades us that our original belief
> about what is right is in fact incorrect, we may be chagrined,
> but we are (or should be) grateful as well. We have been spared
> doing what is wrong.'

Sauerborn et al. (1998, pp9-10) accepted combining quantity of life and disability but suggested that the DALY could go further and capture the possibility of weighting the impact that death or illness (for specific diseases) may have in terms of the reduced chances of survival or quality of life on others (regardless of age). Their suggestion

is particularly relevant for infectious diseases. However, given the paucity of accounting for co-morbidity in DALYs (Mansourian, 1996, p334), it is unlikely that potential disease is likely to be considered yet.

One of the problems of not accounting for co-morbidity, pointed out by Sayers and Fliedner (1997, p383), is that any future intervention planned on the hypothesis of a single disease may fail to make the anticipated impact because DALYs are not 'logically linked to the realities of disease and disability'. Therefore there are concerns that use of DALY weights will overestimate the impact that health services can have because, once one cause of disease is eliminated, the DALY approach assumes that all people with that disease are treated successfully and return to zero DALYs lost (or to the value of the treated state) because it is implicitly assumed there is no co-morbidity.

That disability weights in DALYs deliberately do not account for how well a person may have adapted to their condition has also been a source of criticism. For example, Anand and Hanson (1997, p694) suggested that 'compensated disability weights' that capture the factors affecting an individual's capacity to cope with their disability (given their circumstances) would be more appropriate. These adaptive weights would more closely reflect the true burden of disability as experienced by the individual. They suggested that indices of income or the provision of transport services may be usable for adjusting disability weights[67]. Rock (2000 p412) also criticises the use of pre-adaptive weights because it causes over-emphasis on rehabilitation (rather than on modifying the person's physical or social environment) and on investment in people who do not have disabilities. She argues that if the DALY quantified 'adaptation', then interventions that extend the lives of people with disabilities would become more cost-effective in comparison with those aimed at preventing disability. Such measurement would be more acceptable to governments pursuing policies of social inclusion.

[67] Murray and Acharya (1997, p713) clarify the impact of choosing adaptive or pre-adaptive weights: using pre-adaptive weights will make prevention and rehabilitation more attractive in cost-effectiveness analysis, but life extension for those with paraplegia will appear to be less cost-effective. Using the adaptive weights will make prevention and rehabilitation much less attractive in cost-effectiveness analysis, but it will make life extension more cost-effective.

Finally, Paalman et al. (1998, p18) queried the approach taken to separate intensity and frequency of disability because this did not allow for any interaction of the two variables. Thus they suggest that permanent or temporary inability to walk may be more or less than a constant value multiplied by time. Perhaps if adaptation occurs over time it would suggest a diminishing or decreasing relationship between intensity of, and time in, a disease state.

4.2 Whose values are, and should be, represented?

4.2.1 Conceptions of 'health', 'disease' and 'disability'

The representation of DALYs averted as a measure of health gain and a desirable goal has been debated. Implicit in much of the criticism is disagreement about who is making judgements about whom and with what consequences. Some stress that DALYs are a measure of disease, and disapprove of its biomedical orientation towards conceptions of health (Barker and Green, 1996, p181; Hanson, 1999, p12; Kindig, 1999, p17). There is so much more to health than the absence of disease, for example the ability to call on support networks during sickness or the ability to think clearly and work through problems that might arise. As Mooney and Wiseman, (2000, p371) state, 'to assume that the BOD [burden of disease] sums it all up, or that whatever else is there is constant across the diseases in all these countries, is surely a false premise'.

Others focus on the limitations of using disability as the basis for the measure. For example, Wasserman (2000) notes the appropriation of the language of the International Classification of Impairments Disability and Handicap (ICIDH) group by Murray and raises two issues. First Wasserman argues that it is the consequences of limitations not the limitations themselves that determine the value of a non-fatal outcome. This view finds much support among those concerned that the DALY weights take no account of context (AbouZahr and Vaughan, 2000, p662; Hanson, 1999, p13; Brock, 1998, p73; Sayers and Fliedner, 1997). Secondly, Wasserman notes that even though the DALY is described in the language of the ICIDH, the 'disabilities' are actually described in terms of disease symptoms rather than common

activity limitations (such as difficulty walking or reading). Perhaps this discrepancy occurs because disability itself is never defined and is assumed to be an illness (Groce et al., 2000, p2).

This links closely with the beginning of a fascinating critique by Rock (2000), whose thesis juxtaposes the trial of Latimer[68] and DALYs to expose and compare implicit beliefs about disability. She begins her criticism with the observation that 'the DALY regards disability as pathology – as disease in need of elimination …. it deems time lived with disability as having less worth than time spent in 'perfect health" (Rock, 2000, p408). She states that, whilst there is no single biomedical model of disability, 'biomedicine does take the desirability of a 'normal' body as its starting point, yet 'perfectly normal' bodies seem in short supply' (p416). She contrasts this with the picture conveyed by DALYs that 'imagines 'normality' as limitless in human populations, and universal but for injustice. The DALY thus portrays 'premature' death and observable pathologies as preventable and even 'unnatural" (Rock, 2000, p415).

Part of the reason for the focus on disease can be gleaned from the view that an outcome measure should be able to measure the impact of 'health action', which Murray and Lopez (undated, p9) defined as 'any set of activities whose primary intent is to improve or maintain health'. This was exemplified by education. Whilst more education is related to improvements in health, they are not the primary aim of education services, and therefore education would not be included as a 'health action'. 'Health' is therefore defined, almost with circularity, as what is delivered by health services which try to control disease.

Musgrove (1997, p2) is concerned about this limited focus of DALYs, albeit from the opposite direction. He argues that improving people's health (through disease control) leads to many benefits on top of health itself, especially in children, such as increased opportunities for education and production as well as better health in adulthood. He therefore claims that measurement using DALYs under-represents the benefits of health care, and particularly so in children. He calls for

68 This was a controversial Canadian murder case, where a father killed his own child because of her severe disabilities.

an understanding of the trade-offs between, for example, graduating from another year of school versus being free of disease for three years. Block et al. (2001, p1538) are of a similar view and stress the necessity of conceptualising the whole range of benefits, to allow a shift of focus away from a sole concentration on the individual and their interaction with health services to a wider variety of actors and other benefits.

One of the strongest criticisms of the descriptive basis of DALYs can been seen in the response of the Dutch Disability Weights Project group to eliciting weights for disease states in The Netherlands. Their approach was to systematise the description of disease states[69] in terms of levels of functioning using an adapted version of the EuroQol group's instrument the EQ5D[70] (Brooks, 1996). An example is given below in Box 6. This also shows that, even though this project group gives the same title to a disease, the description of the disease differs from that in the original DALY exercise[71] and the dementia is also described as severe[72].

69 They also changed the range of disease states considered, to include fewer tropical diseases and more chronic diseases seen in affluent societies (Stouthard et al., 1997, p9).

70 See Appendix 2 for an introduction to the EQ-5D. The adapted version referred to here varies from the official EQ5D version in a number of ways: it is based on six, rather than five, dimensions by adding a dimension on 'cognitive functioning'; the levels of severity in some dimensions are mixed, e.g. described as 'no or some problems', whereas in the original version no problems and some problems are separate levels of severity; a percentage figure is used to indicate severity in two dimensions, which is not done in the EQ5D; and the wording of the 'self care' dimension is changed.

71 Dementia was described in the DALY exercise as 'An individual with multiple cognitive deficits that include memory impairment and aphasia (difficulty producing the names of individuals and objects) and apraxia (impaired ability to execute motor activities despite intact motor abilities, sensory function, and comprehension of the required task).'

72 This was because the researchers did not believe it was possible for raters to assign a single weight to the condition in its entirety because to achieve that it would be necessary for all respondents to have an equal command of all the following information: familiarity with all the sequelae; the contribution of each disability to total morbidity burden due to the condition and therefore knowledge about the prevalence, incidence and length of illness; and be able to arrive at an average weight for the entire disorder. Because researchers thought this was implausible (Stouthard, 1997, p13), diagnostic categories were instead 'staged' by severity levels.

Box 6 **Example of the descriptive system used for dementia in the Dutch Disability Weights Project**

DEMENTIA

Patients with severe dementia (permanent supervision required)

- Some problems in walking about (50 percent*) or confined to bed (50 percent*)
- Unable to wash or dress self
- Unable to perform usual activities (e.g. work, study, housework, family or leisure activities)
- No pain or discomfort
- Moderately (50 percent*) or extremely (50 percent*) anxious or depressed
- Extreme problems in cognitive functioning (e.g. memory, concentration, coherence, IQ)

Note: *The percentages varied for the different indicator conditions. However, neither a justification for their inclusion nor an explanation of their intended meaning was given in Stouthard et al. (1997).

The reason for Stouthard et al.'s (1997) departure from purely disease-based descriptions was that even though medical 'experts' were thought to have a real insight into the consequences of disease, it was considered 'strongly unlikely that all the members of the panel would have a comparable level of expertise' and the move towards generic descriptors was a way of standardising the information considered.

This idea has received some support from Nygaard (2000 p121) who had participated in one of the WHO valuation exercises (with just the disease based descriptors) and confessed that 'I caught myself using my own assessments of the various diseases, and not the ones provided … and I believe this might also have been the case for other participants. Hence it is difficult to know what was actually valued'. Whether this is more or less likely with the original DALY description of disease states or with the Dutch approach is a moot point.

4.2.2 The role of 'experts'

Without 'experts' DALYs would have crumbled. 'Experts' have been involved at all stages, from their inception, through to their design, estimation, valuation and interpretation. 'Experts' have been used to validate or refute numbers, conclusions and policies. Musgrove (2000, p111) wrote that the very tight deadline for producing estimates for the World Development Report (World Bank, 1993) meant there were no realistic options other than to include expert opinions. However, their involvement in particular stages of the construction of DALYs has attracted a lot of attention. There has been no criticism of their use in providing information about the distribution of mortality or morbidity of disease or even of the distribution of severity of disease across populations[73]. However, there has been particular criticism about reliance on them for providing values concerning the intensity of disease states, as well as for age weighting, discounting and the concept of health adopted.

Many critics state that the 'experts' were not an appropriate group of people to provide valuations of the impact of disease on disability because:

- 'experts' do not necessarily have better judgement (Rock, 2000, p414; Hanson, 1999, p20);
- 'experts' do not reflect societal preferences (Paalman et al., 1998, p18; Bobadilla, 1996, p10); or at least the assumption that they do has not been validated (Bobadilla, 1996, p10; Üstün et al., 1999, p111);
- 'experts' lack personal experience (Paalman et al., 1998) and therefore their values are likely to differ from those of the general population and those of people experiencing the particular difficulty (AbouZahr, 1999, p123; Nygaard, 2000, p118);
- 'experts' have 'systematic biases' (Brock, 1998, p75);
- there was no representation from an organisation of people with disabilities (Groce et al., 2000);
- values of life with disability differ between able-bodied people and people with disabilities (Groce et al., 2000);

---

73 This does not mean there has been no disagreement with the estimates given!

- medical professionals inevitably see need in terms of medicine rather than in social, economic or rehabilitation terms (Groce et al., 2000);

- the choice of 'experts' had precluded national/local values (Morrow and Bryant, 1994, p163).

Barker and Green (1996 p182) bemoan the fact that 'we are not told how these ('experts') were chosen; one might guess that every effort was made to span the regions of the world in choosing 'experts' of various nationalities, but we do not know if a similar effort was made too in terms of gender or class'. Indeed several authors have questioned why only 40 percent of the 'experts' were women (e.g. Hanson, 1999, p12) and have pointed out that, given that the 'experts' were public health professionals sufficiently well known to receive an invitation to an international consensus meeting in Geneva, it is unlikely that they were from lower social classes. In fact, AbouZahr (1999, p123) goes further and claims that DALYs 'can only be used by people with a relatively high degree of literacy and understanding of basic economic concepts. It is therefore by definition, an exclusive process driven by academic rather than experiential knowledge'. Nygaard (2000, p122) agrees and also highlights the limitation imposed by the need for panel members to speak English.

Nygaard's (2000) paper is particularly interesting because it allows us to consider how expert the 'experts' may feel. She writes that in practice when trying to make assessments of severity, 'one soon realises that the degree of severity, even for fairly objective dimensions of a health condition, depends on contextual factors'. She states that the global burden of disease project recognised this and therefore asked raters to consider the 'average social response or milieu' but she asks: 'Is it possible to describe an 'average social response or milieu', thereby including contextual issues, in a reliable way'?

If it is accepted that such values should not be predicated on 'expert views', then whose views should be solicited? The majority of papers suggest that either the general public (Williams, 1999, p6) or people with disabilities or who are patients (Paalman et al., 1998, p18; Hanson, 1999, p20) would be the most appropriate. The main arguments in favour of these alternatives are because the policy context

focuses on society in general and because of the value given to personal experience. Nygaard (2000, p122), however, points to the problems of using any of these views as a route to concluding that the whole exercise is too difficult and that a single index should be abandoned.

Perhaps one of reasons for the strong reactions against the use of 'experts' has been the 'express disavowal' of including community preferences (AbouZahr, 1999, p123). Rock (2000, p412) questions why Murray suggested that if people with disabilities report happiness it is because they suffer from false consciousness, especially given that no evidence was provided for such assertions (Wasserman, 2000). Rock (2000) went further and dissected Murray's comparison of slavery and disability, which she states 'implies that people with disabilities are unfairly trapped in their 'diseased bodies'…(and) that people with disabilities form a class of people who, like slaves, cannot provide valid information about how they experience the world'. Thus, all question whether Murray's 'expert' opinion on the role of 'experts' is sufficient.

4.2.3 The universality of disability weights

A number of researchers have voiced concerns over the assumption that the weights assigned to disabilities can be applied meaningfully across the world (Nygaard, 2000, p123; Sayers and Fliedner, 1997; McMichael, 1995, p555; Groce et al., 2000, pp4-5). There are questions both about why the assumption is made and how likely it is to be true.

Wasserman (2000) cites two propositions of egalitarianism[74] put forward by Murray and suggests that 'taken together, these propositions ensure that no personal or cultural characteristic of an individual (wealth, religion, intelligence, race, beauty or socio-economic status) shall be taken into consideration when determining the 'burden of disability' represented by his or her health status'. This approach runs counter to the conceptualisation of disability sponsored by the WHO in the International Classification of Impairments, Disability and Handicap, and does not even attempt to account for handicap, despite any claims by Murray to the contrary (Wasserman, 2000).

---

74 The second and third principles in Section 2.2.

'Why should we assume that one person's perception of her quality of life is commensurable with another person's, especially if the two people are separated by language and culture?' asks Wasserman (2000). The burden of disease can only be judged in the context of the culture of the individual, including the economic, family and social circumstances of the individual and as such it is likely to result in complex population-wide heterogeneity (Sayers and Fliedner, 1997, p383). The unpublished evidence cited by Murray that there was close agreement between groups of health professionals on the values assigned to the indicator conditions does not signify that real disability is uniform across countries (Üstün et al., 1999, p112).

Indeed, most authors draw on evidence to suggest the contrary. Brock (1998, p73), for example, considered that, in a setting in which most labour was manual, limitations in physical functioning would be likely to have greater importance than they would in a setting in which most individuals are engaged in non-physical, knowledge-based occupations, where certain cognitive disabilities are of greater importance. He suggested such differences would be 'magnified when summary measures of population health are employed for international comparisons among very disparate countries'.

Groce (1999) litters her text with additional examples of diversity and wonders why the DALY 'ignores an entire body of research and advocacy that has been developed over the past 20 years' based on the thesis that the impact of disablement is dependent on context. Using some of the same indicator conditions as the DALY Mark 2, Sadana (cited in Hanson, 1999) asked 40 women aged between 17-54 in Cambodia to give weights for disability for a range of health conditions from the global burden of disease. She found that most women gave higher disability weights than the global burden of disease exercise especially for those associated with stigma or shame for women (e.g. vitiligo on the face and infertility) and that this was consistent across education and age groups. Whilst it should be noted that Sadana used different methods to the global burden of disease exercise, she did conclude that it was important to assess community/population values.

Üstün et al. (1999, p111) appear particularly concerned about the presumption of universality when the numbers were derived from a

single PTO exercise carried out at the WHO in Geneva. They suggested (p115) that a further reason why DALY disability weights may not be universal is that in addition to the description of disabling effects of disease/conditions, valuations will also be driven by mood and public opinion. They called for 'more systematic testing across cultures and different informant groups, with other forms of measurement' and therefore asked groups of around 15 people from 14 countries[75] the same questions. The groups were comprised of:

- medical professionals (e.g. physician, psychiatrist, psychologist, nurse);
- allied health professionals (e.g. social worker, case worker);
- policy makers or opinion leaders in the area of disability services;
- individuals with a disabling physical condition and their care-givers;
- individuals with a disabling health condition related to alcohol, drugs or mental health, and their care-givers.

Results showed that the conditions at the extreme ends of the ranking exercise showed less variability between informants than for the middle-ranked conditions, which focussed mainly on conditions related to mental health. Üstün et al. (1999) also found that differences between countries in disability weights were statistically significant for 13 out of 17 conditions (but with no differences for quadriplegia, paraplegia, amputation below the knee, and mild mental retardation). Their main conclusions were that:

'the relative burden of different health conditions, in terms of disability, is fairly similar across the world …(but that)…there are systematic, and in some cases, pronounced differences between cultures and informant groups. These differences are large enough to shed doubt on the assumption of universality

75 | **High-income** | **Middle-income** | **Low-income** |
|---|---|---|
| Canada | Greece | China |
| Japan | Romania | Egypt |
| Luxembourg | Tunisia | India |
| The Netherlands | Turkey | Nigeria |
| Spain | | |
| UK | | |

of the disability rankings, and subsequently on the weights, and they are large enough to be investigated further in a systematic way' (Üstün et al., 1999, p115).

Such differences have led to suggestions that cross-country comparisons would be vitiated and interpretations, at best, rendered meaningless (Sayers and Fliedner, 1997). Nygaard (2000, p123) went further and suggested that it:

'could lead to governments using their resources for programmes that give the biggest improvements (i.e. reductions) in the burden of disease, as determined by the GBD [global burden of disease] project, in order to score well in WHO comparisons of national health improvements. However, national considerations might well indicate quite different improvements in health and health care'.

4.2.4 Gender bias

It has been acknowledged that the use of the DALY methodology has brought reproductive health (especially with regard to women) more prominently onto the global health and development agenda[76] (AbouZahr and Vaughan, 2000, p657). Yu and Sarri (1997) have used DALYs to document changing patterns of female diseases and causes of death. But, there has nevertheless been an interesting group of papers published presenting a gender-based criticism of DALYs. They centre on: the types and implications of missing data; gender sensitivities in the construction of the DALY; and the problems of focussing only on disease.

DALYs rely on existing reported data. Therefore, if that information is gender biased, then estimates of DALYs will be too. Sundby (1999, p281) argues that more health data are processed for men than for women and that, when there are no gender disaggregations, the norm is often the adult man. Under-reporting can be a particular problem in places, due to a 'culture of silence' (Sauerborn et al., 1998, p8) especially if deaths result from induced abortion or suicide

76 See Section 4.4 of this monograph for a summary of the critiques about using estimates of burdens of disease to fix policy agendas.

(Hanson, 1999). AbouZahr and Vaughan (2000, p659) note that many reproductive health problems only surface when there are serious complications or through direct questioning. Given that maternal causes of ill health are such a large proportion of the overall burden of disease, the imprecision of current estimates in maternal health undermine confidence in the estimates of women's overall health burden (Hanson, 1999).

Another type of under-reporting is built into the model of diseases. A number of absences from the list of diseases captured by DALYs were found by comparing the list with the definition of reproductive health provided at the International Conference on Population and Development (AbouZahr and Vaughan, 2000, p658). These omissions included gynaecological morbidity due to sexually transmitted diseases, menstrual disorders, female genital mutilation, psychological disabilities due to sexuality and reproduction and stillbirths. It was also noted that different forms of gender related violence are missing such as forced sex and female infanticide (AbouZahr and Vaughan, 2000) as well as the impact of contraceptives on women's lifetime health (Sundby, 1999).

A particular problem, highlighted in the section on conceptions of health, is the over-reliance on averting DALYs through medical interventions rather than broader multi-sectoral developments. If DALYs focussed on health rather than disease, resources would be more likely to be directed to poverty alleviation than medical interventions, and the increasing feminisation of poverty means that failure to address this issue adversely affects women, and increasingly so (Sundby, 1999, p282). Curlin and Tinker (1995, p345) draw on the marginalisation of older women, especially widows, in Bangladesh to highlight the relationship between social exclusion and chronic diseases or malnutrition and have called for disease to be viewed in its social context.

Finally, there have been some interesting positions taken about potential biases in the methods used to estimate DALYs, which might underestimate the true relationship between gender and health. For example:
- many issues in reproductive health lie between health and social well-being (e.g. empowerment and reproductive rights, circumci-

sion, domestic violence, illegal abortion or lack of labour regula-
tions for working pregnant or lactating women) and are therefore
not captured in an approach that favours distinct diseases (Sundby,
1999, p283);

- the estimated gap between the burden of disease of women and
that of men is compressed in two ways: by accounting and adjust-
ing for the riskier behaviour of men; and because the additional
gains in life years receive less weight in women due to discounting
(Hanson, 1999, p22);

- given that one of the reasons justifying age weighting was welfare
interdependence, the age profile used to reflect value of women by
age may be underestimated relative to men as young girls in many
societies have extensive responsibilities for childcare and food
preparation, and it is not clear that women's responsibilities for the
welfare of others decline as they reach old age (Sundby, 1999,
p283).

4.2.5 Human rights

The ethical positioning of the construction and use of DALYs has
been contested in two broad areas concerning human rights: the right
to life and the right to self-determination. Debates around DALYs
and the right to life range across: the right of a foetus to be considered
alive; the right to be born; and the 'rule of rescue'.

Lal (1994) contrasts studies implying very different values attached
to the life of a prematurely born infant (for whom death even after one
minute would mean a lifetime's DALYs had been lost) and the value
attached by both QALYs and DALYs to the life of an unborn foetus
(who, if it died prior to birth, has no impact on DALYs lost). He
shows that it matters crucially when a person comes into existence, as
once a person exists then all future DALYs and QALYs count but,
until that time, none do. He comments:

'the beginning of a person seems inherently vague, so it seems
wrong to attach great importance to the moment when it
occurs … if we doubt there should be a large difference
between the value of saving a premature baby and the value of
saving a foetus, that may make us doubt that the value of sav-

ing a baby is really all the QALYs in the rest of her life' (p7).

This view is part of a much wider debate but it does allow consideration of how well the assumptions of QALYs and DALYs reflect national or state laws on the right to life, whatever one's position. For example, the UK's current legal position allows pregnancies to be terminated[77] but only on certain grounds[78], and it is clear that the positioning of DALYs lies with the 'pro-choice' rather than 'pro-life' (anti-abortion) campaigners.

Considering the role that measuring DALYs has on determining the cost-effectiveness of pre-natal testing, highlights the ethical judgements being made about the quality and value of human life. This is particularly significant given recent announcements of the mapping of the human genome. If genetic tests will enable parents to identify foetuses susceptible to disease or disability, then the tests may also have a more 'eugenic' connotation for public policy. It will mean that health care professionals can identify children and adults who are currently healthy but who have particular susceptibilities and assign a discounted value to their lives. With respect to the future birth of children, it could mean that children likely to be born with disabilities have less right to life[79].

The 'rule of rescue' is the third argument raised in connection with the right to life. Glick (1994, p179) argues that when an identifiable person's life is threatened by a treatable condition we 'will almost

---

77 By a registered medical practitioner, in a hospital or licensed clinic (except in emergencies).
78 Including where: the pregnancy has not exceeded its 24th week and the continuance of the pregnancy would involve risk, greater than if the pregnancy were terminated, of injury to the mental and physical health of the woman or of any existing children of her family; the termination is necessary to prevent grave permanent injury to the physical or mental health of the woman (no time limit); the condition of the pregnancy would involve risk of grave physical and mental injury to the woman or to the life of the woman, greater than if the pregnancy were terminated (no time limit); or there is substantial risk that if the child were born it would suffer from such physical or mental abnormalities as to be seriously handicapped (no limit). See http://www.gn.apc.org/nac/newsite/reference.html
79 For example, there would be less reason to allocate health resources to a pregnant woman with a child more likely to be disabled than not, other things being equal (Wasserman, personal communication).

invariably refuse to stand idly by'. Because DALYs have been equated with reflecting preferences in decision-making, then the DALY should account for such rules. As they do not, it is argued that prioritisation exercises based on cost per DALY averted are likely to fail (Glick, 1994). However, on this point, Musgrove (1994) draws the analogy of doctors choosing how to allocate their time to a group of people each considered a medical emergency. Because this happens, it is argued that it is not unethical *per se* to ask 'How long will the saved individual live as a result and in what condition?' and thereby to choose who to treat. Nevertheless, he believes that, these questions tend to arise for very ill patients who are already hospitalised, when it is urgent to consider how much 'life' can really be 'saved' by further interventions, taking into account both quality and duration. In those circumstances, he argues that 'cost-effectiveness is a reminder that the rule of rescue should not be invoked to justify every effort to keep a disabled person alive for only a short while longer'.

Regarding rights to self-determination, Wasserman (2000) cites the view that the great majority of people with disabilities regard their lives as well worth living and to suggest otherwise is 'disrespectful in its exaggeration of their burdens and its disregard of their testimony'. It also clearly contravenes notions of individual consumer sovereignty if such values are used to control private decision-making within a public health system. Indeed a classical liberal positioning would lead one to question the right of the state to choose on behalf of an individual. 'If people knowing the risks choose a risky and legal activity with a high probability of shortening their lives, and bear its financial costs themselves, their decision is to be respected' (Lal, undated, p3). However, when individuals do not bear the costs themselves and are subsidised by the state or others[80], the issue becomes clouded. This becomes even more difficult if one purpose of state intervention is to focus on equity, such that questions like 'When should we allocate resources to produce 'best outcomes' and when should we give people fair chances at some benefit?' are raised (Daniels, 1998, pp58-9). Given that there are no principled solutions yet for these issues, it is

---

80 Similar to the moral hazard problem with insurance.

not surprising that there is concern that the results from the global burden of disease exercise are being used as a replacement for ethically sensitive deliberations across the world (Daniels, 1998).

4.2.6 Integrating equity[81]

Gilson (1998, p1892) argues that any health care allocations disproportionally favouring disadvantaged social groups, such as the poor, are acceptable with DALYs only to the extent that they maximise aggregate health status gain given existing resources. She contrasts this with an equity-driven approach that promotes a broader basis for decision-making in three respects:

> 'First, it requires consideration of the broader policy action required to promote health equity. Second, the concern for equity also points towards the broader policy goals that might be pursued through health policy. Third, the pursuit of equity forces consideration of decision-making procedures in society, and the extent to which they allow for broad representation and so expand choice' (Gilson, 1998, pp1893-4).

Pro-equity policies therefore emphasise 'the need to redistribute societal and health resources to address the particular needs of the poor and other vulnerable groups such as the old and chronically ill' (Gilson, 1998, p1894). However, an allocation rule based on minimisation of DALYs lost for a community will give less weight to saving the life of a disabled person than of an able-bodied person. In doing so, 'the policy of giving priority to saving the able-bodied would be to compound the disadvantage of the disabled: those who are already worse off because of a disability would be made more worse off because of being discriminated against in the allocation of heath care. It heaps a handicap on a person who is already worse off' (Sen, 1997, p217)[82].

81 Clearly equity could relate to many factors such as age, gender, degree of disability etc. As much of this has already been covered, I focus primarily on equity of income.
82 See pages 214-218 of Sen (1997)for a formalisation of the argument, based on ordinal comparisons, which shows that pursuit of DALYs increases inequality of capabilities and that the converse (a compensatory policy of giving more care to disabled people) will decrease inequality.

Drawing on some of the arguments from the disability rights **93** movement, that DALYs discriminate against those unable to return to perfect health, prompts consideration of whether this can also be said of the poor. If a poor person is less likely to return to full health than a rich person (other things being equal), then resources based on maximising health will discriminate against the poor. If so, then Gilson's critique is particularly strong, considering that it is the rich elite who have been involved in developing and valuing DALYs.

Others have been at pains to argue how the two objectives of equity and efficiency can work together. For example, Musgrove (1994) argues that equity and cost-effectiveness are entirely compatible goals in the global burden of disease exercise, given certain conditions[83], although he does point out the discrimination against older people because they have less life remaining than young people, and against people who live in less accessible places because of the greater marginal costs of their accessing health services.

One of the potential solutions to some of these issues is to create an additional set of explicit equity weights. To date there have been three suggestions:

- using the concept of a 'fair innings' based on age, to replace the weights implicit in the global burden of disease exercise Williams (1999, p7);
- using weights related to income levels (DfID, undated);
- societal valuation could be reflected as a multiplication of: change in utility from an intervention; weighting for the severity of the initial condition; and a weight for 'potential for health'[84] (Nord et al., 1999).

83 These include when: two people with the same health problem at the same age are treated identically and the only distinction is the expected survival of men and women; no account is taken of people's income, because what is measured is health rather than its consequences for productivity; there is no discrimination by race or other factors; and no account is taken of co-morbidity or of the health risks that face a person beyond the immediate problem for which the cost-effectiveness of treatment is considered. Under such circumstances, Musgrove (1994) argues that the health of an African child is not judged to be less valuable than that of a child in a rich country.

84 This is designed to counter the discrimination between patients with different capacities to benefit associated with QALYs and DALYs. Basically, it devalues large health gains relative to smaller health gains.

4.3 Quality of data

Early critics voiced concerns over the availability of the different aspects of data required for the calculation of DALYs. For example, in his review of Murray and Lopez (1994a), McMichael (1995, p555) felt that:

> 'much of the assignment of a disability-time score to various non-fatal illnesses has relied on guesswork, since there are few systematic data available to assist in this Herculean task'.

Most critics concede (see for example, Sauerborn et al., 1998) that Murray and Lopez are well aware of the deficiencies of data[85] (probably more so than the critics) and that, as Paalman et al. (1998, p15) wrote, it 'is an ongoing exercise, and data are improving as time goes by'. Despite this, there appears to be a general feeling of unease as the sources and implications of all the assumptions and estimations are not clear (Sauerborn et al., 1998, pp1-2). One improvement called for by Paalman et al. (1998, p16) was for confidence intervals, or at least upper and lower limits, to be attached to the estimates of global burden of disease. Alternatively, as 'experts' had often provided guesstimates for the range of required data along with an indication of how reliable they felt their estimates were, it would have been interesting to see such information presented. Without this, the quality of all estimates is assumed to be the same.

More recently, AbouZahr and Vaughan (2000, p662) have made a number of recommendations for improving the quality of data in the area of reproductive health. In addition to developing descriptions and measures of reproductive health as part of a cross-national research agenda they suggested improving the underlying epidemiology by:

- setting up a series of country case studies bringing together and analysing all the available information on reproductive health;
- standardising analysis of existing demographic and epidemiological data collected in longitudinal studies;

85 See for example, Lopez's own critiques of cause of death statistics in Lopez (1987; 1990) and Ruzicka and Lopez (1990). Musgrove (2000, p114) writes 'it is probably safe to say that every single issue raised by Paalman et al. (1998) was debated extensively, and sometimes acrimoniously, within the Bank and WHO; arguments between orthodox economists and public health specialists were often particularly sharp'.

- funding longitudinal studies to detect the incidence of reproductive ill-health and the risk of subsequent complications.

4.4 The usefulness of DALYs for decision-making

4.4.1 Costly and time-consuming to estimate

A number of authors have commented on the resources required and available to undertake the calculation of DALYs for burden of disease exercises and cost-effectiveness analyses. Concern has been voiced that burden of disease exercises are 'cumbersome' (Hyder and Morrow, 1999) as well as being 'expensive and time consuming, often taking key ministry personnel away from vital day to day activities' (DfID, undated). This mirrors concerns of Barker and Green (1996, p182), who asked whether DALYs were worth the bother:

> 'in some countries resources are so short that even the most minimal health care programmes are beyond reach. In that situation, can health planners feel justified in standing back and carrying out the necessary research to construct DALYs?'

Others have gone further and suggested that the wide range of scientific, analytical and technical capabilities required for policy analysis 'are rare in most ministries of health in less developed countries', and required much support and capacity building in such countries (Mosley, 1994, p28). There is evidence to suggest that, even when exercises have been undertaken, policymakers in Uganda perceived the burden of disease methods to be complicated and difficult to understand[86] (Kapiriri and Norheim, 2000). Given these difficulties, it increases the likelihood of Barker and Green's (1996, p182) concern that DALYs will be imported off-the-peg, so that global or regional estimates will simply be applied to an individual country for decision-making rather than estimated locally.

Budgets and resources obviously differ across countries. In the high-income setting of Australia it is interesting to note that the burden of disease exercise was not believed to exceed a 'modest' one to two

---

[86] On the other hand, the advantages of the DALY method included its helpful role in advocacy and resource mobilisation, and in providing criteria for resource allocation.

person years of available analytical resources (Vos and Mathers, 1998, p516). However, Vos and Mathers (1998) also felt that far more resources needed to be invested in the cost-effectiveness analysis.

4.4.2 International versus national priorities

It is helpful to remember that the World Bank[87] was the major funder of the 1993 World Development Report (World Bank, 1993) that launched DALYs globally. The World Bank has become the leading source of funding in the health sector amongst international agencies (see Ugalde and Jackson, 1995, citing Buse; Reich, 1995, citing Walt, 1994, p494) alongside its role as a development institution. Therefore the desire for some form of international prioritisation exercise was a key factor in the thinking and process of development of the global burden of disease project and DALYs.

Subsequently there has been much emphasis given to extending the global burden of disease framework to national settings to aid prioritisation of national health expenditures (Bobadilla and Cowley, 1995; Lozano et al., 1995). However, critics have questioned whether the DALY approach can ever be used to reflect a nation's priorities. For example, Ugalde and Jackson (1995, p527) have argued that, as the World Bank is not a democratic institution, its controlling interests lie with the largest stockholders, principally the US, UK, France, Germany and Japan. They question how such a bank can both profit from and assist the developing world[88]. Any prioritisation exercise

87 In conjunction with the Rockefeller Foundation and Edna McConnel Clark foundation.

88 It is important to note that the World Bank is not set up as a profit making institution, although the International Bank for Reconstruction and Development has earned a net income every year since 1948 (which the Board of Governors decides how to allocate to the reserves, as surplus, or for distribution). The International Bank for Reconstruction and Development accounts for around three quarters of the World Bank's international lending and it is here that voting power is linked to members' capital subscriptions. The International Development Association accounts for the remaining quarter of the World Bank's lending and is directed at interest-free loans (known as 'credits', but charged at a 1 percent rate of interest, to cover administration costs) to those countries with a Gross National Product per person less than US$885 and who lack creditworthiness to borrow on market terms but who demonstrate good policy performance (judged on more than 20 criteria). Readers are directed to the World Bank website for more information: www.worldbank.org

set up and supported by the World Bank, they argue, must necessarily reflect (at least partially) the desire to profit from the future. Others have argued for democracy and accountability to national populations rather than international organisations:

> 'in my opinion, national priorities on the use of health care resources should be based within each country, and be determined by open debate and a democratic process. In this debate, the global DALY weights might be one piece of information. Decision-makers, however, should be accountable to the people of their own country for their decisions and not to WHO or the World Bank or any other international organisation' (Nygaard, 2000, p123)

It is important that national settings are considered aside from global decision-making agendas not only for democracy and local accountability but also because national settings have different epidemiological and economic needs that are not recognised in the gross aggregation of world or regional data. Thus the burden of disease as well as the costs and effectiveness of interventions will differ and the 'cultural, economic, political, environmental, infrastructural and behavioural differences will all impact on the package finally chosen' (Paalman et al., 1998, pp14-15). So, Paalman et al. ask, how can international prioritisation exercises set relevant national priorities when the results of any prioritisation exercise are 'location, time and group specific'? Murray et al. (1994, p189) imply that this is merely a function of data availability and recommend moving towards cost-effectiveness production functions as a way of understanding and increasing the relevance of cost-effectiveness ratios produced in one context to another. Whilst this is certainly a useful way forward, it does not necessarily question the nature of the production function across countries and to do so would require research nationally rather than simply the global application of models of production functions.

4.4.3 Burden of disease versus cost-effectiveness analysis

Separately there have been some particularly strong criticisms of the use of DALYs in the burden of disease approach to decision-making. The criticisms have mainly been directed at the approach rather than

the DALY *per se*, but at times it is difficult to separate out the two because the DALY focuses on disease. Both Williams (1999; 2000) and Mooney and Wiseman (2000) have called for an end to the use of burden of disease exercises as a resource allocation procedure for health services and for setting research priorities. Both advocate the use of economic evaluations based on available technologies, with Mooney and Wiseman (2000) also calling for a more socially inclusive view of outcomes.

Their argument stems from the premise that burden of disease studies consider only part of the problem and have nothing to contribute to the solutions. Monitoring disease is mistaken, they argue, and money should be re-directed to assessing the impacts of interventions. Williams (1999) argues that there 'is no link, either in theory or practice, between the total burden of disease and our capacity to benefit'. He highlights 'the crucial difference between conceptualising a problem in terms of totals and averages, and conceptualising a problem in terms of what can be done at the margin'. The latter is where information gathering should be concentrated and is what, he says, 'we so desperately lack'. We should only be interested in interventions, their effectiveness and cost, but that has nothing to do with total burden of disease. This view has also been accepted in a recent guide for DfID staff (DfID, undated) that concluded:

> 'just estimating DALYs lost through burden of disease is not
> very helpful. It tells us what we already know or suspect. At
> best it might be carried out periodically to assess overall health
> progress. At worst it may be an expensive luxury which poor
> countries can ill afford'

Whilst Murray and Lopez have not viewed burden of disease results as the sole input to decision-making, they have also noted the forced reliance on estimates because of the paucity of evidence on the cost-effectiveness/benefit of health sector interventions. This is not a good enough reason, argues Williams, and indeed continued funding of burden of disease studies may become a self-fulfilling prophecy as scarce research money is taken away from economic evaluations of health interventions (Williams, 2000, p83).

Williams (1999) has accepted that burden of disease estimates

could be an appropriate tool to consider how money for research and development should be spent. He writes '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' (p2). However, he has three criticisms: first, that the argument falls if the objective of the health system is also to alleviate symptoms or arrest disease progression; second, that this approach does not account for desires to reduce side-effects of drugs; and finally that it does not account for the cost of the expected impact of the research in reducing overall burden. He suggests (1999, p31) that:

'(a) cheap and successful …(research programme)... attacking a disease which does not impose a major burden in terms of the health of the whole population, may nevertheless do more good for population health than an expensive and unsuccessful research effort directed at some disease which imposes an enormous burden in whole population terms'.

5 PLANS FOR THE FUTURE DEVELOPMENT OF THE DALY

There have been a number of indications in the literature that the DALY is undergoing further development. Bevan (1998, p1404) and Arnesen and Nord (1999a, p1424) have written that a new version of the DALY is expected to be launched in 2003. Murray and Lopez (undated, p15) and Murray, Salomon and Mathers (2000) have outlined ongoing work surrounding the definition and valuation of disability in five projects. One has involved modifying and refining the protocol for eliciting weights, using a convenience sample of international public health practitioners[89]. The modified protocol has been used amongst ten groups based in the US, Mexico, Brazil, the Maghreb countries (Morocco, Algeria and Tunisia), Japan, the Netherlands, and four multi-national groups of health care practitioners. They have provided valuations for 15-22 disability states – with a set common to all exercises – using a multi-method approach with internal consistency checks and group discussions (Murray and Lopez, undated, p15).

The Disability Weights Project for Diseases in the Netherlands[90] (Stouthard et al., 1997; 2000) and the European Disability Weights Project constitute further examples of DALY development. The European Disability Weights Project began in March 1998, following completion of the Dutch project, and was funded by the European Union. The aims included:

- establishing a comprehensive list of the disability weights associated with the various diseases that constitute the major part of the burden in Europe;
- improving, refining and validating the methodology for estimating the disability weights;
- investigating the cross-national stability of the disability weights (interpreting the differences if found);
- comparing the estimates of burden of disease across countries[91].

---

89 No details of the changes are given.
90 Also see http://www.bisociety.dms.it/Projects/BIOMED/Disability_weights.htm
91 See the website www.eur.nl/fgg/mgz/mgzprojects/BIOMED.html for the European Disability Weights Project Newsletter.

This project has partners in England[92], Sweden[93], Spain, Norway, the Netherlands and Denmark. Early correspondence on these projects highlighted dissatisfaction with the PTO questions as they stood[94] – see Box 3 earlier – (Arnesen and Nord, 1999b; Essink-Bott et al., 1999) and the group has made changes, using only one person trade-off question and dropping the requirement of forced consistency.

The final two related ways in which the measurement and valuation of DALYs are being developed include: a multi-informant validation of the global burden of disease weights, in 14 countries (Murray and Lopez, undated, p15); and the WHO multi-country survey of health and responsiveness (Üstün et al., 2001). Whilst Murray and Lopez gave no details of changes to the descriptive system or to the protocol for the first study, two were helpful in identifying the rough direction. AbouZahr and Vaughan (2000, p660) indicated that the following domains had been selected for a pilot study to measure health status in developing as well as developed country settings: mobility; self-care; usual activities; pain/discomfort; anxiety/depression; cognition; sensory; energy/vitality; shame/embarrassment; participation (or overall handicap in a given environment and socio-economic context). Nygaard (2000, p119) also indicated the likely adoption of a modified EuroQol format 'where the health condition is defined by 5-8 psycho-social and physical dimensions …(at) five levels of severity', which she states 'were put together by panels of 'experts' on different disease conditions, under the auspices of WHO, in order to obtain consensus among a group of 'experts". Box 7 shows an example of a new disease state piloted. It can be seen that there has

---

92 See http://spp3.bham.ac.uk/hsmc/HEF/hef%20research%209.htm for an outline of this part of the project.
93 See
http://www.phs.ki.se/socmed/research/se/projects/estimating_disability_weights_.html for the Swedish research project outline.
94 Arnesen and Nord (1999a, p1424): 'In the new procedure, the devaluation of life in disabled people was made explicit. An international research group that intended to use the procedure to establish disability weightings for Europe recently became aware of the offensiveness and lack of validity of the method and is now finding a new valuation method'.

Box 7 **Example of a new DALY disease state being piloted**

Description and dimensions of cervical cancer

Cervical cancer, marked by irregular bleeding, pelvic pain and malodorous vaginal discharge*

X	A few problems with moving around
X	A few problems with self-care
XX	Some problems with performing usual activities
XX	Some pain and discomfort
XXX	Much anxiety or depression
	No cognitive problems
XXX	Much shame and embarrassment
XX	Some limitations to social participation

Note: *The *X*s represent the degree of severity of each dimension of the condition on a scale of (blank) = no problems, to a maximum *XXXX* = severe problems.

Source: Nygaard (2000, p119).

been a return to a dimension based description, but that the levels of severity have been standardised, and that there are a total of 25,000 possible states using this example. This approach draws heavily from the framework of the EuroQol EQ5D questionnaire.

Sadana et al. (2000, p62) later outlined a whole range of potential dimensions considered by WHO as a basis for separating out 'health' (e.g. pain) and 'well-being' (e.g. participation in usual activities). It was argued that these ideas were differentiated because 'the ability to engage in usual activities does not describe non-fatal health *per se*, but limitations or performance in this area may be associated with a non-fatal health state' and they continued that 'we would prefer to assess non-fatal health directly'.

The state of play was set out recently in a series of discussion papers which form part of the development of WHO's survey of health and responsiveness (Üstün et al., 2001; Murray et al., undated; Salomon et al., 2001; Sadana et al., 2002;). The domains that formed

Figure 6 **Domains for assessing health and related well-being**

Health domains	Health-related domains	
● Vision	● Self-care	*Daily activities including eating*
● Hearing		
● Speech	● Usual activities	*Household activities; work or school activities*
● Digestion		
● Bodily excretion		
● Fertility	● Social functioning	*Interpersonal relations*
● Sexual activity		
● Skin & disfigurement	● Participation	*Societal participation including discrimination and stigma*
● Breathing		
● Pain		
● Affect		
● Sleep		
● Energy/vitality		
● Cognition		
● Communication		
● Mobility and dexterity		

Source: Üstün et al. (2001, p12).

the basis of the newly developed questionnaire based tool for survey interviewing are outlined in Figure 6.

The questions chosen to represent each of these domains were taken from existing survey tools, with many coming from the WHO Disability Assessment Schedule (WHODAS II)[95] domains. This questionnaire was piloted and then implemented in 71 surveys across 61 countries, using samples of between 5-10,000 respondents. Within the exercise 11 health states were valued using a visual analogue scale where 0 is death and 100 is perfect health[96]. These states

95 The WHODAS II covers six domains of health; cognition, self-care, mobility, interpersonal relations, daily activities at work and in the household, and social participation and impact. Further details can be found on the website www.who.int/m/topics/whodas/en/index.html
96 Separate surveys using multiple methods for valuing health states (including PTO, time trade-off and standard gamble) were also conducted on samples with high levels of education.

included the respondent's own health and ten other 'health conditions', based on a seven domain health state description (including; mobility, self-care, pain, affect, cognition, usual activities and vision). Initial results concerning validity, reliability and cross-cultural comparability[97] are currently being produced. The current commitment of the WHO is to launch a World Health Survey to monitor health and performance of health systems over time using a revised version of this questionnaire instrument (Üstün et al., 2001, p59).

Finally, Murray et al. (2000) have made a tentative suggestion that new summary measures of health could be developed if an approach applied the principle of the veil of ignorance (Rawls, 1971) to choose which of two populations (with differing characteristics) an individual would prefer to be in. Murray proposes that the population that the individual chooses is considered healthier. Nord (1999) offered an operationalisation of this idea for gaining a weight for severe asthma. He assumed two equal sized groups of people, some of whom are healthy, some of whom have asthma, and some of whom have a fatal disease, as outlined in Table 8. Each person would be asked 'Behind a veil of ignorance, to which cohort, A or B, would you rather belong'? This was described as a variant of the PTO and probability trade-off based standard gamble technique[98]. If the numbers in Table 8 represent the median point of indifference, such that 50 percent of respon-

97 Of particular note within this exercise has been the attempt to make results across countries comparable. This stems from Murray's repeated concerns that self-reports of health are an unreliable basis for comparison because they do not reflect objective physical measurement and because any valuation exercises are biased systematically because people use scales differently in different countries (Murray et al., undated; Saloman et al., 2001; Sadana et al., 2000). 'Calibration tests', such as using vision charts to test near and distant vision and tests of verbal fluency and recall, are being used as 'objective' measures used to adjust for 'biases' in self-reports. 'Standard case vignettes' are also being valued to compare how the same health condition is rated by people in different settings. Clearly the practical, conceptual and statistical mechanics of these approaches need to be evaluated when full details emerge.

98 The standard gamble technique offers respondents either (a) a certain outcome health state I for t years; or (b) a probability p of full health for t years but with a probability $(1-p)$ of immediate death. The probability p is varied until the respondent is indifferent between (a) and (b), and this value of the probability p is used to weight health state I.

Table 8 **Scenario for valuing different distributions of fatal and non-fatal disease**

Cohort	Healthy	Asthma (at specified level) from age 40	Fatal disease	Sum
A	80%	20%	0%	100%
B	95%	0%	5%	100%

dents would rather be in cohort A and 50 percent would rather be in cohort B, then the disability weight for suffering asthma from age 40 onwards would be calculated by dividing 5 by 20, i.e. 0.25, and subtracting the result from 1 (assuming that 'healthy' is set to 1), i.e. 0.75 in this case.

Nord points out that the valuations made in this way could incorporate concerns that an individual has for their own risk of fatal disease versus asthma as well as the distribution of healthiness across the population, and suggests that respondents need to be directed to think in one of these directions.

6 DISCUSSION

The following discussion compares the relative merits of DALYs and QALYs and debates which might be a better tool. I propose that the answer will depend on who is using the measure, where, for what purpose and from which viewpoint. I pose three questions:

- are DALYs more widely applicable than QALYs?
- are DALYs better than QALYs as an outcome measure in economic evaluation?
- when are DALYs or QALYs more likely to feed into decision-making?

Contemplating QALYs, I specifically consider the EQ5D approach to their calculation[99] because it is one of the few generic measures which bases its index values on population derived preferences for alternative health states in relation to death and it is the only health index developed from the beginning for use in more than one country.

6.1 Are DALYs more widely applicable than QALYs?

The response to this question draws on how QALYs and DALYs have been (and can be) applied in practice, as well as how widely held the conceptual basis of either instrument is likely to be. Section 3.1 showed that DALYs lost due to burden of disease were most consistently applied by the original developers of the measure. However when others used them as an outcome measure in cost-effectiveness analyses there was much variability in the assumptions and values used. Thus their use in cost-effectiveness analysis would appear to suggest that a DALY is not a DALY is not a DALY. DALYs differ in the age weights and discount rates used, as well as the sources and methods of deriving disability weights. This latter re-scaling has some resonance with Weinstein's (1988) concerns about QALYs that disease specific re-scaling of QALYs over-sensitises them for the purpose of decision-making. Thus both DALYs and QALYs have a potential for variation in their application.

---

99 See Appendix 2 for a description of this instrument and the approach to deriving the quality weights for QALYs.

There may be more or less variation in the types of DALYs used in practice compared with QALYs. More variation might be expected with DALYs because more assumptions are contained in the 'base case' that people might choose to change, for example the weights for age and future time. As many authors do not make their assumptions explicit, this additional scope for variation causes problems for comparison. Variation is less likely to occur in QALYs because assumptions for weighting time, for example, are generally presented separately, e.g. as 'QALY s discounted at x percent'. However, less variation may occur in the way DALYs are calculated in practice because:

- a smaller range of people are calculating DALYs. The WHO or Harvard University global burden of disease groups and the World Bank have been involved in the majority of applications to date;

- DALYs are used more frequently for national prioritisation exercises undertaken at one point in time, whereas QALYs are more frequently calculated in studies of the cost-effectiveness of individual health interventions over time;

- the focus of the DALY is on disease and disability rather than on a broader concept of health or well-being;

- the DALY has only used one method of valuation and one set of valuations whereas EQ5D based QALYs have used several approaches with a variety of general populations in different countries;

- the DALY burden calculations are based on an assessment of one intervention rather than of a series of potential interventions[100].

Whether DALYs or EQ5D QALYs cover a universal definition of health or have a more universally held conceptual base is an interesting question. Both measures have been applied in several countries, although only the DALY has been used as a global measure. However, wider application does not necessarily reflect universally held views. If DALYs focussed exclusively on disease rather than disability or some aspects of handicap, it might be easier to argue that DALYs had

---

100 However, it is curious that we do not know what the DALY based calculations for treated and untreated populations include as the intervention. It is also curious that treatments and the effects of treatment are considered to be the same across the world, with the only variation occurring in rates of treatment.

108   focussed on a conception of health used globally (albeit a very reductive conception) and practised widely in hospitals and health centres across the world. However, moving from descriptions of disease to valuing the impact of disease on disability or handicap changes this. It also brings the DALY closer to the QALY. Indeed it is interesting to see how close to the EQ5D questionnaire layout the new questionnaires for operationalising the measurement of DALYs are.

It therefore becomes important to consider the approaches taken in developing both instruments. Most notably, neither instrument has evolved from a study of conceptions of health in the countries within which the instrument was designed for use. At best they have relied on a review of past literature. Therefore the views of researchers predominate in the selection of the conception and description of health adopted in both instruments (Fox-Rushby and Parker, 1994). Indeed, Murray states '*I* have chosen to measure disability not handicap' (my italics) (Murray, 1996a, p33).

Starting from researcher based interpretations of 'health', both QALYs and DALYs then value descriptions of health/disease states. The DALY relies on 'experts' and the QALY more frequently on random samples of different lay populations. Williams argues (1999, p53) that the ability to develop local scoring systems is one of the strengths of the EQ5D QALY[101]. However, I doubt whether either approach, based on only local valuations can ever be considered universal (or even widely held) given that the conception and definition of health adopted in either instrument did not attempt to access a variety of views across the globe in the first place.

It is interesting to note that the EuroQol group, whilst supporting the translation of the EQ5D into many languages, has moved to the position where it argues that the translation guidelines it produces can only be used to establish semantic equivalence and not, for example, conceptual equivalence (Herdman et al., forthcoming). Indeed members of the EuroQol group have written about their concern that the cross-cultural equivalence and universality of health-related quality of

---

101 Although I understand that with the new descriptive system planned for the DALY, regional based scoring systems will be developed for that measure too and that national valuations will also be possible.

life questionnaires is assumed rather than investigated (Herdman et al., 1997). They later offered an approach to help assess the extent of equivalence of health-related quality of life measures (Herdman et al., 1998). The estimation of DALYs Mark 2 did not confront these issues as the valuation exercises were developed only in English and the international 'experts' who provided values used English as the common working language during valuation. However, this does not imply that conceptual interpretations were identical. Secondly, given that DALYs are now being operationalised as a survey tool, all the same issues concerning translation, equivalence and testing will apply. Unfortunately the WHO group appear to be concentrating on scale rather than conceptual equivalence, which is the last form of equivalence that should be tested (Herdman et al., 1997; 1998).

The extent to which either QALYs or DALYs represent universal conceptions of health is likely only to be an accident … or is it? In a series of essays written from a universalist's perspective, Wa Thiong'o (1993) suggests that what is written about or propounded as 'universal' or 'global' is often highly questionable. He writes, 'I am suspicious of the uses of the word and the concept universal. For very often, this has meant the West generalising its experience in history as the universal experience of the world' (p26). Maybe what is required, as he suggests, is a movement of what is considered to be 'the centre'. A more representative view of what constitutes health or the impact that disease can have on people's lives may change the way we measure health. The measures may then be capable of becoming genuinely universal, rather than the imposed universe of well educated, wealthy, white, middle class professionals from the US or other high- income countries (Guyatt, 1993; Fox-Rushby and Parker, 1995). Clearly this would require a radical change in the thinking and practice of research by those developing and using QALYs and DALYs.

6.2 Are DALYs better than QALYs as an outcome measure in economic evaluation?

In addressing this question, it is important to clarify whether values are likely to differ between the two approaches and to outline the mean-

ing of 'better', which draws on:

- types of applications (e.g. estimations from literature; application to evaluations of effectiveness);
- methodological issues (e.g. validity, reliability, sensitivity);
- inclusiveness (e.g. what is measured, whose values count).

The values given to disease states with DALYs and QALYs are likely to differ. Nord (personal communication) states that DALYs have a much stronger upper end compression (values closer to full health) than most versions of QALYs. It can be seen, for example, that the median values of the first four out of seven disability classes in Mark 2 DALYs (see Table 2 in Section 2.3.4) cover 30 percent of the scale, with the bottom three disability classes spread across the remaining 70 percent of the scale. Compare this with the EQ5D classification[102], where the first two out of three parts of the scale use 40 percent of the valuation space between 0 and 1 and a quick calculation shows Nord is correct. Of course, this does not show which approach might be 'right', just that they are likely to give different values. It is also the case that the QALY approach allows for the possibility that health states can be valued worse than dead whereas DALYs do not. Finally, as both the methods of valuation and the characteristics of the valuers differ between the DALY and QALY approaches, there is evidence to suggest that values are also likely to differ (Nord, 1992; Dolan et al., 1996).

I found only one example of a microeconomic evaluation using DALYs in a high-income country and this dealt with leishmaniasis amongst immigrant populations in the US (Muenning et al., 1999). Most evaluations using DALYs come from low-income countries and are for interventions to reduce infectious disease. The construction of the DALY is having an impact on the outcome measures used in economic evaluations from low- and middle-income countries. For example, in the period 1993-1996 only three academic papers reported cost per DALY averted for interventions to control communicable diseases (Walker and Fox-Rushby, 2000), whereas Appendix 5 shows

---

102 Using a time trade-off tariff derived from a random sample of 3,000 members of the UK population (Dolan et al., 1996).

that between 1997-2000 a further 11 papers using DALYs were pub-
lished and hence, where researchers once reported cost per death-
averted or life-year saved, there are now increasing reports of cost per
DALY averted (e.g. Marseille et al., 1998). It is also clear that the cal-
culations of DALYs are all estimated, at least in part, using tables from
the burden of disease series (for the disability weight attached to dif-
ferent sequelae, for the time spent in the sequelae states, or for the dis-
ability weight attached to treated and untreated disease), see for exam-
ple Goodman et al. (1999).

One of the most significant explanations for the current pattern of
use of DALYs and QALYs is that the DALY has not yet been opera-
tionalised for evaluating the effectiveness of specified interventions. It
has largely been a tool to estimate the potential impact of an unspeci-
fied intervention on a specific disease and the only way to access
DALY values is through the burden of disease group numbers.
However, as Fox-Rushby and Hanson (2001) highlight:

'none of the estimates of disability provided in the current
Global Burden of Disease series relates to any specific inter-
vention. This creates a problem for using DALYs in cost-effec-
tiveness analysis if researchers base their estimates on those pro-
vided in the Murray and Lopez (1996a-c)[103] books because
there is no way of distinguishing alternative interventions [for
the same disease] using the existing disability weights'.

A second reason for the pattern is because QALY measures are
rarely translated for use in low- or middle-income countries and so
there is little that researchers in these settings can draw on for use in
cost-effectiveness/utility analysis.

When considering the comparable evidence concerning the
validity, reliability and sensitivity of DALYs versus QALYs, it is clear
that there is almost no published evidence for DALYs but an array of
data for QALYs[104]. Indeed the lack of openness regarding the

---

103 Same references as in this monograph.
104 See for example, copies of all the annual EuroQol plenary group conference issues
and listing of published papers, available from Frank de Charro, Business Manager of
the EuroQol group http://www.eurqol.org

design[105] and measurement properties of DALYs is likely to be a significant problem for researchers choosing outcome measures for economic evaluations[106].

Given that significant criticisms of the DALY have focussed on the exclusive reliance on 'experts' to estimate changes in health state before and after treatment, it is interesting to compare this with the sources used to calculate QALYs. In 1992, a literature review by Gerard showed that 8 percent of papers used the researchers' own values for QALYs, 39 percent of studies used clinicians' views and 41 percent the views of the general community. She also showed that 25 percent of studies measured (rather than estimated) changes in health outcomes following an intervention (Gerard, 1992). More recently, it has been shown that the number of studies not reporting the source of values has fallen from 23 percent to 11 percent but that large proportions of values are still coming from health professionals (16 percent) and researchers (30 percent) (Gerard et al., 1999). Thus whilst QALY calculations are more frequently based on community preferences than DALYs, there are still many QALY values that are not. Gerard et al. (1999, p233) were highly critical of this and suggested that a study not reporting community preferences was 'probably not worth publishing'.

There are problems of how inclusive DALYs can be in their current state as an outcome indicator for economic evaluation. For example, they do not include side-effects, co-morbidity or ability to adapt to conditions before and after treatments. One of the questions on the EQ5D questionnaire for measuring health-related quality of life asks whether people have no problems, some problems or are unable to

---

105 There is a paucity of data and explanation. For example: what were the response rates by question in the valuation exercise; how did the disability weights change after each individual had deliberated on their value publicly within the focus group; what was the distribution of values; and how would results have changed if weights given had been divided into a six or eight disability classes instead of seven or if different size intervals had been used (see Table 2 in Section 2.3.4)?
106 The most comprehensive evidence concerning reliability and validity of DALYs' disease weights is presented by Stouthard et al. (1997), but this does not relate to the estimates of the global burden of disease exercise.

perform their usual activities. The impact of side-effects and co-morbidity are therefore subsumed within the answers. As people adapt to living with a disease they may adapt to their circumstance by, for example, changing their expectation of what is considered usual or what they feel to be important[107]. In either case this could lead a person stating that they have no problems in undertaking their usual activities. Because QALYs do allow for these issues, they are more inclusive than DALYs[108]. Theoretically, DALYs could include these if either the focus were changed from disease to interventions and the interventions were specified, or if a multi-dimensional measurement tool were operationalised and tested so that disease-specific health states tied to specific interventions could be measured or estimated[109], or if DALYs moved away from measuring disease to measuring health.

Section 4.2 listed a number of criticisms of DALYs concerning their bias against particular populations. Are QALYs considered to be any less biased than DALYs? Certainly there are several criticisms of the in-built ageism of QALYs. For example, Harris (1987) argued that QALYs discriminate against older people because of their shorter life expectancy. More recently, Johanesson and Johansson (1997) have suggested that as people value QALYs more if they gain them at younger ages rather than at older ages, QALYs in fact discriminate against the young by ignoring that factor. There is no reason to expect that the criticisms that using DALYs biases resources away from women, unborn babies and the poor[110] do not apply equally to QALYs if, for example, the same type of gender-biased data are used,

---

107 See for example Sprangers and Schwartz (1999), who outline a new theoretical approach to examining change in health-related quality of life scores over time. In particular, they argue that experience of illness may change an individual's internal standards or conceptualisation of quality of life.

108 This is only the case for a questionnaire-based measure such as the EQ5D. For a scenario-based approach to QALY valuation, it would depend on what was written in the scenario.

109 See Gerard et al. (1999) for an example of how women linked descriptions of breast screening to EQ5D classifications and hence adopted the values provided by the EuroQol group. Kirsch and McGuire (2000) also elicit health state valuations from the general population for the New York Heart Association classification.

110 See Sections 4.2.4 to 4.2.6 above.

114   or if the beginning of life is thought of in the same way, or if poor people are less able or likely to return to full health.

Many of the criticisms of DALYs that are made by the disability rights movement are also levelled at QALYs. For example, Koch (2000) equates DALYs and QALYs in their broad aims and criticises both for disregarding the sanctity of human life. Both, he argues, would allocate funds away from people like Stephen Hawking and Christopher Reeves, who face challenges of severe physical limitations, if the only alternative was giving the same funds to perfectly healthy people, because 'prospective, health-related quality of life constructs – and the culture that supports them – assume a restricted life is necessarily less full than one lived in a state of physical normalcy irrespective of any other context or condition' (Koch, 2000, p425). Indeed the core criticism, and centre of the legal case, that prevented the introduction of the state of Oregon's proposal for health care rationing based on cost per QALY was that it discriminated against those who were disabled and could not be returned to full health after life-saving treatments[111] (Ubel et al., 1999, p739).

The underlying problem that these criticisms highlight is that the focus of decision-makers is drawn to maximising 'health', rather than a more general definition of utility, and acceptance of the distribution of resources across people. This implies willingness to value 'health' above all other outcomes. If the desired objective function is more than just 'health', however, it implies that the best strategies for improving utility and efficiency may be forfeited. Moving to a broader measurement of well-being and, as Broom (2001) suggests, distinguishing measured change from the value of that change would allow clarification of the impact of alternative value bases and hence alternative distributions of well-being to be included. As objective functions are likely to differ from country to country, it suggests we should

---

111 See also the open letter on the internet from the Consortium for Citizens with Disabilities (17th July 2000, p4) to DeParle of the US Health Care Financing Administration, that stated 'both the QALY and DALY have serious flaws when .. (they are)... premised on the antiquated notion that once disabled, quality of life is poor, when in fact, with adequate health services, assistive technologies, and other supports, the quality of life of a person with even a significant disability can be quite high'.

be particularly cautious about transferring such measurement tech-
nologies under the auspices of improving efficiency.

6.3 DALYs or QALYs in decision-making?

Decision-makers and their needs are not homogenous. Therefore I con-
sider a variety of settings: at international, national and sub-national lev-
els; and in high-, middle- and low-income countries. Specifically, I con-
sider how QALYs and DALYs may feed into decisions about where to
target expenditure in the health sector. Decision-makers' preferences for
QALYs or DALYs will depend on the way in which decisions are cur-
rently made in any setting, the availability of data, and views on the rel-
evance of burden of disease and cost-effectiveness in priority setting.

There are a number of actors at the international level interested in
using QALYs and DALYs for decision-making, including:

- international agencies, such as the World Bank and World Health
  Organisation;
- multi-national multi-agency initiatives like the Global Alliance for
  Vaccines Initiative (GAVI);
- non-governmental organisations, like Save the Children and
  Oxfam;
- bilateral aid agencies, such as the UK Department for
  International Development;
- multi-national pharmaceutical companies;
- international research organisations.

The common link is that each organisation is interested in fund-
ing, providing or influencing health care in several countries.
Comparison of the outcomes of intervening in disease areas can be
part of their decision-making and DALYs offer an easy route to this.

A vast amount of data on the burden of disease has been compiled
by the global burden of disease group, along with estimates of the
impact of treatment on DALYs. A broad brush approach can quickly
assume what intervention might be used and estimates of effectiveness
made. Indeed a cursory glance through the pages of Jamison et al.
(1993) will provide global estimates of the cost per DALY averted for
a range of named interventions.

116

Compare this with data on cost per QALY gained and the difference is stark. QALYs were developed to measure and value change from interventions and not to assess burden of disease, and therefore they have not been used widely in burden of disease studies. Secondly they have been developed and used predominantly in high-income countries. To date the EQ5D survey instrument has been translated for use in 19 countries[112], with the section on self-assessed health translated most frequently for use in international clinical trials[113]. However, even the EuroQol group is cautious about how comparable results from new translations are across countries. Therefore, for international agencies considering allocation of expenditure across low- and middle-income countries, QALYs currently offer little help.

It is the case that, even when information is not available, decisions still have to be made. This is a key reason for the push to developing, and continuing the use of, the DALY for decision-making by international organisations. It is also important to recognise that international organisations have less incentive to respond to the needs of local populations because not only are they interested in comparisons across diverse populations but also because they are not held accountable to local populations. The role of the WHO and of the World Bank in its lending for health sector development and sector-wide programme planning is also key because the focus remains on 'health', rather than a broader notion of well-being.

There is a disturbing focus on disease burden by international agencies. Burden of disease studies are really only helpful in knowing the size of the problem. They do not suggest what might be done or

112 Languages into which the EQ5D instrument have been translated include: Afrikaans, Catalan, Croatian, Czech, Danish, Dutch, Finnish, French, German, Hungarian, Italian, Japanese, Norwegian, Polish, Spanish, Swedish, Turkish, UK English. Cultural adaptations have additionally been prepared for Argentina, Australia, Austria, Canada, Mexico, Peru, South Africa, US and Venezuela. Seventeen further translations and adaptations are awaiting approval by the EuroQol group for use in other countries. (Source: EuroQol website, http://www.euroqol.org reported on 20th July 2000).

113 In order to get EQ5D scores and the facility to calculate QALYs from such questionnaires it is necessary to adopt one of the valuation scoring systems produced from existing valuation surveys, which are available only from a limited set of countries.

where to start, or anything about the decision to provide care for one disease in terms of what is lost in not providing care for another disease. Knowing the size of the disease burden is helpful if a disease is to be eradicated (e.g. polio) or to ascertain the size of current and future resource use for palliative care given current practice. However, a choice to eradicate polio or continue palliative care also involves an implicit choice not to put the resources somewhere else, for example into reducing measles or promoting healthy life styles. Should the WHO's Macroeconomic Commission on Health succeed in raising sufficient global funding for health care in developing countries to trillions rather than the millions of US$ currently allocated (Sachs, 2001), it may also help to know the size of the burden of disease (to gain a sense of the possible orders of magnitude for change). However, much of this data is now available and yet precious research and planning resources continue to be channelled into burden of disease calculations rather than learning about the costs and effects of health interventions. Does this matter?

There is astonishingly little data available on the efficiency of health interventions. Walker and Fox-Rushby (2000) showed that only 107 papers were published on communicable disease interventions in developing countries in academic journals between 1984-1997, which is an average of one per low- or middle-income country! In addition, fewer than 35 percent of these used an outcome indicator suitable for comparison across diseases (including a mix of death averted, DALYs or a monetary valuation) and only 7 percent of papers considered the generalisability of their findings to other contexts.

This rather frustrating situation for decision-makers clearly needs to change. It is important to direct research to decisional information about which interventions are most efficient to provide, where and when. As all DALYs are calculated on the basis of estimation (rather than direct measurement) it does suggest that DALYs might provide a useful route to increasing the utility of existing economic evaluations for international agencies – as long as these are tied to specific interventions and not just burden of disease. Alternatively, researchers might consider making similar kinds of estimations using an EQ5D QALY based exercise but it is likely to take longer. Those agencies dis-

agreeing with the equity implications, the narrow disease focus or the lack of consideration of context in calculating DALYs urgently need to consider how to invest in changing this situation. It also suggests that evaluators need to develop techniques for assessing their findings in the light of their spatial and temporal context, and to aim to find out how and why costs and effects vary, to improve the validity of transferring or adapting results and models across countries.

Considerations for using DALYs and QALYs at the national level need to take account of current approaches to decision-making in the health sector, as there may or may not be a demand for outcome measures combining quantity and quality of life. Where there is a demand, its nature with respect to these specific tools is interesting to contemplate, as like much of health care, the suppliers and demanders are not always distinct. For example, academic advisors recommend use of their own instruments and may also be employed by government ministries of health to conduct or advise on the evaluation of medical technologies. Once these ministries or other organisations have invested in the research and development or use of an instrument, there is likely to be some unwillingness to change to using another type of outcome measure to aid policy-making.

For decision-making concerning resource allocation, DALYs can be considered a substitute for QALYs[114]. The degree of substitutability is likely to be related to the number of substitutes available (i.e. the number of valid, reliable and culturally appropriate methods for combining quantity and quality of life, in addition to the EQ5D) and the speed with which one can be substituted for another. Finally, the tastes and preferences of decision-makers are also key.

Table 9 delineates potential scenarios of the demand for DALYs in relation to QALYs. It considers differences between high-, middle- and low-income countries and why the demand for DALYs in deci-

114 I distinguish this from the demand for DALYs as a research tool where, at least in the short term, I think that DALYs and QALYs are more likely to be complements than substitutes. This is because testing the reliability and validity of new instruments is a comparative process where the relative performance of new and old measures is compared. It will be particularly relevant should a questionnaire version for measuring and valuing the 'D' part of DALYs become available.

sion-making may be higher or lower relative to QALYs. At present, it would appear there is likely to be less demand for DALYs than for QALYs by decision-makers in high-income countries. This is because there are a number of good quality and well-researched substitutes for DALYs as a result of a substantial investment. This considerable knowledge will take a significant time to replicate in a new instrument, both in terms of the measure itself and as a measure of the effectiveness of various interventions.

The situation is very different in low-income countries, as they are required to calculate DALYs averted as part of the process of applying for health sector loans from the World Bank. Should there be insufficient local expertise to carry out the calculations, external consultants can be hired to help complete the procedure (this may even be part of the loan). Low-income countries have little research money and may also represent poor potential markets for pharmaceutical companies. Therefore it would seem unlikely that QALY based measures will receive priority for development there. Clearly if negotiating positions between the World Bank and individual countries change, or the World Bank chooses to develop other approaches, or if other organisations sponsor the development of culturally relevant national tools, then this situation might change.

Middle-income countries such as Mexico have the potential to develop either or both of QALYs and DALYs. Links with World Bank funding have facilitated the calculation of DALYs. Indeed, Mexico has undertaken one of the most extensive analyses of burden of disease and cost-effectiveness to date[115]. However, it also has a number of research projects developing QALYs and other quality of life outcome measures, which may have the potential to develop alternative approaches for resource allocation (see for example: de Icaza et al., 1997; Caso- Marasco et al., 1996; Lara-Munoz et al., 1995). Research

---

115 It will be fascinating the see the impact of DALYs on decision-making in Mexico as, at the time of writing (May 2002) the Minister of Health was Dr Julio Frenk, who was the director of the Evidence and Information for Policy cluster at WHO until May 2001. He was therefore responsible, with Dr Christopher Murray, for leading the development of tools to assess the performance of health systems.

Table 9 **Potential demand for DALYs relative to QALYs by decision-makers in different countries**

Reasons for:	High-income country e.g. UK	Middle-income country e.g. Mexico	Low-income country e.g. Tanzania
High demand for DALYs relative to QALYs	– If DALYs[116] become viewed as an indicator of health system performance – If decision-makers prefer disease- rather than intervention-based decision-making – Government becomes concerned about using DALYs for world ranking of health system	– Limited development effort invested in QALYs – Cultural relevance of existing QALY questionnaire instruments questioned – World Bank ties loans to investment plans requiring calculation of DALYs – Developers of DALYs employed in Ministry of Health or as advisors – DALYs wanted as an indicator of health system performance	– World Bank ties loans to investment plans requiring calculation of DALYs – Multi-national organisations sponsor calculation of DALYs but not QALYs – Measurement using QALYs too expensive
Low demand for DALYs relative to QALYs	– Many substitutes for DALYs available – Database on cost per QALY exists already for many interventions – Instruments used to calculate QALYs were developed and tested with UK populations – Developers of QALYs employed by or advise Ministry of Health – Ministry of Health has already invested in research developing QALYs – DALYs are not usable to measure effectiveness – Performance of the currently developing DALY is not substantially better than existing tools	– Research developing or adapting QALYs shows sign of promise and research funding is increased – Results of health sector prioritisation exercises using DALYs are rejected – Promising shared market (with high-income countries) for pharmaceuticals – Cultural relevance of DALYs questioned	– Results of health sector prioritisation exercises using DALYs are rejected – Negotiating position with World Bank improves – Strong political movement for culturally relevant measures to be developed, backed up with research funds – Increase in number of pharmaceutical sponsored trials of cost per QALY

116 Or disability adjusted life expectancy (DALE).

income and potential private markets may have a significant impact on research in such middle-income countries, in particular leading to investigation of culturally appropriate national instruments.

DALYs have been used at the sub-national level to set priorities. The only published examples I could find were from the UK (Bowie, 1998), Australia (Vos and Begg, 1999) and Tanzania (Finlay et al., 1995; The Economist, 2002). It is likely that the issues identified at the national level, exemplified in Table 9, are similar and possibly accentuated at this level. For example, given that per capita spending on health in Tanzania ranges between US$4-7 (Finlay, 1996, p7), it is not surprising that the Tanzanian project was funded from money external to the country (although, it is not clear how much of this might be in the form of a repayable loan).

7 CONCLUSIONS

In drawing conclusions about what DALYs have achieved and what might have been achieved without them, I draw on their four principal uses as:

- a measure of the burden of disease;
- a director of research and development;
- a measure of effectiveness[117]; and
- a unit of benefit in cost-effectiveness analysis for micro and sectoral evaluations.

Finally I outline key research questions for the future.

The construction of DALYs has had a major impact on thinking about the burden of disease across the globe. It provided the first estimates of the impact of disease on death and disability, using one measure to combine both effects. This had never before been done across so many diseases or so many regions of the world. The impact was to highlight the burden of some diseases in a new way. For example the DALY estimates stressed the size of the burden of mental health. Results from the global burden of disease exercise have been provided in an eminently usable way and the results have been cited by a myriad of agencies, often calling for new resources to prevent, treat or research a specific disease area. However, the major problems with the use of the DALY as a measure of the burden of disease are:

- the quality of the evidence;
- its inability to account for co-morbidity other than by adding up the impact of each non-fatal disease in an individual;
- the implied equity weighting introduced by the 'aspirational' life expectancy.

The first two concern the reliability of the data. The first can be dealt with to some extent by understanding the impact of data uncertainty through sensitivity analysis. The second is much harder to account for within the current DALY framework because the focus is on diseases rather than individuals or communities. If the focus were to change to individuals then, in theory, it would be possible to consider the marginal effect of each disease, although in practice this would be a difficult and costly exercise.

---

117 DALYs or disability adjusted life expectancy (DALE).

The third point, concerning equity weights is tricky, but relevant for international comparisons. Murray's arguments concerning the implication of not capturing an ideal length of life (i.e. undervaluing lives in poor countries relative to richer countries because of differential life expectancies) are understandable, but Williams shows clearly the impact of the implied equity weights and that they do not match the stated desires (Williams, 1999). It would therefore be better to clarify the impact of the value system imposed by separating it out from the estimation of quantity of life years and disability. This would also allow others the option to adopt different valuation approaches.

The value of burden of disease studies to decision-makers (and ultimately to citizens) must be questioned. They have been shown to be costly endeavours that tie up scarce human resources. Results of DALY burden exercises tell us only about the size of the problem, and have nothing to do with the solutions. Indeed, to find solutions (deciding which interventions should be funded) requires a complete re-estimation of life expectancy with and without interventions, specific to the location of the intervention (and therefore without the aspirational life tables used in DALYs), as well as knowledge of the setting-specific costs. The burden of disease data tell us nothing about this and the business of ascertaining the useful information needs to begin.

The startling paucity of data on cost-effectiveness of health interventions in developing countries, and the concerns voiced over transferring cost-effectiveness results across settings, have been used to explain why burden of disease studies are used. But why is the money not being transferred to providing useful decisional information? Why are some countries being required to undertake more burden of disease studies prior to undertaking any analysis of the variations in cost-effectiveness of alternative interventions? There is an urgent need for useful data on the costs and cost-effectiveness of alternative health interventions.

There has been both less criticism of, and less work directed at, the use of DALYs to help decide research agendas. The calculation of DALYs has helped change or cement decisions to invest in particular areas, for example recommending moving some funding of research

from HIV to TB. However, the main problem with using DALYs in this context is the focus on disease and not the treated person who may experience side-effects (which could be worse than the disease itself). The Ad Hoc Committee on Health Research Relating to Future Intervention Options (1996, p85) highlighted the need for more information about the effectiveness of single, and packages of, health interventions. However, it only ranked the value of such information (along with cost) as fifth of five priorities for future research[118].

As a measure of effectiveness of health interventions, the development of DALYs has encouraged a wide variety of governments across the world to think about the impact of interventions on morbidity and mortality, although in practice many low-income countries still only use mortality and therefore only the YLL part of DALYs. However, the availability of disability weights does allow an easy route to estimating YLDs. All that is needed are estimates of prevalence, life tables, the size of the population treated and consideration of the average life history of a patient with and without treatment, and DALYs averted from an intervention can then be estimated. The main problem with using this as decisional information is that none of the disability weights that have been created are related to any particular intervention. It therefore means that unless evaluators use different DALY weights for different interventions, the only way of differentiating the impact of alternative interventions for the same disease is through the traditional epidemiological parameters (mortality and morbidity) and the proportion of a population treated.

Murray argues that patient reported data are biased for a range of reasons, e.g. adaptation to disease, and that patients are unable to report the 'truth'. This is reflected both in the original development of DALYs as a tool for estimation by health professionals rather than for measurement in the population and the more recent direction of

---

118 The four higher research priorities were: development of simple, affordable methods of monitoring the status and trends of mortality and morbidity; development of indicators of population health status that take into account morbidity, mortality and disability; measurement of the degree to which research findings are implemented, why and how to improve it; and development of cheap and reliable methods of monitoring the patterns of use of health care services.

'adjusting' self-reports by 'objective' biological measures. This is a major departure from the alternative approaches that measure health-related quality of life and QALYs. It also serves to highlight some of the values held within the estimation of disability encompassed in DALYs. The views of patients are excluded from the calculation of the burden of disease and the impacts of interventions, as both are considered the preserve of the 'expert' professional. This is a key criticism upheld by a range of disability rights groups as well as by some of the 'experts' who were involved in the DALY valuation exercises. The main problems are that DALYs are too narrowly focussed, they do not account for the context of illness and there are concerns about the quality of data. For this to change, DALYs have to move away from a biomedical model of health, to one that is centred on people and not only disease.

The ability to move away from a bio-medical model to a more holistic, socialised and context oriented conception of health or well-being that can be used to measure change from health interventions across the world is, however, not easy. There is no universal agreement on what constitutes an adequate measure of health. Given the process of its development and the conception of health adopted, the WHO-QOL instrument probably comes closest to this at present (WHO-QOL, 1998a,b; Bowden and Fox-Rushby, in press) but unfortunately it does not provide information that allows life years to be adjusted for quality. Whilst the EQ5D does allow this, there is concern about how appropriate that instrument would be for populations in low- and middle-income countries. There are also increasing questions of how culturally fair the assessment processes are. The desire for international comparisons could lead to one of several developments including:

- translation and adaptation of instruments like the EQ5D, with particular attention paid to evaluation of different types of equivalence;
- development of the WHOQOL into an index suitable for quality adjusting survival data (plus further translation and adaptation);
- development of different national questionnaires designed to tap universally agreed dimensions of health (these may include very

different questions across countries about pain, but presupposes research on the universality of dimensions of 'health') which are then valued on a utility scale;

- developing broader measures of well-being rather than just one with a limited focus on health.

This range of suggestions indicates that the value of an operationalised DALY measurement tool is likely to be much less in those countries that have already invested in QALY-type measures, e.g. the UK, but also that much methodological development of an operationalised DALY would be needed in a very wide range of countries before accepting results as valid, reliable or equivalent across countries. The potential cost of developing such measures also explains, in addition to requirements made by the World Bank, why low-income countries are more likely to use DALYs.

The main value of DALYs is to organisations like the World Bank and World Health Organisation that have to make international comparisons and decide which types of health services to fund or recommend. However, the calculation of DALYs exacts a toll. Most prioritisation exercises in lower income countries, whether based on burden of disease or cost-effectiveness, require foreign assistance in terms of personnel and funding, and most take on the assumptions embodied within the construction of DALYs. At the national level the value for decision-making of using DALYs relative to QALYs will depend on:

- the development of views about what constitutes an appropriate approach to prioritising health care resources;
- a country's need for funding by the World Bank; and
- the extent of current knowledge about QALYs and DALYs.

The use of DALYs as an outcome measure in cost-effectiveness analysis raises further issues because of the implications for resource allocation. One of the main concerns that has arisen is the extent to which specific assumptions are embroiled inextricably within the construction of DALYs. For example, the current methods and interpretations lead valuers and decision-makers into discrimination against those with disability when considering the value of life-enhancing interventions, and therefore may increase inequities in health and access to health care.

Separating out the measurement and valuation aspects of DALYs would provide two benefits: a measurement tool that could be developed and tested more easily; and an aid to assessment of the impact of alternative approaches to valuation. It would mean, for example, using base case DALYs without weights for age or time preference, so that different value systems that weight the distribution of potential health gain and/or initial levels of health may be considered. This would also allow evidence to be built up and then added, for example on time preference rates. It might also encourage evaluators to examine the sensitivity of their results to uncertainties in the data. Finally, when the objective function of health spending differs between countries, this approach could also be used to compare the effect of different national and international priorities.

The research agenda is large. I would suggest that there is a pressing need to move away from funding more of the same burden of disease studies to instead funding cost-effectiveness analyses and for these analyses to focus specifically on understanding the variation in cost-effectiveness of health interventions across settings. Clearly for this to have an impact supposes that cost-effectiveness analyses have a value through impact on decision-making. Whether they do have such an impact is an important question itself. In the short term, in low- and middle-income countries, the DALY could help as an outcome measure for this exercise but this must be accompanied by a focus on interventions rather than disease, and an operationalisation of methods to allow for alternative valuations of the distribution of health gain across populations.

In the medium term, better information is needed on the generalisability of existing methods for measuring and valuing gains in health and well-being from interventions across settings, rather than relying on estimation by health professionals sitting in offices in Geneva or elsewhere.

In the longer term (beginning now), we need to see a radical change in the way that measures of health (or well-being) that claim to be universal, or are used globally, are developed. Starting with people rather than disease is key. This does not exclude disease but will add to it the context in which diseases thrive and people live, and so is

128   more likely to provide usable interventions that people believe improve their well-being.

REFERENCES

Abbasi K (1999a) The World Bank and world health: interview with Richard Feacham *Br Med J* 318, 1206-1208

Abbasi K (1999b) The World Bank and world health: focus on South Asia – II: India and Pakistan *Br Med J* 318, 1132-1135

AbouZahr C (1999) Disability adjusted life years (DALYs) and reproductive health: a critical analysis *Rep Hlth Matters* 7, 14, 118-129

AbouZahr C, Vaughan P (2000) Assessing the burden of sexual and reproductive ill-health: questions regarding the use of disability-adjusted life years *Bull WHO*, 78, 5, 655-666

Ad Hoc Committee on Health Research Relating to Future Intervention Options (1996) *Investing in Health Research and Development*, World Health Organisation, Geneva

Aiyer S, Jamison DT, Londono J-L (1995) Health policy in Latin America: progress, problems and policy options *Cuad de Econ* 32, 95, 11-28

Akhavan D, Musgrove P, Abrantes A, Gusmao Rd'A (1999) Cost-effective malaria control in Brazil: cost-effectiveness of a malaria control programme in the Amazon Basin, Brazil, 1988-1996 *Soc Sci Med* 49, 10, 1385-99

Anand S, Hanson K (1997) Disability-adjusted life years: a critical review *J Hlth Econ* 16, 6, 685-702

Anand S, Hanson K (1998) DALYs: efficiency versus equity *Wrld Dev* 26, 2, 307-310

Andrews G (2000) *Managing Scarcity: a Worked Example Using Burden and Efficacy* presentation for the national workshop on co-morbidity, Canberra ACT http://www.unsw.edu.au/clients/crufad/burden/burden.htm

Andrews G, Mathers C, Sanderson K (1998a) The burden of disease *Med J Aust* 169, 3, 156-158

Andrews G, Sanderson K, Beard J (1998b) Burden of disease: methods for calculating the disability from mental disorder *Br J Psychiatry* 173, 123-131

Appelby L, Shaw J, Amos T, Dennehy J (1997) Global burden of disease *Lancet*, letter 350(9071) (141-145)

Arnesen T, Nord E (1999a) The value of DALY life: problems with ethics and validity of disability adjusted life years *Br Med J* 319(7222), 1423-5

Arnesen T, Nord E (1999b) The problems with disability weights *Br Med J*, email letter 3rd December

Ashby J, O'Hanlon M, Buxton M (1994) The time trade-off technique: how do the valuations of breast cancer patients compare to those of other groups? *Qual Life Res* 3, 4, 257-265

Barandregt JJ, Bonneux L, van der Maas PJ (1996) DALYs: the age weights on balance *Bull WHO* 74, 4, 439-43

Barker C, Green A (1996) Opening the debate on DALYs *Hlth Pol & Plan* 11, 2, 179-183

Berman S (1995) Otitis media in developing countries *Pediatrics* 96, 126-131

Bevan G (1998) Allocating the WHO's resources rationally *Br Med J* 316(7142), 1403-1404

Bevan G, Hollinghurst S, Bowie C (1999) Disability-adjusted life years: an introduction to their objectives, methods and potential *Eurohealth* 5, 2, 28-30

Block MAG, Sandiford P, Ruis JA, Rovira J (2001) Beyond health gain: the range of health system benefits expressed by social groups in Mexico and Central America *Soc Sci Med* 52, 1537-1550

Blue I, Harpham T (1994) The World Bank World Development Report 1993: Investing in Health: Reveals the burden of common mental disorders, but ignores its implications *Br J Psych* 165, 1, 9-12

Bobadilla J-L (1996) Searching for essential health services in low- and middle- income countries *Draft Report for Human Development Department*, World Bank

Bobadilla J-L, Cowley P (1995) Designing and implementing packages of essential health services *J. Int Dev*, 7, 3, 543-554

Bobadilla J-L, Cowley P, Musgrove P, Saxenian H (1994) Design, content and financing of an essential national package of health services *Bull WHO* 72, 4, 653-662

REFERENCES

Bowden A, Fox-Rushby JA (in press) A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East and South America *Soc Sci Med*

Bowie C (ed.) (1998) *The South and West DALYs Project: Developing a Powerful Tool for Health Care Planning* Somerset Health Authority

Bowling A (1997) *Research Methods in Health: Investigating Health and Health Services* Open University Press

Bremberg S (1998) Health promotion in school age children *Scand J Soc Med* 26, 2, 81-4

Brock DW (1998) Ethical issues in the development of summary measures of population health status. In Field MJ, Gold MR (eds.) (1998) *Summarizing Population Health: Directions for the Development of Population Metrics,* Division of Health Care Services, Institute of Medicine, National Academy Press, Washington

Brooks R (1996) EuroQol: the current state of play *Hlth Pol* 37, 1, 53-72

Broome J (2001) *Health Measurement and Decision*, presentation at Health Economists Study Group, University of Oxford, January 2001

Buse KD (1993) *The World Bank and International Health Policy: Genesis, Evolution and Implications,* Master's Thesis, London School of Economics

Bussbach JJV, Jessing DJ, de Charro FT (1993) The utility of health at different stages of life: a quantitative approach *Soc Sci Med* 37, 2, 153-158

Cairns JA, van der Pol MM (1997) Saving future lives: a comparison of three discounting models *Hlth Econ* 6, 4, 341-350

Carter R, Vos T, Stone C, Hocking J, Mihalopoulos C, Peacock S (2000) *Priority Setting in Cancer Control: Using DALYs in an Evidence-Based Approach to Program Budgeting and Marginal Analysis*, presentation at the International Conference on Priorities in Health Care, Amsterdam, November 2000

Caso-Marasco A, Vargas-Ruiz R, Salas-Villagomez A, Begona-Infante C (1996) Double-blind study of a multi-vitamin complex supplemented with ginseng extract *Drugs Exp Clin Res* 22, 6, 323-9

Cassels A, Janovsky K (1998) Better health in developing countries: are sector-wide approaches the way of the future? *Lancet* 352, 1777-1779

Chan MS (1997) The global burden of intestinal nematode infection – fifty years on *Parasitol Today* 13, 11, 438-443

Coleman PG, Goodman C, Mills A (1999) Rebound mortality and the cost-effectiveness of malaria control: potential impact of increased mortality in late childhood following the introduction of insecticide treated nets *Trop Med Int Health* 4, 3, 175-86

Consortium for Citizens with Disabilities (2000) Open letter to Honorable Nancy-Ann DeParle, Health Care Financing Administration, Washington DC www.resna.org/ata/ccdcomments.htm

Coons SJ, Rao Sumati, Keininger DL, Hays RD (2000) A comparative review of generic quality of life instruments *Pharmacoecon* 17, 1, 13-35

Curlin P, Tinker A (1995) Women's health, *Inf Dis Clin Nrth Am* 9, 335-51

Daniels N (1998) Distributive justice and the use of summary measures of population health status. In: Field MJ, Gold MR (eds.) (1998) *Summarizing Population Health: Directions for the Development of Population Metrics* Division of Health Care Services, Institute of Medicine, National Academy Press, Washington

De Icaza E, Arredondo A, Calderon C, Hernandez G (1997) Changes in the quality of life of patients with end-stage renal disease treated with high-efficiency bicarbonate hemodialysis in Mexico *Renal Failure* 19, 1, 99-110

Deen JL, Vos T, Huttly SR, Tulloch J (1999) Injuries and non-communicable diseases: emerging health problems of children in developing countries *Bull WHO* 77, 6, 518-524

Department for International Development (DfID) (1999) *Target Strategy Paper: Better Health for Poor People,* UK

Department for International Development (DfID) (undated) *What Roles Should the DALY Play? – Its Use and Potential Abuse?* A briefing note for health professionals, UK

Dolan P, Gudex C (1995) Time preference, duration and health state valuation *Hlth Econ* 4, 4, 289-300

REFERENCES

Dolan P, Gudex C, Kind P, Williams A (1995) *A Social Tariff for EuroQol: Results from a UK General Population Survey* Discussion paper no 138, Centre for Health Economics, University of York, York

Dolan P, Gudex C, Kind P, Williams A. (1996) The time trade-off method: results from a general population study *Hlth Econ*, 5, 141-154

Donaldson C, Birch S, Gafni A (2002) The distribution problem in economic evaluation: income and the valuation of costs and consequences of health care programmes *Hlth Econ* 11, 1, 55-70

Eisenberg L (1997a) Psychiatry and health in low income populations *Compr Psychiatry* 38, 2, 69-73

Eisenberg L (1997b) Global burden of disease – letter *Lancet* 350, 9071, 141-145

Elbasha EH (2000) Discrete time representation of the formula for calculating DALYs *Hlth Econ* 9, 4, 353-366

Essink-Bott M-L, Krabbe PFM, Bonsel GJ, Aaronson NK (1997) An empirical comparison of four generic health status measures *Med Care* 35, 5, 522-537

Essink-Bott M-L, Stouthard M, Bonsel G, Gunning-Schepers L, van der Maas P (1999) The problems with disability weights *Br Med J* 319, e-letter, 2nd December

Fèvre EM, Barnish G (1999) Malaria treatment policies: when and how should they be changed? *Ann Trop Med Parasitol* 93, 6, 549-560

Field MJ, Gold MR (eds.) (1998) *Summarizing Population Health: Directions for the Development of Population Metrics* Division of Health Care Services, Institute of Medicine, National Academy Press, Washington

Finlay JF (1996) The World Development Report 1993: Investing in Health. Is it practical? *BRIDGE,* Spring Edition, 6-7

Finlay JF, Law MM, Gelmon LJ, de Savigny D (1995) A new Canadian health care initiative in Tanzania *CMAJ* 153, 1081-5

Foege W (1994) Preventive medicine and public health *JAMA* 271(121), 1704-5

134

Foster S, Phillips M (1998) Economics and its contribution to the fight against malaria *Ann Trop Med Parasitol* 92, 4, 391-8

Fox-Rushby J, Hanson K (2001) Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis *Hlth Pol & Plan* 16, 3, 326-331

Fox-Rushby J, Parker M (1995) Culture and the measurement of health-related quality of life *Eur Rev App Psy,* 45, 4, 257-263

Friedman C (2000) *Cars, Strawberries and DALYs: Assessing Burden of Disease Among Working Americans,* presentation to National Committee on Vital and Health Statistics, Washington Feb 24 http://ncvhs.hhs.gov/000224tr.htm

Gafni A (1995) Time in health: can we measure individuals' 'pure time preferences'? *Med Dec Making* 15, 31-37

Gerard K (1992) Cost-utility in practice: a policy maker's guide to the state of the art *Hlth Pol* 21, 249-279

Gerard K, Smoker I, Seymour J (1999) Raising the quality of cost-utility analyses: lessons learnt and still to learn *Hlth Pol* 46, 217-238

Gilson L (1998) In defence and pursuit of equity *Soc Sci Med* 47, 12, 1891-1896

Gilson L, Mkanje R, Grosskurth H, Mosha F, Picard J, Gavyole A, Todd J, Payaud P, Swai R, Fransen L, Mabey D, Mills A, Hayes R (1997) Cost-effectiveness of improved treatment services for sexually transmitted diseases in preventing HIV-1 infection in Mwanza Region, Tanzania *Lancet* 350(9094), 1805-9

Glick S (1994) Disability-adjusted life years – ethical implications. In Bankowski Z and Bryant JH (eds.) *Poverty, Vulnerability, the Value of Human Life, and the Emergence of Bioethics: a Global Agenda for Bioethics* highlights and papers of the XXVIIIth CIOMS conference, Ixtapa, Guerrero State, Mexico, 17-22 April, 1994

Global Forum for Health Research (2000) *The 10/90 Report on Health Research 2000,* World Health Organisation, Geneva

Goodman C, Coleman P, Mills P (1999) Cost-effectiveness of malaria control in sub-Saharan Africa *Lancet* 354(9176) 378-85

REFERENCES

Greikspoor A, Sondorp E, Vos T (1998) Cost-effectiveness analysis of humanitarian relief interventions: visceral leishmaniasis treatment in the Sudan *Hlth Pol & Plan* 14, 1, 70-76

Groce NE (1999) Framing disability issues in local concepts and beliefs *Asia Pac Dis Rehab Jl* 10, 1, 1-9

Groce NE, Chamie M, Me A (2000) Measuring the quality of life: rethinking the World Bank's disability adjusted life years *Disab World* 3, June/July edition p1-8 www.disabilityworld.org/June-July2000/International/DALY.html

Gross CP, Anderson GF, Powe NR (1999) The relation between funding by the National Institutes of Health and the burden of disease *New Eng J Med* 340, 24, 1881-7

Guerrant DI, Moore SR, Lima AA, Patrick PD, Schorling JB, Guerrant RL (1999) Association of early childhood diarrhoea and cryptosporidiosis with impaired physical fitness and cognitive function four-seven years later in a poor urban community in northeast Brazil *Am J Trop Med Hyg* 61, 5, 707-13

Gupta PC, Sankaranarayannan R, Ferlay J (1994) Cancer deaths in India: is the model-based approach valid? *Bull WHO* 72, 943-944

Guyatt G (1993) The philosophy of health-related quality of life translation *Qual Life Res*, 2, 461-466

Gwatkin DR, Guillot M, Heuveline P (1999) The burden of disease among the global poor *Lancet* 354(9178), 586

Hanson K (1999) *Measuring Up: Gender, Burden and Priority Setting Techniques in the Health Sector,* Discussion paper, Harvard University

Harris J (1987) Qualifying the value of life *J Med Eth* 13, 117-123

Herdman M, Fox-Rushby J, Badia X (1997) Equivalence and the translation and adaptation of health-related quality of life measures *Qual Life Res* 6, 237-247

Herdman M, Fox-Rushby J, Badia X (1998) A model of equivalence in the cultural adaptation of HRQL instruments: the universalist approach *Qual Life Res*, 7, 4, 323-335

136   Herdman M, Fox-Rushby J, Rabin R, Badia X, Selai C (forthcoming) Developing and translating different language versions of the EQ-5DBrooks R, Rabin R, de Charro F (eds.) *The Measurement and Valuation of Health Status Using the EQ-5D: European Perspective* (Evidence from the EuroQol BIOMED research program), Kluwer Academic Publishers, Dordrecht, The Netherlands

Hinman AR (1997) Quantitative policy analysis and public health policy: a macro and micro view *Am J Prev Med*, 13, 1, 6-11

Hollinghurst S, Bevan G, Bowie C (2000) Disease by disability adjusted life years *Hlth Care Man Sci*, 3, 9-21

Homedes N (1995) *The disability-adjusted life year (DALY) definition, measurement and potential use* Human Capital Development and Operations Policy Working Paper, World Bank

Horton S, Sanghvi T, Phillips M, Fielder J, Perez-Escamillia R, Lutter C, Rivera A, Segall-Correa A-M (1996) Breastfeeding promotion and priority setting in health *Hlth Pol & Plan* 11, 2, 156-168

Hyder AA, Morrow RH (1999) Steady state assumptions in DALYs: effect on estimates of HIV impact *J Epid Com Hlth* 53, 1, 43-5

Jamison DT (1993) Disease control priorities in developing countries: an overview. In Jamison DT, Mosley WH, Measham AR, Bobadilla JL (eds.) (1993) *Disease Control Priorities in Developing Countries* The World Bank, Oxford University Press

Jamison DT (1997) Investing in health *Fin & Dev* 30, 3, 2-5

Jamison DT, Frenk J, Knaul F (1998) International collective action in health: objectives, functions and rationale *Lancet* 351(9101), 514-517

Jamison DT, Mosley WH, Measham AR, Bobadilla JL (1993) *Disease Control Priorities in Developing Countries* The World Bank, Oxford University Press

Jamison DT, Saxenian H, Bergevin Y (1995) Investing in health wisely: the role of needs-based technology assessment *Int J Tech Ass Hlth Care* 11, 4, 673-84

Jelsma J, Chivaura V, de Weerdt W, de Cock P. (2000) A bridge between cultures: a report on the process of translating the EQ-5D instrument into Shona *Sth Af J Physio* 56, 3-9

Johanesson M (1992) On the discounting of gained life-years in cost-effectiveness analysis *Int J Tech Ass Hlth Care* 8, 2, 359-364

Johanesson M, Johansson P-O (1997) Is the valuation of a QALY gained independent of age? Some empirical evidence *J Hlth Econ*, 16, 589-599

Jonsson B (1998) The economic impact of diabetes *Diabetes Care* 21, suppl 3, C7-10

Kale OO (1998) Onchocerciasis: the burden of disease *Ann-Trop Med Parasitol* 92, suppl 1, S101-15

Kamm FM (1993) *Morality/Mortality Vol 1: Death and Whom to Save from It* Oxford University Press

Kapiriri L, Norheim O (2000) *Health Planners' Perceptions of the Usability of the Burden of Disease (BOD) in Priority Setting: the Ugandan Experience* presentation at the International Conference on Priorities in Health Care, Amsterdam, November 2000

Kind P, Rosser R, Williams A (1982) Valuation of quality of life: some psychometric evidence. In Jones Lee M (ed.) *The Value of Life and Safety* North Holland

Kindig DA (1999) Purchasing population health: aligning financial incentives to improve health outcomes *Nursing Outlook* 47, 1, 15-22

Kirsch K, McGuire A (2000) Establishing health state valuations for disease specific states: an example from heart disease *Hlth Econ* 9, 2, 149-158

Koch T (2000) Life quality vs the 'quality of life': assumptions underlying prospective quality of life instruments in health care planning *Soc Sci Med* 51 3, 419-428

Lal D (1994) *The Role of the Public and Private Sectors in Health Financing* Human Resources Development and Operations Policy Working Paper, World Bank, June 1994

Lancet (1993) World Bank's cure for donor fatigue *Lancet* 342(8863) 63-64

Lancet (1997) From what will we die in 2020? *Lancet* 349, 1263

Lara-Munoz MC, Ponce de Leon S, de al Fuente R (1995) Conceptualisation and measurement of the quality of life of cancer patients *Rev. Invest. Clin*, 47, 4, 315-327

138

Laurell AC, Arellano LO (1996) Market commodities and poor relief: the World Bank proposal for health *Int J H Serv* 26, 1, 1-18

Lopez AD (1987) Who dies of what? A comparative analysis of mortality conditions in developed countries around 1987 *Rapp Trim Stat Sanit Mond* 43, 2, 105-114

Lopez AD (1990) Causes of death: an assessment of global patterns of mortality around 1985 *Rapport Trimestriel de Statistiques Sanitares Mondiales* 43, 2, 91-104

Lozano R, Murray CJL, Frenk J, Bobadilla J-L (1995) Burden of diseases assessment and health system reform: results of a study in Mexico *J Int Dev* 7, 3, 555-564

Macran S (in press) Test-retest performance of the EQ5D. In Brooks R, Rabin R, de Charro F (eds.) *The Measurement and Valuation of Health Status Using the EQ-5D: European Perspective* (Evidence from the EuroQol BIOMED research program), Kluwer Academic Publishers, Dordrecht, The Netherlands

Mansourian BG, for DALY Review Group (1996) WHO Advisory Committee on Health Research (AHCR) News: DALY review, *Bull WHO*, 74, 333-4

Mara DD, Feacham RGA (1999) Water- and excreta-related diseases: unitary environmental classification *J Env Eng-asce* 125, 4, 334-339

Marseille E (1996) Cost-effectiveness of cataract surgery in a public health eye care programme in Nepal *Bull WHO* 74, 3, 319-24

Marseille E, Kahn JG, Mmiro F, Guay L, Musoke P, Fowler MG, Jackson JB (1999) Cost effectiveness of single dose nevirapine regimen for mothers and babies to decrease vertical HIV-1 transmission in sub-Saharan Africa *Lancet* 354(9181), 830-9

Marseille E, Khan JG, Saba J (1998) Cost-effectiveness of antiviral drug therapy to reduce mother-to-child HIV transmission in sub-Saharan Africa *AIDS* 12, 8, 939-48

Martens WJ, Niessen LW, Rotmans J, Jetten TH, McMichael AJ (1995) Potential impact of global climate change on malaria risk *Environ Health Perspect* 103, 5, 458-64

Mathers C, Vos T, Stevenson C (1999) *The Burden of Disease and Injury in Australia: Summary Report* Australian Institute of Health and Welfare, Canberra AIHW cat. No. PHE18

Mathers CD, Sadana R, Salomon JA, Murray CJL, Lopez AD (2000) *Estimates of DALE for 191 countries: methods and results*, Global Programme on Evidence for Health Policy Working Paper No 16, WHO, Geneva

McMichael AJ (1995) Book review of Murray CJL and Lopez AD (eds.) 'Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Interventions Packages' *Pop Stud* 49, 3, 554-5

McNamee (1993) Health costs *Lancet* 342(8863), 105-106

Meerding WJ, Polder JJ, Koopmanschap A, van der Maas P (1998) Demographic and epidemiological determinants of healthcare costs in the Netherlands: cost of illness study *Br Med J* 317, 111-115

Meltzer MI, Rigau-Perez JG, Clark GG, Reiter P, Gubler DJ (1998) Using disability-adjusted life years to assess the economic impact of dengue in Puerto Rico: 1984-1994 *Am J Trop Med Hyg*, 59, 2, 265-271

Miller MA (1998) An assessment of the value of Haemophilus influenzae type b conjugate vaccine in *Asia Pediatr Infect Dis* J 17, suppl 9, S152-9

Mooney G, Wiseman V (2000) Burden of disease and priority setting *Hlth Econ* 9, 5, 369-372

Morrow RH, Bryant JH (1994) Measuring and valuing human life: cost-effectiveness, equity and other ethics-based issues. In Bankowski Z and Bryant JH (eds.) *Poverty, Vulnerability, and the Value of Human Life: a Global Agenda for Bioethics* highlights and papers of the XXVIIIth CIOMS conference, Ixtapa, Guerrero State, Mexico, 17-22 April, 159-174

Mosley WH (1994) Population change, health planning and human resource development in the health sector *Wrld Hlth Stats Quart* 47, 26-30

Muennig P, Pallin D, Sell RL, Chan MS (1999) The cost-effectiveness of strategies for the treatment of intestinal parasites in immigrants *New Eng J Med*, 340, 10, 773-9

Murray CJL (1990) Rational approaches to priority setting in international health *J Trop Med Hyg* 93, 5, 303-311

Murray CJL (1994) Quantifying the burden of disease: the technical basis for disability adjusted life years. In Murray CJL, Lopez AD (eds.) (1994a) *Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Interventions* World Health Organisation, Geneva, 3-19

Murray CJL (1995a) Towards an analytical approach to health sector reform *Health Policy* 32, 1-3, 93-109

Murray CJL (1995b) Towards an analytical approach to health sector reform. In Berman P (ed.) *Health Sector Reform in Developing Countries: Making Health Development Sustainable* Harvard University Press, 121-142

Murray CJL (1996) Rethinking DALYs. In Murray CJL and Lopez AD (eds.) *The Global Burden of Disease: a Comprehensive Assessment of Mortality and Disability for Diseases, Injuries and Risk Factors in 1990 and Projected to 2020* Harvard University Press, Cambridge, 1-97

Murray CJL (1998) Testimony of Christopher Murray to the Senate Committee on Labor and Human Resources Subcommittee on Public Safety, March 3 1998 http://www.senate.gov/~labor/hear/03038hrg/murray.htm

Murray CJL, Acharya AK (1997) Understanding DALYs *J Hlth Econ* 16, 6, 703-730

Murray CJL, Evans D, Acharya A, Baltussen R (2000) Development of WHO guidelines on generalised cost-effectiveness analysis *Health Economics*, 9, 2, 235-251

Murray CJL, Kreuser J, Whang W (1994) Cost-effectiveness analysis and policy choices: investing in health systems. In Murray CJL, Lopez AD (eds.) (1994a) *Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Interventions* World Health Organisation, Geneva, 181-192

Murray CJL, Lopez AD (eds.) (1994a) *Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Interventions* World Health Organisation, Geneva

Murray CJL, Lopez AD (1994b) Quantifying disability: data, methods and results. In Murray CJL, Lopez AD (eds.) (1994a) *Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Interventions* World Health Organisation, Geneva, 55-96

REFERENCES

Murray CJL, Lopez AD (eds.) (1996a) *The Global Burden of Disease: a Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020,* The Global Burden of Disease and Injury Series, Volume I, Harvard School of Public Health on behalf of the World Health Organisation and The World Bank, Harvard University Press

Murray CJL, Lopez AD (1996b) *Global Health Statistics: a Compendium of Incidence, Prevalence and Mortality Estimates for over 200 Conditions Cambridge,* Harvard University Press

Murray CJL, Lopez AD (eds.) (1996c) *Health Dimensions of Sex and Reproduction: the Global Burden of Sexually Transmitted Diseases, HIV, Maternal Conditions, Perinatal Disorders, and Congenital Anomalies* The Global Burden of Disease and Injury Series, Volume III, Harvard School of Public Health on behalf of the World Health Organisation and The World Bank, Harvard University Press

Murray CJL, Lopez AD (1996d) Evidence-based health policy – lessons from the global burden of disease study *Science* 274(5288), 740-743

Murray CJL, Lopez AD (1996e) The incremental effect of age-weighting on YLLs, YLDs and DALYs: a response. Point of view *Bull WHO* 74, 4, 445-446

Murray CJL, Lopez AD (1997a) The utility of DALYs for public health policy and research: a reply *Bull WHO* 75, 4, 377-381

Murray CJL, Lopez AD (1997b) Mortality by cause for eight regions of the world: global burden of disease study *Lancet* 349(9061), 1269-1276

Murray CJL Lopez AD (1997c) Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: global burden of disease study *Lancet* 349(9062) 1347-1352

Murray CJL, Lopez AD (1997d) Alternative projections of mortality and disability by cause 1990-2020: global burden of disease study *Lancet* 349(9064), 1498-1504

Murray CJL, Lopez AD (1997e) Global mortality, disability and the contribution of risk factors: global burden of disease study *Lancet* 349(9063), 1436-42

142

Murray CJL, Lopez AD (undated) *Progress and directions in refining the global burden of disease approach: a response to Williams* Global programme on evidence for health policy, Working Paper No 1, WHO, Geneva

Murray CJL, Salomon JA, Mathers C (2000) *A critical examination of summary measures of population health* Discussion Paper No 2, Global Programme on Evidence for Health Policy, WHO, Geneva

Murray CJL, Tandon A, Salomon J, Mathers CD (2001) Enhancing cross-population comparability of survey results *GPE Discussion Paper* No 35, WHO, Geneva

Musgrove P (1994) *Cost-Effectiveness and Health Sector Reform* Human Capital Development and Operations Policy, World Bank Working Paper No 43, Washington

Musgrove P (1995) Cost-effectiveness and socialisation of health care. In Berman P (ed.) *Health Sector Reform in Developing Countries: Making Health Development Sustainable* Harvard University Press 143-159

Musgrove P (1997) Poor nutrition and chronic disease *SCN News*, No 14

Musgrove P (2000) A critical review of 'A critical review': the methodology of the 1993 World Development Report, 'Investing in Health' *Hlth Pol & Plan* 15, 1, 110-115

Nord E (1992) Methods for quality adjustment of life years *Soc Sci & Med* 34, 559-569

Nord E (1999) *My Goodness and Yours: a History, and Some Possible Future, of DALY Meanings and Valuation Procedures* paper presented for WHO's global conference on summary measures of population health, Marrakech, Dec 6-9

Nord E, Pinto JL, Richardson J, Menzel P, Ubel P (1999) Incorporating societal concerns for fairness in numerical valuations of health programmes *Hlth Econ* 8, 25-39

Nygaard E (2000) Is it feasible or desirable to measure burdens of disease as a single number? *Rep Hlth Matters* 8, 15, 117-125

Oortwijn WJ, Vondeling H, Bouter L (1998) The use of societal criteria in priority setting for health technology assessment in The Netherlands. Initial experiences and future challenges *Int J Tech Ass Hth Care* 14, 226-236

REFERENCES

Oswald JW (2000) *A Model of Using Cost-Effectiveness Estimates for Priority Setting and Resource Allocation for a Defined Population* presentation at the international Conference on Priorities in Health Care, Amsterdam, November 2000

Paalman M, Bekedam H, Hawken L, Nyheim D (1998) A critical review of priority setting in the health sector: the methodology of the 1993 World Development Report *Hlth Pol & Plan* 13, 1, 13-31

Politi C, Carrin G, Evans D, Kuzoe FA, Cattand PD (1995) Cost-effectiveness of alternative treatments of African gambinese trypanosomiasis in Uganda *Hlth Econ* 4, 4, 273-87

Prabhakar L (1996) Tuberculosis – the continuing scourge of India *Ind J Med Res* 103, 19-25

Preston SH (1993) 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 DC, National Academy Press

Rahi JS, Gilbert CE, Foster A, Minassian D (1999) Measuring the burden of childhood blindness – commentary *Br J Opthalmol* 83, 4, 387-388

Ramaiah KD, Kumar KN, Ramu K, Pani SP, Das PK (1997) Functional impairment caused by lymphatic filariasis in rural areas of South India *Trop Med Int Health* 2, 9, 832-8

Rawls J (1971) *A Theory of Justice* Harvard University Press, Cambridge, Mass.

Reich MR (1995) The politics of agenda setting in international health: child health versus adult health in developing countries *J Int Dev*, 7, 3, 489-502

Remme JHF (1998) Onchocerciasis. In Murray CJL, Lopez AD (eds.) (in press) *The Gobal Epidemiology of Infectious Diseases* Cambridge MA, Harvard University

Reynolds Whyte S, Ingstad B (1995) Disability and culture: an overview. In Reynolds Whyte S, Ingstad B (eds.) (1995) *Disability and Culture* University of California Press, Berkeley

Rock M (2000) Discounted lives? Weighing disability when measuring health and ruling on 'compassionate' murder *Soc Sci Med* 51, 407-417

144

Rosser R (1990) *Is it Too Late to Prevent the Abuse of Quality of Life Measurement?* Annual Public Health Lecture Series, London School of Hygiene and Tropical Medicine

Ruzicka LT, Lopez AD (1990) The use of death statistics for health situation assessment: national and international experiences *Rapp Trim Stat Sanit Mond* 43, 4, 249-258

Sachs J (2001) *The Links of Public Health and Economic Development* Office of Health Economics Annual Lecture, London

Sadana R, Mathers C, Lopez AD, Murray CJL, Iburg K (2000) *Comparative Analyses of More than 50 Household Surveys on Health Status* Global Programme on Evidence for Health Policy Working Paper No 15, World Health Organisation

Sadana R, Tandon A, Murray CJL, Serdobova I, Cao Y, Xie WJ, Chatterji S, Üstün TB (2002) *Describing Population Health in Six Domains: Comparable Results from 66 Household Surveys* GPE Discussion Paper, No 43, WHO, Geneva

Salomon JA, Tandon A, Murray CJL (2001) *Using Vignettes to Improve Cross-population Comparability of Health Surveys: Concepts, Design, and Evaluation Techniques* GPE Discussion Paper, No 41, WHO, Geneva

Sauerborn R, Schmidt CM, Baltussen R, Snow R, Gauss T, Vargas V, Somerfield J (1998) *Towards Community Based Assessment of the Burden of Illness* presented at the Global Forum on Health Research, Geneva, June 26-28 1998

Sayers BM, Fliedner TM (1997) The critique of DALYs: a counter reply *Bull WHO* 75, 4, 383-4

Sen A (1997) *On Economic Inequality: Expanded Edition with a Substantial Annex by James Foster and Amartya Sen* Oxford University Press, Oxford

Shepard DS (1994) Economic analysis of investment priorities for measles control *J Infect Dis* 170, suppl 1, S56-62

Shore MF (1999) Replacing cost with value *Harv Rev Psych* 6, 6, 334-336

Sprangers MAG, Schwartz CE (1999) Integrating response shift into health-related quality of life research: a theoretical model *Soc Sci Med*, 48, 11, 1507-1515

REFERENCES

Squire SB (1997) Tuberculosis control programmes: developing countries *Eur Respir Monogr* 2, 4, 116-143

Stouthard MEA, Essink-Bott M-L, Bonsel GJ, Barendregt JJ, Kramer PGN, van de Water HPA, Gunning-Schepers LJ, van de Maas PJ (1997) *Disability Weights for Diseases in The Netherlands,* Department of Public Health, Erasmus University, Rotterdam

Stouthard MEA, Essink-Bott M-L, Bonsel GJ on behalf of the Dutch Disability Weights (DDW) Group (2000) Disability weights for diseases: a modified protocol and results for a Western European region *Eur J Pub Hlth* 10, 24-30

Sundby J (1999) Are women disfavoured in the estimation of disability adjusted life years and the global burden of disease? *Scand J Pub Hlth* 27, 4, 279-285

Taylor SM (1997) Review of Murray CJL and Lopez AD (eds.) 'Global comparative assessments in the health sector: disease burden, expenditures and interventions packages' *Prof Geog*, 49, 1, 155

Tengs TO, Adams ME, Pliskin JS, Safran DG, Siegel JE, Weinstein MC, Graham JD (1995) Five hundered life-saving interventions and their cost-effectiveness *Risk Anal* 15, 3, 369-90

Thaineua V, Sirinirund P, Tanbanjong A, Lallement M, Soucat A, Lamboray JL (1998) From research to practice: use of short course zidovudine to prevent mother-to-child HIV transmission in the context of routine health care in Northern Thailand *Southeast Asian J Trop Med Pub Hlth* 29, 3, 429-42

The Economist (2002) For 80 cents more, Aug 17-23, 2002 p20-22

Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q (1996) Multi-attribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2 *Med Care* 34, 7, 702-722

Tsuchiya A (1999) Age-related preferences and age weighting health benefits *Soc Sci Med* 48, 2, 267-276

Ubel PA, Richardson J, Menzel P (2000) Societal value, the person trade-off, and the dilemma of whose values to measure for cost-effectiveness analysis *Hlth Econ* 9, 127-136

Ubel PA, Richardson J, Prades J-LP (1999) Life-saving treatments and disabilities *Int J Tech Ass Hlth Care* 15, 4, 738-748

Ugalde A, Jackson JT (1995) The World Bank and international health policy: a critical review *J Int Dev* 7, 3, 525-542

United Nations (1982) *United Nations, 1982 Model Life Tables* Department of International Economic and Social Affairs, United Nations, 76-117

Üstün TB, Chatterji S, Villanueva M, Bendih L, Çelik C, Sadana R, Valentine N, Ortiz J, Tandon A, Saloman J, Cao Y, Jun XW, Ozaltin E, Mathers C, Murray CJL (2001) *WHO Multi-country Survey Study on Health and Responsiveness* GPE Discussion Paper No 37, WHO, Geneva

Üstün TB, Rehm J, Chatterji S, Saxena S, Trotter R, Room R, Bickenbach J (1999) Multiple-informant ranking of the disability effects of different health conditions in 14 countries *Lancet* 354(9173), 111-5

Van Dam CJ (1995) HIV, STD and their current impact on reproductive health: the need for control of sexually transmitted diseases *Int J Gynaecl Obstet* 50, suppl 2, S121-S129

Varley RC, Tarvid J, Chao DN (1998) A reassessment of the cost-effectiveness of water and sanitation interventions in programmes for controlling childhood diarrhoea *Bull WHO* 76, 6, 617-631

Varmus H (1997) Statement to Department of Health and Human Services, before the Subcommittee on Public Health and Safety, Committee on Labor and Human Resources, United States Senate, May 1st http://www.nih.gov/about/director/tetim2.htm

Varmus H (1999) Evaluating the burden of disease and spending the research dollars of the National Institutes of Health *New Eng J Med* 340, 1914-1915

Victora CG, Van Haecke P (1998) Vitamin K prophylaxis in less developed countries: policy issues and relevance to breastfeeding promotion *Am J Pub Hlth* 88, 2, 203-9

Vos T, Begg S (1999) *The Victorian Burden of Disease Study: Mortality* Department of Human Services, Victoria, Australia

Vos T, Mathers C (1998) Appropriate yardsticks for measuring population health *Aust N Z J Pub Hlth* 22, 4, 515-7

Wa Thiong'o N (1993) *Moving the Centre: the Struggle for Cultural Freedoms* East African Educational Publishers, Nairobi

Walker D, Fox-Rushby J (2000) Economic evaluation of communicable disease interventions in developing countries: a critical review of the published literature *Hlth Econ* 9, 8, 681-698

Walt G (1994) *Health Policy: an Introduction to Process and Power* London, Zed Books

Wasserman D (2000) *Genetic Testing, Disability, and Quality of Life* research proposal submitted to and accepted by the Ethical, Legal and Social Implications (ELSI) Programme, National Institute for Human Genome Research, US. Accessible from dw79@umail.umd.edu

Weinstein MC (1988) A QALY is a QALY is a QALY – Or is it? *J Hlth Econ*, 7, 3, 289-290

WHOQOL Group (1998a) The World Health Organisation quality of life assessment (WHOQOL): development and general psychometric properties *Soc Sci Med* 46, 12, 1569-1585

WHOQOL Group (1998b) Development of the World Health Organisation WHOQOL-BREF quality of life assessment *Psych Med*, 28, 3, 551-558

Williams A (1999) Calculating the global burden of disease: time for a strategic re-appraisal *Hlth Econ* 8, 1-8

Williams A (2000) Comments on the response by Murray and Lopez *Hlth Econ* 9, 1, 83-86

World Bank (1993) *World Development Report: Investing in Health*, Oxford University Press, Oxford

World Health Organisation (1980) International classification of impairments, disabilities and handicaps: a manual of classification, WHO, Geneva

World Health Organisation (2000) *The World Health Report: Health Systems: Improving Performance,* WHO, Geneva www.dinf.ne.jp/doc/prdl/othr/apdrj/z13jo0300/z13jo0301.html

Yu M-Y, Sarri R (1997) Women's health status and gender inequality in China *Soc Sci Med* 45, 12, 1885-1898

APPENDIX 1
METHOD AND RESULTS OF SEARCH FOR DALY BIBLIOGRAPHY

Method

The search for literature involved four methods: search of electronic data bases; contact with researchers known to be working in the field; search of the world wide web (on 31st March and 10th November 2000); and a hand search of selected journals in year 2000 journals.

Databases searched: Medline (Jan1991- March 2000), Econlit (1969-Dec 1999), Embase (1989 – Jan 2000), CINAHL (Nursing and Allied Health) (1982 – Jan 2000), Psychological abstracts (1989 – Dec 1999), SIGLE (Grey literature in Europe) (1980 – June 1999), Sociological abstracts (1986 – Dec 1999), BIDS ISI (1999 – March 2000), Aslib Index to Theses, and the world wide web.

Search terms used for all above: Disability Adjusted Life Yea**, DALY*

Authors searched: Murray C, Murray CJ, Murray CJL, Lopez A, Lopez AD, Jamison D, Jamison DT, Vos T, Bevan G, Bonsel G.

Researchers contacted, and replied: Sadana R, Nord E, Bonsel G, Bevan G, Evans D, Essink-Bott M-L, Hanson K, Musgrove P, Mathers C, Wasserman D, Gold L, Robinson S.

Specific additional websites searched: World Bank, World Health Organisation, Harvard University, London School of Economics. Year 2000 journals included: Health Economics, Health Policy and Planning; Social Science and Medicine, Bulletin WHO.

Papers selected

Papers were selected from the electronic search on the basis of a review of abstracts or titles (where abstracts were not available). Papers were excluded if not in English. Of the total 912 references found in the search, 331 were selected after the following exclusions: 357 related to the wrong authors (e.g. Daly) or to authors writing on other subjects and the remainder were duplicated references. A further 83 were then rejected on review as disability adjusted life years were not mentioned or mentioned only in passing, or authors were writing on a related but not directly relevant subject. The final bibliography for DALYs can be obtained directly from the author at Julia.Fox-Rushby@LSHTM.ac.uk. The bibliography is divided into ten sections. These are as follows (with the number of references read in regular type, and number not found but judged relevant in italics and * in the bibliography): journal papers (n= 142, *13*), books (n=5), chapters

Table A1 **Results of the literature search**

	Total number of references found (all abstracts or titles reviewed)	Total references selected for full review
Disability Adjusted Life Years	320	208
Specific authors	550	56
Websites	Not recorded	25
Personal collection, not available from other searches	9	9
From search of references	33	33
Total	Not applicable	331

in books (n=8, *1*), letters in journals, book reviews/editorials (n=14, *1*), discussion/working papers (n=13, *3*), reports (n=16, *7*), items in newsletters (n=8), items on websites (n=25, *3*), and 'other' (n=8).

APPENDIX 2
AN INTRODUCTION TO EQ5D QALYS

The EQ-5D is a generic instrument for describing, measuring and valuing health outcomes that was originally designed for use along-side other measures of health and disease in five countries (United Kingdom, The Netherlands, Sweden, Norway and Finland). The EuroQol Group, an eclectic collection of academic researchers[1] led initially by the University of York economist Professor Alan Williams[2], designed it.

The EuroQol website[3] describes the EQ5D as applicable to a wide variety of health conditions and treatments and providing a 'simple descriptive profile and a single index value for health status that can be used in clinical and economic evaluation of health care as well as population health surveys'. The questionnaire is comprised of three main sections that can be used independently:

1. The descriptive health outcome measure

This is based on five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each with three levels of severity (e.g. none, some/moderate, extreme/unable). It is designed to be self-completed by respondents as a postal questionnaire, within clinics or as part of an interview. It takes about two minutes to complete.

The dimensions of EQ-5D were selected on the basis of a review of other generic HRQOL measures at that time (EuroQol Group, 1991) and the experiences of the EuroQol Group members. There was general agreement that the following dimensions were included: mobility, daily activities and self-care, psychological functioning, social and role performance, and pain or other health problems (Gudex, 1996). The selection of levels took place during Group discussions and involved on-the-spot translations by the developers for key words in all languages of development. The levels were designed with the aim of covering a wide range of severity and to ensure ordinality (Gudex, 1996). The first descriptive system emerged in 1988 with six dimensions. Following initial experimentation, a large survey of lay concepts in the UK, and further deliberation about the impact of dimensions and levels on the number of potential health

---

1 Of whom I am one!

2 Williams has been instrumental in the development and promotion of QALYs for health care decision-making in the UK.

3 http://www.euroqol.org/

Table A2 **EQ5D classification system and associated coefficients**   **151**

Dimension	Level	Statement	Coefficient
Mobility	1√	No problems walking	0
	2	Some problems walking about	0.069
	3	Confined to bed	0.314
Self-care	1√	No problems with self-care	0
	2	Some problems washing or dressing self	0.104
	3	Unable to wash or dress self	0.214
Usual activities	1	No problems with performing usual activities (e.g. work, study, housework, family or leisure activities)	0
	2√	Some problems performing usual activities	0.036
	3	Unable to perform usual activities	0.094
Pain/ discomfort	1	No pain or discomfort	0
	2√	Moderate pain or discomfort	0.123
	3	Extreme pain or discomfort	0.386
Anxiety/ depression	1	Not anxious or depressed	0
	2	Moderately anxious or depressed	0.071
	3√	Extremely anxious or depressed	0.269

Source: Adapted from Dolan et al. (1995).

states (and hence requirements for valuation), the current EQ-5D was developed (Gudex, 1996).

2. The visual analogue rating scale
This is a vertical 20cm scale with end points labelled at 'best/worst imaginable health' with points numbered along its length between 0 and 100. It takes about two minutes to complete, and it most often used along with the descriptive health outcome measure above.

3. The valuation questionnaire
This part consists of a minimum of 14 health states that are valued, along with the states dead and unconscious, using either a visual analogue scale or the time trade-off approach. It is not designed for use in clinical trials

or population surveys and is used only by researchers wishing to develop valuations of health to create an index. There currently exist 14 valuation sets from eight different countries using samples of the general population.

Most research only uses the first two parts of EQ5D. Results from the first part provide a profile of an individual's health state. The value for any one state is 'read off' using the results from one of the existing valuation sets. Hence the value of change between two states is also easily calculated. Alternatively, using a specific valuation set in the UK, Dolan et al. (1995) provided a formula to facilitate calculation, whilst at the same time showing the basis from which the values had been calculated.

Table A2 shows a series of statements and a pattern of statements that might have been ticked by a respondent in the questionnaire. It also shows the coefficients for each statement. In this example, the health state is known as 11223.

To calculate the index score for health state 11223, the following algorithm is used:

Full health	= 1.0
Constant term	– 0.081
Mobility (1)	– 0
Self care (1)	– 0
Usual activities (2)	– 0.036
Pain/discomfort (2)	– 0.123
Anxiety/depression (3)	– 0.236
N3	– 0.269
Estimated value	**= 0.255**

Three particular aspects to note are that: a constant is needed; the coefficient for any dysfunction is subtracted from 1.0 and that an additional weighting (N3) is subtracted for any health state which includes a level 3 severity.

The psychometric performance of the EQ5D has been compared with other preference and non-preference based health-related quality of life measures. Essink-Bott et al. (1997) compared the Nottingham Health Profile, the Medical Outcomes Study 36-item short form, and the COOP/WONCA charts with the EQ5D in a group of migraine sufferers and matched control group in the Netherlands. They found that no one instrument performed worse or better than any other. Coons et al. (2000) reviewed published evidence comparing the EQ5D with two other pref-

erence-based measures (the quality of well-being index and health utilities index) according to six criteria[4]. With respect to reliability and validity they concluded that, whilst the supporting evidence for the EQ5D's reliability and validity was acceptable, evidence from validity testing was strongest for the quality of well-being index. More recently Macran (in press) has argued that in a group of arthritis sufferers and general population sample in the UK the EQ5D exhibited acceptable measurement properties over a three-month period and that the EQ VAS was sensitive to self-reported clinical change.

The EQ-5D has been used in a range of studies including population surveys, as a measure of change in patient health over time and as a measure of the impact of health interventions for use in economic analysis. In addition to the five countries in which it was developed, the EQ-5D has also been used in a variety of countries in Europe, Asia, Latin America, Australasia and Africa. Its use in an increasing number of settings is leading to greater questioning about what the EQ-5D is measuring in each setting (e.g. Jelsma et al. 2000).

Further details of how to use the EuroQol can be gained from the EuroQol business management through the website.

---

4 Including: the conceptual and measurement model; reliability; validity; respondent and administrative burden; alternative forms; and cultural and language adaptations.

APPENDIX 3
FORMULAS FOR CALCULATING DALYS

The formulas here are taken from Murray (1996, pp65-6)[5].

$$YLLs[r, K, ß] = \frac{KCe^{ra}}{(r+ß)^2} \{e^{-(r+ß)(L+a)}[-(r+ß)(L+a)-1] - e^{-(r+ß)a}[-(r+ß)a - 1]\}$$

$$+ \frac{(1 - K)}{r}(1 - e^{-rL})$$

where:

K = age weighting modulation factor
C = constant
r = discount rate expressed as a decimal
a = age of death
ß = parameter from the age weighting function
L = standard expectation of life at age a

The formula for YLDs[r, K, ß] differs in the addition of D (the disability weight) and in different interpretations of a and L, and is shown below:

$$YLDs[r, K, ß] = D\{\frac{KCe^{ra}}{(r+ß)^2} \{e^{-(r+ß)(L+a)}[-(r+ß)(L+a)-1]$$

$$- e^{-(r+ß)a}[-(r+ß)a - 1]\} + \frac{(1-K)}{r}(1 - e^{-rL})\}$$

where:

K = age weighting modulation factor
C = constant
r = discount rate expressed as a decimal
a = age of onset of disability
ß = parameter from the age weighting function
L = duration of disability
D = disability weight

5 This broadly follows Murray's recommended nomenclature of YLLs[r,K] in order to ensure that DALYs are compared on a like basis. However, given that anyone could choose a different age weighting, I have added the rate of age-weighting.

With the discount rate set to zero and a uniform age weighting, the length of life lost due to death at a particular age is L (i.e. the life expectancy at age a) and the years of life lived with disability as a result of ill-health commencing at age a is the product of the disability weight and the duration of disability at age a, that is DL.

Tabular Review of the Use of DALYs in Studies Calculating the Burden of Disease

Authors	Hollinghurst et al.	Mara and Feacham	Mathers et al.
Year	2000	1999	1999
Aim of paper	To estimate the 'avoidable'[6] (using 3 approaches to define this) burden of disease as a proportion of total burden of disease	To reappraise aging classification of water and excreta related diseases given newly emergent pathogens	To calculate the burden of disease in Australia
Disease focus	Heart disease, stroke, cataract, benign prostatic hyperplasia, arthritis of the hip, peptic ulcer, depressive illness	Water and excreta related diseases	176 diseases, injuries and risk factors
Interventions considered	A range of treatments	Not applicable	Not applicable
Geographical focus	South and West region in England	Global	Australia
Population considered	General population in regional health authority (6.3 million)	All ages	All Australians in 1996
DALY type	A development and variant of DALYs. Neither formula was used directly	Mark 2	Mark 2
Disability weight	GBD weights for untreated conditions for some and other data (cited) to weight or re-weight other conditions	0.066	Different weights depending on condition
Age weight	No	0.04	No
Discount rate	3%	3%	3%
DALYs [0.03,1, 0.04] presented?	No	Yes	No

Which life expectancy?	82.5 (females) 80 (males) SEYLL	82.5 (females) 80 (males)	81.5 for males and 85.7 for females at birth
Presentation of DALYs	Total/'avoidable' DALYs by disease. Avoidable DALYs/ 100,000 by cause and health authority	Total DALYs, and percent contributions from water/excreta related DALYs	YLL and YLD, age standardised, DALYs per 1,000 by disease, total burden by: gender, age, socio-economic group. Attributable DALYs as a proportion of total DALYs
Source of disability weights	GBD (where possible)	GBD 1996	Dutch Disability Weights, and if not available then GBD weights
Source of life expectancy	GBD	GBD 1996	Australian Cohort Life Expectancies for 1996
Source of sequelae	Mostly from additionally cited evidence	GBD 1996	Dutch study and GBD
Sensitivity analysis on DALY assumptions?	Yes. Discount rate, age weighting and the interaction between them	Not included in this paper	Discount rate
Funders	Regional Health Authority in the South and West of England	Not stated	Australian Institute of Health and Welfare
Main conclusion	Prioritise treating depression in general practice, and R&D in preventing heart disease in the elderly. More data needed on severity	Combined classifications of water/excreta-related diseases is more useful because of the bi-causal nature of many diseases	YLLs account for 57 percent total burden and 51 percent in females. Depression is the leading cause of non-fatal disease and ischaemic heart disease for total burden. Leading risk factor is tobacco smoking. More information on cost-effectiveness needed

Note: GBD = global burden of disease.

6 This is a variant of the burden of disease by accounting for the proportion of treatment deemed successful.

Tabular Review of the Use of DALYs in Studies Calculating the Burden of Disease (*continued*)

Authors	Gwatkin et al.	Meltzer et al.	Stouthard et al.
Year	1999	1998	1997
Aim of paper	To estimate the burden of disease among the richest and poorest 20% of the global population	To calculate DALYs lost to dengue during epidemic and non-epidemic years (1984–94)	To investigate reliability, validity and usability of the disability weighting method. To develop and apply a set of coherent weights
Disease focus	(Non) Communicable diseases and injury	Dengue	32 indicator conditions from 52 diagnostic groups + activities of daily living limitations in the elderly
Interventions	Not applicable	Not applicable	Not applicable
Geographical focus	20% richest/poorest countries (average per capita income adjusted for purchasing power)	Puerto Rico	The Netherlands
Population considered	The poor	Country	45 doctors, 15 lay people with academic background but no medical knowledge
DALY type	Mark 2	Mark 1	A variant of Mark 2
Disability weight	No disability weights used – only deaths were converted to DALYs.	0.81	New weights developed for different set of indicator conditions more relevant to Dutch population[2]
Age weight	Not stated	0.04	Not applicable
Discount rate	Not stated	3%	Not applicable
DALYs [0.03, 1, 0.04] presented?	Not stated (although did cite GBD approach)	Yes	Not applicable
Which life expectancy?	Cohort	44 years, at age of 33	Not applicable

Presentation of DALYs	Percent of total DALY loss	DALYs lost / million pop for each year, using 3 year moving average. Mean, min, max, 5th and 95th percentiles DALYs lost by year. Log base 10 of DALYs for all diseases	Just the disability weights were presented, therefore no DALYs
Source of disability weights	Not applicable	Class 5 of World Bank classification. Number of days in severity class estimated from literature and local surveys	4 group exercises
Source of life expectancy	UN and World Bank sources on life expectancy and numbers in population cohorts. Life expectancy converted to death rates via life tables and 'adjusted' in line with figures in GBD	US Bureau of Census 1996, Statistical Abstract of the US 1996	Not applicable
Source for sequelae	Not applicable	GBD and added classic dengue, not only dengue hemorrhagic fever (DHF) as in GBD	Full disease staging was given to weighting panels
Sensitivity analysis on DALY assumptions?	No	Disability weight, time lost by severity condition, multiplication factor used for case load	95 percent confidence intervals provided for each disease weight
Funders	World Bank, Global Forum for Health Research, Governments of Norway and Switzerland, World Health Organisation	Not stated	Dutch Ministry of Health, Welfare and Sports
Main conclusion	Health policy based on average estimates of GBD is likely to shift resources away from poor to rich countries	Using only DHF underestimates the burden of disease. Classic dengue is the cause of greatest burden	Weights for 175 disease stages from 52 diseases provided. Testing of reliability/validity needed

Note: GBD = global burden of disease.

2 There were a number of fundamental differences: more and different indicator conditions were weighted; some 'conditions' rather than diseases were weighted; disease stage was described in terms of a variant of the EQ5D (6 dimensions with 3 levels per dimension, including an additional cognitive dimension); the addition of a rating scale exercise following the person trade-off tasks; the aim of discussion of personal valuations was not to reach a consensus but to allow well-considered valuations; and eventually 11 classes between 0 and 1 were outlined.

Tabular Review of the Use of DALYs in Studies Calculating the Burden of Disease (*continued*)

Authors	Chan	Martens et al.
Year	1997	1995
Aim of paper	Quantification of the disease burden from intestinal nematode infections, and predicting cost-effectiveness of treating school children	To model the effects of projected changes in temperature and precipitation on mosquito and parasite characteristics and their potential impact on malaria risk
Disease focus	Intestinal nematode infections	Malaria
Intervention focus	School based chemotherapy	Not applicable
Geographical focus	Global estimates, as well as consideration of a hypothetical population	Global, particularly high and low endemic malarious areas
Population considered	A 'hypothetical community, assumed to have a typically Chinese population distribution'	General population in high and low endemic areas
DALY type	Mark 1	Not stated (probably Mark 1)
Disability weight	Not stated	0.6, for 7 days, twice a year
Age weight	The GBD approach to DALYs is explained, although it is not clear for the hypothetical example	Yes
Discount rate	The GBD approach to DALYs is explained, although it is not clear for the hypothetical example	3%
DALYs [0.03, 1, 0.04]] presented?	Yes for global estimates, unclear for hypothetical example	Not stated
Which life expectancy?	82.5 (females) 80 (males) for global estimates, and presumably estimates from China for the cost-effectiveness estimates	Present upper life expectancy for humans
Presentation of DALYs	DALYs lost per 100,000 of the total (adult and child) population	DALYs per 1,000 population

Source(s) of disability weights	GBD for global estimates, unclear for hypothetical example	Not stated
Source(s) of life expectancy	GBD for global estimates, not stated for hypothetical example	Not stated, but Coale et al. (1983) in references
Source(s) for sequelae	Adapted from de Silva et al. (their ref 21) and given in detail in paper	Not stated
Sensitivity analysis on DALY assumptions?	Referenced another paper (de Silva et al.)	No, only on climatic and vector characteristics
Funders	Wellcome Trust, Edna McConnell Clark Foundation	Dutch National Research Programme on Global Air Pollution and Climatic Change
Main conclusions	Ability to calculate burden, costs of provision are important to aiding operationally realistic services, when cheap cost-effective drugs are available	There is a real risk of reintroducing malaria into nonmalarial areas and increasing the incidence in malaria endemic areas, which will increase the DALYs lost

Note: GBD = global burden of disease.

APPENDIX 5

Tabular Review of the Use of DALYs in Cost-effectiveness Analyses

Authors	Goodman et al.	Muennig et al.	Akhavan et al.
Year	1999	1999	1999
Disease focus	Malaria	Intestinal parasites	Malaria
Interventions considered	Alternative methods of: prevention in childhood and pregnancy; treatment	Albendazole	Monitoring + stratification of population by risk, new diagnosis and treatment protocols, training, supplies of anti-malarials
Geographical focus	Sub-Saharan Africa	US	Brazil
Population considered		Immigrants to US from Asia, Middle East, sub-Saharan Africa, Eastern Europe, Latin America and Caribbean	Amazon basin (16-18 million people)
Aim of paper	Provide estimates of the main prevention and treatment interventions in Sub-Saharan Africa	Compare watchful waiting with universal treatment with albendazole for 5 days	Evaluate the malaria control project during the 7 years of the World Bank loan
DALY type	Mark 2? (if used GBD)	Mark 2	Unclear, but probably Mark 2, as 1996 is cited following age/discounting weights
Disability weight	Actual weights not stated	Not stated	*Privax* – 0.22 *Pfalciparum* – 0.375
Age weight	No	No	0.04
Discount rate	3%	3%	3%
DALYs [0.03, 1, 0.04] presented?	No	No	Presumably, not clear
Which life expectancy?	West African life table with life expectancy at birth of 50 years for low- and middle-income countries and a general pattern life table with life expectancy at birth of 65 years for higher-income countries	Unclear, but possibly same as cited in Murray 1996	80

Presentation of DALYs	Cost per DALY averted	DALYs averted, cost per DALY averted	DALYs per case, total DALYs, (net) cost per DALY
Source of disability weights	GBD 1996	Murray and Lopez 1996	Not stated
Source of life expectancy	UN Model Life tables for developing countries	Possibly Murray 1996	Life tables in Murray 1994, for burden of disease
Source of sequelae	GBD 1996	Murray 1996 and other cited references	Mix of assumptions from literature
Sensitivity analysis on DALY assumptions?	No	No	Yes, age weighting, discounting, case fatality rates
What were results compared to?	World Bank's range of 'good buys'. Estimated total cost of full coverage by interventions as a % of public health budget (using Tanzanian data). Cost per DALY averted for oral rehydration therapy, measles vaccination, other interventions	Costs of smoking cessation. Cost per DALY by area of immigration. Cost of vaccinating against Hep A in health care workers	Cost-effectiveness of treatment and prevention strategies in project compared. Cost-effectiveness ratios compared over time and with ratios presented in 1993 World Development Report
Funders	World Bank	Not stated	Not stated
Main conclusion	All interventions reviewed are cost-effective ways of reducing mortality from malaria	Watchful waiting is more cost-effective per DALY averted	Most DALYs averted by preventing death. Most health gain is from treatment, which is cheaper than prevention, especially where *P.falciparum* prevalent

Note: GBD = global burden of disease.

Tabular Review of the Use of DALYs in Cost-effectiveness Analyses (*continued*)

Authors	Coleman et al.	Greikspoor et al.	Marseille et al.	Varley et al.
Year	1999	1999	1999	1998
Disease focus	Malaria	Visceral Leishmaniasis	HIV-1	Childhood diarrhoea
Interventions considered	Insecticide treated mosquito nets (ITN)	Diagnosis and treatment (with pentavelent antimonium) in hospital and health posts vs no treatment	HIVNET 012 nevirapine regimen to reduce mother/child transmission	Hard/software interventions for water and sanitation
Geographical focus	Sub-Saharan Africa	War-torn area of Western Upper Nile, Sudan	Sub-Saharan Africa	Slum/peri-urban areas with few formal services
Population considered	Hypothetical population of children <10 years in Sub-Saharan Africa	Estimated 300,000 population during 1990-1994	Hypothetical cohort of 20,000 pregnant women	Assumed a city of 1 million inhabitants
Aim of paper	To ascertain at what level of rebound impregnated nets would no longer be cost-effective or effective	To evaluate cost-effectiveness of Médcin Sans Frontiers programme	Cost-effectiveness analysis of: targeted and universal coverage with HIVNET 012; HIVNET with other short-course anti-retroviral interventions	Cost-effectiveness analysis of hardware and software programmes separately and jointly using financial costs
DALY type	Mark 2	Mark 2	Not clear	Mark 2
Disability weight	Not stated	Not used (only YLL)	None cited – just YLLs	0.6
Age weight	No	0.04	Not stated	0.04
Discount rate	3%	3%	3%	3%
DALYs [0.03, 1, 0.04] presented?	No	Yes	Not clear	Yes
Life expectancy	50, at birth	80	44	80
Presentation of DALYs	DALYs averted per child aged 1-119 months, cost per DALYs averted	Average/marginal cost per DALY averted, total DALYs averted	Cost per DALY averted, DALYs saved	Cost per DALY saved, from deaths / morbidity

Source of disability weights	Murray and Lopez, 1996, and other studies cited	Not applicable		Used information from DALY group
Source of life expectancy	UN standard West African life table for women	Coale and Demeny's table from Murray 1996. Also model life table South level 8 and 12, in sensitivity analysis	Haub and Cornelius (1998) World Population Data Sheet = average life expectancy for 18 countries of East Africa	Used information from DALY group
Source of sequelae	Murray and Lopez, 1996, and other studies cited	Not applicable	Chin (1989)	Used information from DALY group
Sensitivity analysis on DALY assumptions?	Yes, for incidence, sequelae, change in mortality	Yes, for life expectancy and case fatality rate	Just life expectancy	Covered changing the case fatality rate and incidence
What were results compared to?	Cut-off points by World Bank for cost per DALY averted estimates	Cut-off points recommended by World Bank and cost per DALY averted for other interventions in other countries	World Bank cost per DALY cut off point. Cost per DALY averted for other interventions in STDs and diseases with high child mortality	Cost per case / death averted
Funders?	WHO, DfID	Not stated	NIAID/NIH and NIMH	USAID
Main conclusion	Age range of rebound crucially affects cost-effectiveness. If it falls amongst 3-6 years olds, ITN may no longer be cost-effective. Long term surveillance crucial	It was a most cost-effective intervention	NIVNET 012 regimen can be highly cost-effective in high sero-prevalence settings	Investing in software is as cost-effective as oral re-hydration, especially where hardware exists

Note: NIAID = National Institute of Allergy and Infectious Disease.

Tabular Review of the Use of DALYs in Cost-effectiveness Analyses (*continued*)

Authors	Victora and van Haecke	Marseille et al.	Miller
Year	1998	1998	1998
Disease focus	Vitamin K deficiency	HIV	Haemophilus Influenzae (Hib) as a cause of bacterial meningitis
Interventions considered	Range of prophylaxis and breastfeeding promotion	Do nothing, 3 different regimens of short-course antiviral regimen	Type b conjugate vaccine
Geographical focus	No one country – global policy	Sub-Saharan Africa	Asia
Population considered	Less developed countries as a whole	A cohort of 100 women	Surviving infants in each birth cohort according to country-specific population and coverage rates for 23 Asian countries
Aim of paper	Cost-effectiveness and need for prophylaxis to reduce vitamin K deficiency	To assess cost-effectiveness of short course treatments to reduce mother to child transmission of HIV	Assesses potential of new Hib vaccine for Asia and barriers to its introduction
DALY type	Mark 1	Mark 2	Mark 1
Disability weight	50% in grade 3, 50% in grade 4	0.123 for HIV and 0.505 for AIDS	Disability adjustment of 25% for each year after recovery
Age weight	Not mentioned	0.04	Not stated
Discount rate	Not mentioned	5% and 3%	Not stated
DALYs [0.03, 1, 0.04] presented?	Unlikely	Yes	Not clear
Life expectancy?	Not stated	Based on an average life span of 52.5 years (midpoint between the life expectancies at birth for men and women in Sub-Saharan Africa)	Not stated

Presentation of DALYs	Total and percent DALYs lost by gender and 3 levels of incidence. Cost per DALY saved	DALYs averted per averted infection, cost per DALY averted	Cost per DALY, cost per DALY as percent GNP per capita
Source of disability weights	Assumption	GBD	Schillinger et al. in GBD series
Source of life expectancy	Thai study and meta analysis. Possibly 1993 WDR too (i.e. Japanese life expectancy) but unclear	1996 World Population Data Sheet, Population Reference Bureau	Not stated
Source of sequelae	Assumptions, one study from Thailand and a meta analysis	Cited 2 papers as justification for assumptions	Not stated
Sensitivity analysis on DALY assumptions?	No	Yes, discount rate	No
What were results compared to?	World Bank 'cut-off' points. Cost-effectiveness of other interventions from around the world to reduce vitamin K deficiency	Decision branches from the model. Results from other cost-effectiveness analysis in low and high mortality areas	Cost per life saved with and without treatment cost savings. Changing price of vaccine. Average cost per fully vaccinated child
Funders	UNICEF evaluation and research office	UNAIDS	Not stated
Main conclusion	Intervention is 'moderately cost-effective'. It should be recommended in developed countries	Antiviral therapy may be cost-effective compared with other health interventions if: HIV prevalence is high; clinical trials confirm estimated efficacies; drug prices are reduced	Countries where Hib vaccine most likely to be efficient are least likely to be able to afford it. Cost-effectiveness is critically dependent on price of vaccine

Note: GBD = global burden of disease.

Tabular Review of the Use of DALYs in Cost-effectiveness Analyses *(continued)*

Authors	Thaineua et al.	Gilson et al.	Marseille
Year	1998	1997	1996
Disease focus	HIV	STDs	Cataracts
Interventions considered	Short course Zidovudine (ZDV) to reduce mother-child transmission	Improved management of STDs by primary health care workers	Out-patient (eye hospital/camp) cataract surgery with local anaesthetic, 7 days follow-up and aphakic spectacles
Geographical focus	Northern Thailand	Mwanza region, Tanzania	Nepal
Population considered	Pregnant women with HIV and population of 517,000 inhabitants in Phayao province	Random cohort of 1,000/12,537 adults aged 15-54	2 million people in Lumbini Zone, south-central Nepal served by a US/Canadian non-governmental organisation
Aim of paper	Delineate the implementation issue of a short course ZDV treatment programme	To assess impact, cost and cost-effectiveness of averting HIV-1 infection with new programme	To compare cost-effectiveness of cataract surgery versus no cataract surgery
DALY type	Mark 1	Mark 1	Used own formula to calculate DALYs
Disability weight	Unclear, possibly only YLLs	0.1 for limited morbidity and 0.9 for severe morbidity	0.39 for blindness and 0.87 following surgery
Age weight	Not stated	0.04	Not mentioned
Discount rate	Not stated	3%	3%
DALYs [0.03, 1, 0.04] presented?	Not clear	Yes	No
Life expectancy	Average life expectancy of 69 years	40.7 and 53.9	Not stated

Presentation of DALYs	Cost per DALY averted	DALYs per infection averted, incremental cost per DALY saved	(Discounted) Cost per DALY
Source of disability weights	Not stated	Unsure, possibly Mulder 1996	Sackett and Torrance 1987. (Un)published studies on visual acuity following treatment and rates of lost/broken glasses
Source of life expectancy	Not given	Tanzanian demographic data and WDR 1993	Annual 1980 mortality rate among non-blind Nepalese for population with same age-sex structure as cataract backlog. National Blindness Survey and others
Source of sequelae	Not mentioned	Unsure, possibly Mulder 1996	Blindness is a steady state, only mortality changing. Following treatment considered data from various follow-up studies
Sensitivity analysis on DALY assumptions?	No	Yes, discount rate, life expectancy	Yes, utility value and mortality rate with blindness, and discount rate
What were results compared to?	Cost per DALY for antenatal and delivery care. Cost per capita per year. Size of health sector budget	Cost-effectiveness for other public health interventions in low income countries and of 'minimum essential clinical services'	Cost per DALY for other interventions from other countries as well as World Bank estimates of cost-effectiveness
Funders	Not stated	AIDS task force of EU, UK Overseas Development Administration, German Centre for International Migration, UK Medical Research Council	Not stated
Main conclusion	It is possible to integrate a short course ZDV into routine mother and child health services	Cost-effectiveness of this intervention compares favourably with childhood immunisation	Cataract surgery even more cost-effective then previously reported. It is one of the most cost-effective public health interventions

Tabular Review of the Use of DALYs in Cost-Effectiveness Analyses (*continued*)

Authors	Horton et al.	Politi et al.	Shepard
Year	1996	1995	1994
Disease focus	Diarrhoeal disease	African Gambinese Trypanosomiasis	Measles
Interventions considered	Breast feeding promotion	Alternative drug therapies	Early one dose vaccine
Geographical focus	Brazil, Honduras, Mexico	Uganda	Varied and worldwide
Population considered	Populations served by 2 hospitals in each country	1,000 hypothetical patients	Hypothetical cohort of 1,000 children from birth to age 59
Aim of paper	Cost-effectiveness analysis of 3 breastfeeding promotion programmes with other health interventions	Cost-effectiveness analysis of switching drugs in treating late stage sleeping sickness	To discuss allocation of resources among implementation, operations research and vaccine research
DALY type	Mark 1 (but only YLLs)	Mark 1	Mark 1
Disability weight	Not stated	Not stated	Not stated – possibly none
Age weight	Not stated but presumably 0.04 because cited 1993 WDR	Not stated	No
Discount rate	3 percent	Not stated	Not stated
DALYs [0.03, 1, 0.04] presented?	Unclear	Not clear	No
Life expectancy	80	SEYLL (81 years at birth)	Not stated
Presentaton of DALYs	Cost per DALY averted	Cost per DALY averted, additional DALYs averted	DALYs averted, cost per DALY averted

Source of disability weights	No weighting for disability	Unclear – possibly only mortality	Not applicable
Source of life expectancy	WDR 1993, and Victora et al. 1987+1989 and other literature	WDR	Unclear
Source of sequelae	Unclear	Unclear	Unclear
Sensitivity analysis on DALY assumptions?	No	No	No
What were results compared to?	Cost per death averted for measles, rotavirus and cholera immunisation, oral re-hydration therapy, hygiene promotion and formula feed. Cost per DALY estimates from 1993 WDR	World Bank's range of 'good buys'. Potential finances available for funding	World Bank cut-off points
Funders	USAID	WHO	WHO and Rockefeller Foundation
Main conclusion	Breastfeeding should be promoted in countries where it is not widely practised in early months and where most women give birth in health facilities	Melarsopol saves lives at low cost, but with side effects. Willingess of donors/government to pay affects choice of intervention	Greater % of existing resources should go on RCTs, to vaccine development for early measles vaccines and before/after studies of existing strategies

Note: WDR = World Development Report.

RECENT OHE PUBLICATIONS

Interpreting and Addressing Inequalities in Health: from Black to Acheson to Blair to…?
by Robert Evans, 2002 (price £7.50)

The Life Cycle of Pharmaceuticals: a Cross-national Perspective
by Patricia Danzon and Jeong Kim, 2002 (price £10.00)

Health Economics: An Introduction to Economic Evaluation (2nd edition)
by Gisela Kobelt, 2002 (price £5.00)

The Links of Public Health and Economic Development
by Professor Jeffrey Sachs, 2001 (price £5.00)

Applied Econometrics for Health Economists – a Practical Guide
by Andrew Jones, 2001 (price £10.00)

Consolidation and Competition in the Pharmaceutical Industry
ed. Hannah Kettler, 2001 (price £10.00)

Don't Look Back? Voluntary and Charitable Finance of Hospitals in Britain, Past and Present
by John Mohan and Martin Gorsky, 2001 (price £10.00)

Capturing the Unexpected Benefits of Medical Research
ed. Clive Pritchard, 2001 (price £10.00)

The Economics of the Private Finance Initiative in the NHS
by Jon Sussex, 2001 (price £10.00)

Why Care about Health Inequality?
by Adam Oliver, 2001 (price £7.50)

Health Care Without Frontiers? The Development of a European Market in Health Services?
by Lyndsay Mountford, 2000 (price £10.00)

Productivity Costs: Principles and Practice in Economic Evaluation
by Clive Pritchard and Mark Sculpher, 2000 (price £10.00)

Improving Population Health in Industrialised Nations
ed. Jon Sussex, 2000 (price £10.00)