QALY maximisation and people's preferences: a methodological review of the literature

Paul Dolan[a,b,*], Rebecca Shaw[c], Aki Tsuchiya[d] and Alan Williams[e]
[a] *Sheffield Health Economics Group, School of Health and Related Research, University of Sheffield, UK*
[b] *Health Economics Research Programme, University of Oslo, Norway*
[c] *Department of Sociology, University of York, UK*
[d] *Sheffield Health Economics Group, University of Sheffield, UK*
[e] *Centre for Health Economics, University of York, UK*

Summary

In cost-utility analysis, the numbers of quality-adjusted life years (QALYs) gained are aggregated according to the sum-ranking (or QALY maximisation) rule. This requires that the social value from health improvements is a simple product of gains in quality of life, length of life and the number of persons treated. The results from a systematic review of the literature suggest that QALY maximisation is descriptively flawed. Rather than being linear in quality and length of life, it would seem that social value diminishes in marginal increments of both. And rather than being neutral to the characteristics of people other than their propensity to generate QALYs, the social value of a health improvement seems to be higher if the person has worse lifetime health prospects and higher if that person has dependents. In addition, there is a desire to reduce inequalities in health. However, there are some uncertainties surrounding the results, particularly in relation to what might be affecting the responses, and there is the need for more studies of the general public that attempt to highlight the relative importance of various key factors. Copyright

Introduction

Cost-utility analysis (CUA) seeks to provide health care policy-makers with information on the health benefits associated with alternative resource allocation decisions. Health benefits in CUA are measured in terms of the number of quality-adjusted life years (QALYs) gained. The QALY is a combination of the value of the health states and their duration, and every QALY is equivalent to one year of life in full health. In CUA, the numbers of QALYs gained are aggregated across individual patients according to the sum-ranking (or QALY maximisation) rule. This requires that the social value from health improvements is a simple product of gains in quality of life, length of life and the number of persons treated.

In this paper, we consider the results of a methodological review of the literature concerning the de facto standard in CUA that the sole objective of health care is to maximise the number of QALYs gained, irrespective of who those QALYs go to and how they are distributed across society. The objective of the review is to search for

*Correspondence to: Sheffield Health Economics Group, School of Health and Related Research, University of Sheffield, 30 Regent Street, Sheffield S1 4DA, UK. E-mail: p.dolan@sheffield.ac.uk

Received 8 March 2003
Accepted 23 January 2004

studies where different ways in which the descriptive validity of this de facto standard is examined. We then seek to draw some general patterns that emerge from the data and to suggest where future research efforts might be directed. In what follows, we describe the basis of QALY maximisation and present the questions addressed in the review. Next we describe the literature search and some descriptive data on the papers reviewed. Subsequent sections open with the theoretical literature on each item, and then report the evidence from the review. Finally, we discuss some of the problems with interpreting the results and the implications of the results for future empirical research.

QALY maximisation

In the simplest case, with no uncertainty and no changes in health over time, an individual's health gain from treatment, QALY_G, can be represented as

$$\mathrm{QALY}_G = T_1 Q_1 - T_0 Q_0 \tag{1a}$$

where T is the number of years, Q represents health state values, and the subscripts 1 and 0 represent health with and without treatment, respectively. Accounting for uncertainly, this becomes

$$\mathrm{QALY}_G = \Sigma_h \Sigma_t p_{1ht} Q_{ht} - \Sigma_h \Sigma_t p_{0ht} Q_{ht} \tag{1b}$$

where p_{1ht} and p_{0ht} represent the probabilities of an individual finding himself in health state h in time period t with and without treatment, respectively. Q_{ht} is the value of health state h at time t. The number of consecutive periods spent in any given state corresponds to the duration element, T in Equation (1a). So, this is the QALY *model* representing individual benefit from treatment. In order to compare interventions that are expected to benefit different numbers of people in different ways, this QALY model will usually be (although it does not have to be) incorporated into an algorithm that implies QALY *maximisation*. Assume there are individuals $i, j, \ldots$, each with a probability p_{hti} of being in state h at time t. Then the sum of p across the relevant population, $\Sigma_i p_{hti}$, will be equivalent to the expected number of people being in state h at time t; that is, n_{ht}. (This procedure allows for $p_{hti} \neq p_{htj}$, but requires the assumption $Q_{hti} = Q_{htj}$, where $j \neq i$.) The expected health gain from the intervention can then be expressed as

$$\text{Aggregated } \mathrm{QALY}_G = \Sigma_h \Sigma_t n_{1ht} Q_{ht} - \Sigma_h \Sigma_t n_{0ht} Q_{ht} \tag{2}$$

For this formula to accurately represent the social value associated with health care interventions all of the relevant parameters must have interval scale properties; that is, social value must be linear in n, as well as linear in Q and T. Linearity in n includes anonymity, since so long as there are n people in state h at time t, it does not matter who these people are. In other words, the algorithm does not distinguish between health gains to individuals i and j. Economists who allow for interpersonal comparisons of welfare typically aggregate benefits according to the sum-ranking rule, which, by implication, assumes linearity in all of the elements in Equation (2).

The question that lies at the heart of this paper is the extent to which people's preferences depart from the assumption of linearity inherent in Equation (2). Therefore, we conducted a literature review that sought to shed light on the following five questions:

1. Is social value a linear function of changes in quality and length of life?
2. Is social value independent of the age of the recipient?
3. Is social value independent of other characteristics?
4. Is social value unaffected by inequalities in health?
5. Is social value independent of how a fixed gain is distributed?

The literature review

The literature search was based on a 'citation pearl growing' method [1]. This is an alternative search strategy suitable for methodological reviews, where the more conventional keyword based search strategies may result in a very large number of irrelevant references. A set of core references need to be identified as the 'initial' references, and based on these the first wave of searches looks for papers where this core literature has been cited, the next wave looks for papers where those identified and included from the first

wave have been cited, and so on, until no further papers are identified. This was undertaken using the citation search facility of the Institute of Scientific Information (ISI) citation indexes and through reference list searching. These databases cover the science (including biomedical science), social sciences (including economics) and arts and humanities literature. The search was restricted to papers in the English language, dated 2001 or earlier.

First, 129 relevant 'initial' references were identified from a review of equity in economic evaluations [2] and from the authors' own collections. Three rounds of searches generated a total of 1739 additional references. Any paper that either allows inferences to be drawn about the relative weight that ought to be given to the health gain of one group or individual *vis-à-vis* another, or discuss these issues theoretically was eligible for inclusion in the review. In addition, given the methodological nature of the review, where a theoretical issue of general interest in health care was discussed in specialist medical journals for educational purposes these were usually excluded. Finally, when an author uses similar data to address the same question above, we have included only the main paper by that author. A total of 247 references were selected on the basis of the title and the abstract and 78 were included in the final review, 64 of which have empirical data.

Table 1 summarises the 64 studies in the review that report empirical data. It shows that the studies that have looked at QALY maximisation have mostly asked respondents to self-complete a questionnaire. Many studies have asked questions of a random sample of the general public, with samples often in excess of 100. However, there are still a few studies that have used relatively small convenience samples. The majority of studies have been carried out in the UK, followed by Europe and North America.

This is a methodological review and, in presenting the results, we have made no attempt to assess the quality of empirical studies. We will simply draw readers' attention to particular aspects of the study design (such as the sample size and composition and the country of origin) from which their own judgements about the quality and relevance of the data can be made but in discussing we will be a little more prescriptive.

Is social value a linear function of changes in quality and length of life?

Harris warns that QALY maximisation may lead to unacceptable discrimination against the elderly, the infirm, and other vulnerable groups in society with lower than average capacity to benefit from treatment [3]. To Harris, provided that the patient wishes to go on living, each life should be valued equally irrespective of how much is left. Whilst it is one thing to say that the size of the benefit is not the only thing that matters, it is another thing entirely to say, as Harris does, that it should not matter at all. People may quite reasonably be interested in differences in health without treatment as well as in the benefits from treatment. Two recent Norwegian commissions on priority setting in health care have identified that an important rationale for government involvement in health care is to provide benefit to those with the worst health prospects [4].

In general terms, and across a range of decision contexts, the empirical evidence currently available suggests that people are willing to sacrifice quality of life gains in order to give priority to the most severely ill [5–7]. Comparing improvements in health that start at different levels of severity but are equal in size, Nord reports a preference for movements starting at lower levels over equidistant improvements starting at higher levels [8]. Dolan asked respondents to trade off severity of the initial condition, with the size of health gain and found a move from 0.2 to 0.4 was equivalent to 0.4 to 0.8 [9]. Ubel replicated Nord's study and obtained similar results in that people wanted to give greater priority to those who were most severely ill [10]. However, when he provided clarification of the consequences of their choice, only 6% of respondents chose to allocate resources to the very ill patient. When the question was framed in terms of self-interest, the number rose to 12%.

In the context of liver transplantation, Ubel and Loewenstein show that only a small minority of respondents chose to give all the organs to the better-prognosis group [11]. However, the larger the prognostic differences between patients, the less likely respondents were to give all patients an equal chance of receiving the organs (also see [12,13]). Ubel and colleagues found that subjects placed equal importance on saving the lives of people with pre-existing paraplegia as compared to

Table 1. Empirical references

Author(s)	Year	Design	Sample	Sample size	Country of study
Abellan–Perpinan and Pinto-Prades	1999	2	3	3	3
Anand and Wailoo	2000	1	1	3	1
Andersson and Lyttkens	1999	2	3	3	3
Block *et al.*	2001	5	1	3	2
Bowling	1996	3	1	4	1
Browning and Thomas	2001	1	1	3	4
Busschbach *et al.*	1993	3	1, 3	2	3
Charny *et al.*	1989	1	1	3	1
Choudhry *et al.*	1997	2	4	3	1
Cookson and Dolan	1999	5	1	2	1
Cropper *et al.*	1994	6	1	4	2
Cuadras-Morato *et al.*	2001	2	3	4	3
Dolan	1998	2	3	2	1
Dolan and Cookson	2000	5	1	2	1
Dolan *et al.*	1999	5	1	3	1
Dolan and Green	1998	3	4	1	1
Dolan and Robinson	2001	2	3	3	1
Dolan and Tsuchiya	2002	3	1	3	1
Dolan *et al.*	2002	3	1	3	1
Edwards *et al.*	1999	1	1, 4	4	1
Emmelin *et al.*	1999	1	4	2	3
Furnham *et al.*	2000	5	3	3	1
Holmes	1997	2	3	3	2
Johannesson and Gerdtham	1996	4	3	4	3
Johanesson and Johansson	1996	6	1	4	3
Johanesson and Johansson	1997	6	1	4	3
Kneeshaw	1999	1	1	4	1
Kuder and Roeder	1995	5	1	2	3
Lewis and Charny	1989	1	1	3	1
Lindholm *et al.*	1996	1	4	2	3
Lindholm *et al.*	1997	1	4	2	3
Lindholm *et al.*	1998	1	4	2	3
Lindholm and Rosen	1998	1	4	3	3
Mooney *et al.*	1995	2	4	3	4
Neuberger *et al.*	1998	1, 3	1, 4	3	1
Nord	1993a	3	4	1	3
Nord	1993b	2	4	2	3
Nord	1995	2	1	1	3
Nord *et al.*	1995a	2	1	3	4
Nord *et al.*	1996	3	1	2	4
Olsen	1994	2	3	3	3
Olsen	2000	1	1	4	3
Ratcliffe	2000	2	4	3	1
Roberts *et al.*	1997	3	1	2	1
Rodriguez and Pinto	2000	4	3	2	3
Rodriguez and Pinto	2002	4	3	2	3
Shmueli	1999	3	1	4	5
Skitka and Tetlock	1992	2	3	3	2
Tsuchiya	2001	3	2	2	5
Tsuchiya *et al.*	2003	3	1	3	1
Ubel	1999	2	2	3	2
Ubel and Loewenstein	1995	2	2	3	2
Ubel and Loewenstein	1996a	2	2	3	2
Ubel and Loewenstein	1996b	2	2	3	2

Table 1 (*continued*)

Author(s)	Year	Design	Sample	Sample size	Country of study
Ubel *et al.*	1996a	2	2	3	2
Ubel *et al.*	1996b	1	2	4	2
Ubel *et al.*	1998	2	2	2	2
Ubel *et al.*	1999a	2	2	3	2
Ubel *et al.*	1999b	2	2	3	2
Ubel *et al.*	2000	2	2, 4	3	2
Ubel *et al.*	2001	1	4	4	2
Verkamp *et al.*	1998	7	4	3	3
Williams	1988	2	4	2	1
Zweibel *et al.*	1993	6	1	3	2

Note: Design: 1 = postal questionnaire; 2 = self-completion questionnaire; 3 = structured interview; 4 = experiment; 5 = focus group discussion; 6 = telephone survey; 7 = ethnography. Sample: 1 = general public (random/quota); 2 = general public (convenience); 3 = students or patients; 4 = health professionals or academic staff. Sample size: 1 = 1–29; 2 = 30–99; 3 = 100–999; 4 = more than 1000. Country of study: 1 = UK; 2 = US and Canada; 3 = Europe; 4 = Australia; 5 = Other.

those who could be returned to perfect health because they did not have pre-existing paraplegia [14]. But interestingly, the same subjects gave lower priority to patients who would experience the onset of paraplegia after having their lives saved. And respondents are not completely insensitive to the size of the health gain: for example, Abellan-Perpiñan and Pinto-Prades found that the smaller the size of the benefit to one patient, the more likely people were to maximise health gain [15].

In addition, the *final* health state is also important, particularly in the context of a patient who cannot be returned to full health after treatment [16,17]. Roberts and colleagues found little support for programmes that provided a prognostic improvement but left patients in relatively poorer states of health [18]. Dolan and Green found that many respondents favoured treating those patients with most to gain, seemingly on the grounds that the health state the worse-off would be in *after* treatment was not sufficiently good enough to warrant giving them priority [19]. In relation to life years, Dolan and Cookson found that people were willing to make health gain trade-offs between patient groups once the differences in the number of life year gained went beyond a certain threshold [20]. Shmueli asked respondents to choose between two people: one who was paralysed and who could be returned to full health for the rest of their life and one who would otherwise die but could be given one month in full health [21]. Sixty-three percent of respondents chose the first person. However, when the question was asked with the second person now living for five years with treatment, only 33% chose the first person. It appears, then, that people may also take account of a threshold level of health after treatment.

Is social value independent of the age of the recipient?

Tsuchiya distinguishes three different types of 'ageism', all of which suggest lower priority for older people [22]. The first favours the young over the old because they have longer life expectancies, the second favours young adults over children and the old because they are more productive, and the third favours the young over the old because the old have had more of a share of life years. The first of these is not relevant here, since the issue is valuing life years (as opposed to whole lives) at different ages. Regarding the second and third types of ageism, Kappel and Sandøe argue that the young should be prioritised for both productivity and 'fair innings' type reasons [23]. That is, other things being equal, we should favour the young; either because resources will generally be more useful when given to young people, or because they have lived less life and therefore 'deserve' the health improvement.

In developing the 'fair innings' argument, Williams suggests that the expected number of QALYs a person enjoys over a lifetime should be taken into account [24]. Williams argues that there is some amount of quality-adjusted

length of life that can be regarded as an ethical entitlement for everybody. Individuals receiving less than this amount 'have in some sense been cheated', whilst anyone getting more than this 'is living on borrowed time'. However, Harris (who was one of the first to use the term the 'fair innings argument' back in 1985) maintains that it is not possible to decide who has had a fair innings without a detailed life history [25]. But it would seem that there are strong ethical reasons for the age weighting of health benefits, and this is something that is currently being done by the World Health Organisation in its calculation of disability adjusted life years, or DALYs [26].

Three empirical studies support the idea of treating people of all ages equally. Respondents to studies by Anand and Wailoo [27], Kuder and Roeder [28] and Zweibel *et al*. [29] did not want to discriminate on the basis of age. However, most studies suggest that health gains to the old are weighted less. Respondents to a number of studies want to give lower priority to older people [30–36]. Browning and Thomas found that the age of a potential recipient of a donor organ was as important a consideration as their prognosis [37]. Respondents in a study by Cropper and colleagues viewed saving one 20 year old as equivalent to saving seven 60 year olds [38]. The study by Johannesson and Johannsson produced results suggesting that saving one 30 year old was equivalent to saving 35 individuals at the age of 70 years old [39]. Johannesson and Johannsson report undiscounted weights for life years gained at ages 30, 50 and 70 of 1.0; 0.22; and 0.1, respectively [40]. Lewis and Charny show that respondents have a very strong preference for 5 year olds over 70 year olds, a strong preference for 35 year olds over 60 year olds and a slight preference for 8 year olds over 2 year olds [41].

Nord and colleagues asked respondents a life-saving question and found 41.9% wanted to give patients equal priority; 40.5% wanted to give less priority to the very old; and 17.6% wanted to give more priority to the young [42]. They then asked the same respondents a quality of life question, and found 75.6% wanted to accord all patients equal priority; 21.5% wanted to give more to the young; and 2.9% more to the old. Nord and colleagues derived weights for ages 10, 20, 60 and 80 of 1.1; 1.0; 0.4 and 0.1 [43]. Busschbach and colleagues asked people aged around 20 and people aged around 70 to assign weights for ages 5, 10, 35, 60 and 70, and obtained implied age weights of 0.2, 1.5, 1.0, 0.7, 0.7, respectively [31]. Replicating this, Tsuchiya found that weights assigned by the two groups were different: for the young, these were 1.8; 1.6; 1.0; 0.5 and 0.6, respectively, and for the old, they were 0.6; 0.8; 1.0; 0.5 and 0.3 [35]. Varekamp and colleagues carried out an enthnographic study of *actual* decision-making [44]. This participant observation of clinicians showed age rationing to be occurring on the basis of capacity to benefit but also because of the 'stage of life' patients were felt to be at.

Is social value independent of other characteristics?

In addition to the assumptions about the linearity of the parameters in Equation (2), aggregation methods in CUA usually ignore (or, more accurately, are neutral towards) any other attributes that may affect the social value attached to health gains. These additional factors relate to the *causes* of the need for treatment and to the *wider consequences* from treatment. Society may wish to take account of the extent to which a person's ill health has been caused by factors beyond his control and the extent to which his ill health is considered as being determined by his chosen 'lifestyle'. LeGrand argues that 'if an individual's ill health results from factors beyond his or her control, then the situation is inequitable; if it results from factors within his or her control, then it is equitable' [45, p. 269]. Similarly, Dolan and Olsen offer theoretical support for taking 'responsibility' into account (but only in very limited circumstances) [46].

In relation to the consequences from treatment, Williams argues that the public would discriminate in favour of people with children over people without children [36]. Looking at benefits to wider society, a paper by Labelle and Hurley consider including externalities derived from interdependent utility functions [47]. They argue that the level of interdependent utility associated with a health intervention will vary depending on the individuals to whom it is targeted: a QALY will be worth more than 1.0 for those individuals for whom society values health improvements more highly. In addition, there is the possibility that people may

wish to compensate certain groups for disadvantages they may face in life, and a number of empirical papers have looked at whether benefits that go to lower socio-economic groups should be given greater weight.

There is empirical evidence that some people (but by no means a clear majority) wish to give less priority to those who are considered to be in some way responsible for their ill health [30,36,41,42,48]. In the Gallup Poll in August 1994, 41% of respondents wanted lower priority for those with self-inflicted illnesses. A survey by the Office of Population, Census and Surveys in May–June, 1995 [49] found similar results. Skitka and Tetlock found that respondents were more likely to deny aid to those responsible for their predicament [50]. Similarly, Neuberger and colleagues found respondents wanting to give lower priority to patients waiting for a liver transplant if this was caused by alcoholic liver disease [51]. Ubel and colleagues observe that even when drug users had a better transplant prognosis than any other patient, respondents only gave them 33% of available organs [52]. In the same study, smokers received 36% of the organs.

Ubel and colleagues note that in response to these questions respondents who had never smoked discriminated the most, and those who were current smokers discriminated the least [53]. Furnham and colleagues found that, in individual interviews followed by focus group discussions, respondents generally favoured non-smokers [54]. Dolan and Tsuchiya found that respondents wanted to give people who have not cared for their own health about half as much weight as those who have cared for their health, and respondents were generally unwilling to reduce inequalities in health that were due to smoking behaviour [55]. However, whilst Dolan and colleagues found that many people were in favour of discriminating against those whose ill health is considered to be partly self-inflicted, this view provoked considerable discussion and dissent [56] (see also [20]). And contrary to many of the results above, Edwards and colleagues found that a clear majority of people did not want to take 'responsibility' into account [57].

There is also evidence that people may not be neutral to the *non-health consequences* of treatment. For example, a person who has family or friends who are dependent upon them might be given priority over someone else that does not. There is evidence that people wish to discriminate in favour of those with dependents, particularly young children [32,36,37,51,56,58]. Edwards and colleagues present an exception to this general consensus, and found that respondents did not want to prioritise those with dependents [57]. In terms of gender, Dolan and colleagues presented respondents with a series of trade-off questions in relation to differences in life expectancy and found that, when the differences were between men and women, the median preference was not to target those with lower average life expectancy (i.e. men) [59]. Using racial groups, Ubel and colleagues found respondents less willing to maximise benefits because of the desire to positively discriminate minority racial groups [60].

A number of studies have looked at socio-economic status. Respondents to Anand and Wailoo's survey did *not* discriminate on the basis of social class (i.e. they did not wish to target the worse off) [27]. In the context of focus group discussions, respondents to Block and colleagues mentioned a range of benefits other than health gain, but equity was rarely mentioned and only 2% wanted to prioritise poorer patients [61]. Charny and colleagues report that 21.5% of respondents preferred a director to an unskilled worker; 28.5% an unskilled worker to a director; 34.5% a teacher to a lorry driver; and 15.2% a lorry driver to a teacher, but about 50% of people did not want to discriminate on the basis of job title [32]. Dolan and colleagues found that after discussion, 23% of respondents wanted lower priority for the rich and 8% wanted higher priority for those with lower educational attainment [56]. Holmes presented the same questions on mobility and speech dysfunction as before and reported values for mobility dysfunction of 0.57 for blue-collar workers and 0.66 for white-collar workers, and, for speech, values of 0.65 and 0.62, respectively [58]. In a study by Mooney and colleagues, 41.3% of respondents chose to prioritise the lower social class [33].

Is social value unaffected by inequalities in health?

Culyer and Wagstaff have argued that equality of health is one of the most important objectives of health care [62]. Economists have used the social welfare function (SWF) to consider the trade-off

between maximising health and equalising health [63]. The SWF accounts for *both* inequalities in health and the absolute health status of individuals, as the level of social welfare decreases with inequalities in health and increases with the health of individuals. In addition, social welfare depends on the degree of society's aversion to inequalities in health, and on the relative weights attached to the health of different groups or individuals. Johannesson proposes 'equity-weighted QALYs', where relative changes in QALYs receive the same weight for all individuals controlling for age and sex [64]. If the objective is to reduce inequalities in future health, it follows that QALY gains should be distributed initially to those who can expect the worst prospective health if they are left untreated. It might also be desirable to take into account the number of QALYs that a person has experienced up to the decision point, for instance, in terms of reducing inequalities in lifetime health (see the discussion of the 'fair-innings' argument above). Further to this, Dolan and Olsen put the concerns for severity and for age weighting in a more complete analytical framework that distinguishes between QALYs gained as a result of *past* health care and those gained 'free' of health care [46].

Lindholm and colleagues have measured people's preferences regarding inequalities in health by using differences in health that exist by socio-economic status. Lindholm and Rosen found that given a choice between two programmes (one which is more effective but does not eliminate inequality and the other which is less effective but eliminates inequality), Swedish politicians are prepared to sacrifice 15 of 100 preventable deaths to achieve equity in death rates between blue and white-collar workers [65]. Lindholm and colleagues report that 27 out of 48 respondents prefer the programme targeted at the blue-collar workers [66]. Emmelin and colleagues observe that 12.2% of respondents want to direct the programme at the socially disadvantaged [67]. However, Lindholm and colleagues observe that these respondents want to equalise outcomes up to a threshold [68] and another study reports a median willingness to sacrifice efficiency at 15% of health gains [69].

Dolan and colleagues presented respondents with differences in life expectancy by social class, and asked them to choose between maximising health and targeting benefits on the lowest social class (but with lower overall health gains as a result) [59]. The median response in this study was to be indifferent between a state of the world where people in the highest and lowest social classes live, on average, to be 80 and 75 and another state of the world where these groups live, on average, to be 78 and 75.5, respectively. In a second question, presenting differences in rates of limiting long-term illness between social classes one and five, the median respondent was indifferent between a decrease in the rate of long-term illness of 7% for both groups and a 2% and 8% reduction. The paper goes on to calculate the implied weights at the initial point, and reports a given gain in life expectancy to the lowest social class to be weighted (by the median respondent) about seven times as highly as an equivalent gain to the highest social class; and a given reduction in long-term illness to be weighted about four times as highly.

In general, the literature is indicative of a general preference for reducing health inequalities – but not at all costs. Andersson and Lyttkens observe that the degree of inequality matters – influenced by the cost in terms of reduced life expectancy for the more fortunate [70]. Johannesson and Gerdtham report a willingness to sacrifice one QALY to the better off to give 0.45 QALYs to the worse off [71]. Dolan and Robinson report inequality neutrality in one study and inequality aversion in the other, and attribute the difference to the effect of loss aversion (i.e. people are reluctant to take benefits away from one person to give them to someone else) [72]. In a study by Cuadras-Morato and colleagues the majority of respondents chose to allocate resources such that differences between final health states is the same for each patient [73].

Is social value independent of how a fixed gain is distributed?

The QALY maximisation model is indifferent between a large benefit to the few and a small benefit to the many, provided that the overall benefits are of equal magnitude. But people may have preferences in favour of one over the other. Olsen suggests that this preference will be affected by four considerations: positive time preference, threshold effects, risk aversion, and 'pure' equity preferences [74]. He notes that time preferences and equity preferences would take respondents in the direction of distributing benefits as widely as possible. If the smaller gain is below a certain

threshold, or people are risk averse, this may take them in the reverse direction i.e. they may prefer to concentrate the health gains to a few.

There is evidence that people are not indifferent to the distribution of the same total gain over populations of different sizes. Choudhry and colleagues found that one half of respondents choose to provide a large increase in life expectancy to the few and about one fifth preferred a small increase in life expectancy for the many [75]. Olsen also found that a majority of respondents were not indifferent between two distributions with the same total gain, although a much smaller percentage preferred to concentrate benefits [74]. He suggests that there is a threshold level of benefits to the larger group above which people prefer to distribute gains to as many people as possible but below which they prefer to concentrate gains (also see [76]). This is reiterated by Rodríguez Miguez and Pinto-Prades who found that participants, on average, prefer health programmes that distribute benefits over a greater number of people, provided that the gain to each patient is sufficiently high, and who report this threshold gain to be over nine years [77]. There is also evidence that people have a strong preference for giving everyone a benefit, if this is possible. In a choice between two screening programmes, Ubel and colleagues found that one half of respondents preferred the programme which could prevent fewer deaths but which could be offered to all citizens [78]. However, in a follow up study, Ubel and colleagues found that many fewer respondents held this preference when neither screening programme could be offered to everyone [79].

Discussion

This review has shown that an increasing number of studies are looking at the descriptive validity of the QALY maximisation rule. In relation to *the linear additivity of quality and length of life*, there seems to be a diminishing marginal social value associated with changes in both Q and T. However, very few of the studies have properly controlled for the possibility that people's preferences may have been contaminated by the belief that there is a diminishing marginal value to incremental health benefits. Some studies focusing on quality of life, like that of Nord [5] and Dolan [9], have explicitly told respondents that the more severely ill patients will gain less benefit from treatment than the less severely ill patients, but more could be done in future studies to get respondents to really appreciate the interval properties of health state valuation scales. For example, respondents could be asked to identify two treatments that they consider to bring about the same individual benefit but which start and finish at different points on the valuation range. Only then would they be asked to consider the social value of these treatments. This is somewhat similar to the design adopted by Dolan and Green [19]. In a similar way, studies that have looked at the linear additivity of life years might have been contaminated by decreasing marginal utility of life years, and possibly even by positive time preference. So, again, future studies should first seek to establish that the duration of individual benefits are considered identical before proceeding to the evaluation of social benefits.

In relation to *the value of life years at different ages*, there have been some studies that suggest that all ages should be treated equally but the majority of the empirical evidence is supportive of giving lesser weight to older people. Whilst perceived differences in productivity across age groups alone are unlikely to have explained the results in these studies, it is often difficult to tell how much of the preference for the young is due to the benefits to the young being greater (or being perceived to be greater) and how much is due to the young having lived for less time. The former explanation is consistent with the QALY maximisation rule while the latter is consistent with the 'fair innings' argument. One possible way to do control for health benefits might be to ask respondents to compare different ages when the benefits do differ, and then to ask them the same question when the benefits are the same; in this way, the difference between what they are being asked to do and what they might naturally do (i.e. answer in a way consistent with the first question) is made explicit (see [80]).

Regarding the *importance of other characteristics*, there is some evidence to support the view that people who are considered to be responsible for their ill health should be given lower priority, but this is never a majority view, and it is certainly an issue that generates much controversy. There is a greater consensus, insofar as there is majority support, for the idea that we should discriminate in favour of those with dependants. The evidence pertaining to peoples' preferences about *inequal-*

ities in health suggests that there is a clear preference for reducing inequalities in health when they are described according to socio-economic status, but no real desire to when they are described according to gender. Finally, in terms of the questions addressed in this review, there is evidence that people are not indifferent concerning *the distribution of a fixed benefit*. Generally, they prefer to disperse benefits as widely as possible but, if the benefits going to any one individual are considered to be too small, they prefer to concentrate benefits amongst fewer people instead. The possibility of threshold effects – where health benefits and/or the final level of health are considered to be too small to give priority to – seem to be pervasive amongst many of the studies reported here (e.g. also in relation to the linearity additivity of Q and T), and so it really is important that empirical research is directed at this issue.

Overall, many empirical studies have been undertaken to look at the societal value of health gains. Most of these studies have used relatively small samples and, to enhance the policy usefulness of the results, we would welcome more large-scale general population studies. However, given that people's preferences are partly constructed during the process of elicitation, such studies must also make some attempts to 'get behind the numbers' so that we can be much clearer than hitherto about what preferences are the result of concerns for equity and what are the result of extraneous factors. It has not been possible from this review to draw any real conclusions about the weight given to these considerations relative to one another. This is because respondents have typically been asked to weigh only one factor at a time against health gain and have rarely been asked to consider more than one factor simultaneously. Moreover, very few studies have actually tried to estimate weights for those factors that a simple QALY maximisation rule might be traded-off against. Therefore, our main recommendation would be for a coherent research programme that attempted to provide some very general conclusions about the relative importance of various key factors and then perhaps to elicit some general weights for them.

Acknowledgements

This review was funded by the NHS Executive R&D Programme, Health Technology Assessment Grant, on 'The validity of aggregation methods in cost-utility analysis'. We would like to thank Suzy Paisley for her assistance with the systematic review and Jan Abel Olsen for his very helpful comments on earlier drafts. The usual disclaimer applies.

References

1. Hartley RJ *et al*. *Online Searching*: *Principles and Practice*. BowkerSaur: Epping, UK, 1990.
2. Sassi F, Archard L, Le Grand J. Equity and the economic evaluation of healthcare. *Health Technol Assess* 2001; **5**(3).
3. Harris J. Unprincipled Qalys—a response to Cubbon. *J Med Ethics* 1991; **17**(4): 185–188.
4. Olsen J. Theories of justice and their implications for priority setting in health care. *J Health Econ* 1997; **17**: 625–640.
5. Nord E. The trade-off between severity of illness and treatment effect in cost-value analysis of health-care. *Health Policy* 1993; **24**(3): 227–238.
6. Ubel PA, Spranca MD, Dekay ML, Hershey JC, Asch DA. Public preferences for prevention versus cure: what if an ounce of prevention is worth only an ounce of cure? *Med Decision Making* 1998; **18**(2): 141–148.
7. Cookson R, Dolan P. Public views on health care rationing: a group discussion study. *Health Policy* 1999; **49**(1–2): 63–74.
8. Nord E. The person-trade-off approach to valuing health-care programs. *Med Decision Making* 1995; **15**(3): 201–208.
9. Dolan P. The measurement of individual utility and social welfare. *J Health Econ* 1998; **17**(1): 39–52.
10. Ubel PA. How stable are people's preferences for giving priority to severely ill patients? *Social Sci Med* 1999; **49**(7): 895–903.
11. Ubel PA, Loewenstein G. The efficacy and equity of retransplantation—an experimental survey of public attitudes. *Health Policy* 1995; **34**(2): 145–151.
12. Ubel PA, Loewenstein G. Distributing scarce livers: the moral reasoning of the general public. *Social Sci Med* 1996; **42**(7): 1049–1055.
13. Ubel PA, Loewenstein G. Public perceptions of the importance of prognosis in allocating transplantable livers to children. *Med Decision Making* 1996; **16**(3): 234–241.
14. Ubel PA, Richardson J, Pinto Prades JL. Life-saving treatments and disabilities—are all QALYs created equal? *Int J Technol Assess Health Care* 1999; **15**(4): 738–748.
15. Abellan-Perpinan JM, Pinto Prades JL. Health state after treatment: a reason for discrimination? *Health Econ* 1999; **8**(8): 701–707.
16. Nord E. The relevance of health state after treatment in prioritizing between different patients. *J Med Ethics* 1993; **19**(1): 37–42.

17. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Econ* 1999; **8**(1): 25–39.
18. Roberts T, Bryan S, Heinbotham C. Public involvement in health care priority setting: an economic perspective. *Health Expectations* 1997; **2**: 235–244.
19. Dolan P, Green C. Using the person trade-off approach to examine differences between individual and social values. *Health Econ* 1998; **7**(4): 307–312.
20. Dolan P, Cookson R. A qualitative study of the extent to which health gains matter when choosing between groups of patients. *Health Pol* 2000; **51**(1): 19–30.
21. Shmueli A. Survival vs. quality of life: a study of the Israeli public priorities in medical care. *Social Sci Med* 1999; **49**(3): 297–302.
22. Tsuchiya A. Age-related preferences and age weighting health benefits. *Social Sci Med* 1999; **48**(2): 267–276.
23. Kappel K, Sandøe P. QALYS, age and fairness. *Bioethics* 1992; **6**(4): 297–316.
24. Williams A. Intergenerational equity: an exploration of the 'fair innings' argument. *Health Econ* 1997; **6**(2): 117–132.
25. Harris J. Does justice require that we be ageist. *Bioethics* 1994; **8**(1): 74–83.
26. Murray C, Acharya A. Understanding DALYs. *J Health Econ* 1997; **16**: 703–730.
27. Anand P, Wailoo A. Utilities versus rights to publicly provided goods: arguments and evidence from health care rationing. *Economica* 2000; **67**(268): 543–577.
28. Kuder LB, Roeder PW. Attitudes toward age-based health-care rationing – a qualitative assessment. *J Aging Health* 1995; **7**(2): 301–327.
29. Zweibel NR, Cassel CK, Karrison T. Public attitudes about the use of chronological age as a criterion for allocating health care resources. *Gerontologist* 1993; **33**: 74–80.
30. Bowling A. Health care rationing: the public's debate. *Br Med J* 1996; **312**(7032): 670–674.
31. Busschbach JJV, Hessing DJ, Decharro FT. The utility of health at different stages in life – a quantitative approach. *Social Sci Med* 1993; **37**(2): 153–158.
32. Charny MC, Lewis PA, Farrow SC. Choosing who shall not be treated in the NHS. *Social Sci Med* 1989; **28**(12): 1331–1338.
33. Mooney G, Jan S, Wiseman V. Examining preferences for allocating health-care gains. *Health Care Anal* 1995; **3**(3): 261–265.
34. Rodriguez E, Pinto JL. The social value of health programmes: is age a relevant factor? *Health Econ* 2000; **9**(7): 611–621.
35. Tsuchiya, A. *The value of health at different ages*. Discussion Paper 184, York, Centre for Health Economics, University of York, 2001.
36. Williams A. Ethics and efficiency in the provision of health care. In *Philosophy and Medical Welfare*, Bell M, Mendus S (eds). Cambridge University Press: Cambridge, 1988.
37. Browning CJ, Thomas SA. Community values and preferences in transplantation organ allocation decisions. *Social Sci Med* 2001; **52**(6): 853–861.
38. Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs – how the public discounts time and age. *J Risk Uncertainty* 1994; **8**(3): 243–265.
39. Johannesson M, Johansson PO. The economics of ageing: on the attitude of Swedish people to the distribution of health care resources between the young and the old. *Health Policy* 1996; **37**(3): 153–161.
40. Johannesson M, Johansson PO. Is the valuation of a QALY gained independent of age? Some empirical evidence. *J Health Econ* 1997; **16**(5): 589–599.
41. Lewis PA, Charny M. Which of 2 individuals do you treat when only their ages are different and you can't treat both. *J Med Ethics* 1989; **15**(1): 28–32.
42. Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs. egalitarianism – an Australian survey of health issues. *Social Sci Med* 1995; **41**(10): 1429–1437.
43. Nord E, Street A, Richardson J, Kuhse H, Singer P. The significance of age and duration of effect in social evaluation of health care. *Health Care Anal* 1996; **4**(2): 103–111.
44. Varekamp I, Krol LJ, Danse JAC. Age rationing for renal transplantation? The role of age in decisions regarding scarce life extending medical resources. *Social Sci Med* 1998; **47**(1): 113–120.
45. LeGrand J. Equity, health and health care. *Social Just Res* 1987; **1**: 257–274.
46. Dolan P, Olsen JA. Equity in health: the importance of different health streams. *J Health Econ* 2001; **20**(5): 823–834.
47. Labelle RJ, Hurley JE. Implications of basing health-care resource allocations on cost-utility analysis in the presence of externalities. *J Health Econ* 1992; **11**(3): 259–277.
48. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. *Health Econ* 2000; **9**(2): 137–148.
49. Kneeshaw J. What does the public think about rationing?: a review of the evidence. In *Rationing*: *Talking and Action in Health Care*, New B (ed.). BMJ: London, 1999.
50. Skitka LJ, Tetlock PE. Allocating scarce resources – a contingency-model of distributive justice. *J Exper Social Psychol* 1992; **28**(6): 491–522.
51. Neuberger J, Adams D, MacMaster P, Maidment A, Speed M. Assessing priorities for allocation of donor liver grafts: survey of public and clinicians. *Bri Med J* 1998; **317**(7152): 172–175.

52. Ubel PA, Baron J, Asch DA. Social responsibility, personal responsibility, and prognosis in public judgments about transplant allocation. *Bioethics* 1999; **13**(1): 57–68.
53. Ubel PA, Jepson C, Baron J, Mohr T, McMorrow S, Asch DA. Allocation of transplantable organs: do people want to punish patients for causing their illness? *Liver Transplant* 2001; **7**(7): 600–607.
54. Furnham A, Simmons K, McClelland A. Decisions concerning the allocation of scarce medical resources. *J Social Behav Personality* 2000; **15**(2): 185–200.
55. Dolan P, Tsuchiya A. *Individual Choice and the Social Welfare Function*: *Some Theoretical Issues and Empirical Evidence from Health*. Paper Series 2003. Ref 03/1. The University of Sheffield, Sheffield.
56. Dolan P, Cookson R, Ferguson B. Effect of discussion and deliberation on the public's views of priority setting in health care: focus group study. *Br Med J* 1999; **318**(7188): 916–919.
57. Edwards RT, Boland A, Wilkinson C, Cohen D, Williams J. Choosing explicit criteria for the prioritisation of elective NHS waiting lists: survey evidence of clinical and lay preferences from Wales. In *Proceedings of the Health Economists' Study Group Meeting*, Birmingham, 1999.
58. Holmes AM. A method to elicit utilities for interpersonal comparisons. *Med Decision Making* 1997; **17**(1): 10–20.
59. Dolan P, Tsuchiya A, Smith P, Shaw R, Williams A. *Determining the parameters in a social welfare function using stated preference data*: *an application to health*. Sheffield Health Economics Group Discussion Paper Series. 2002. Ref 02/2. The University of Sheffield: Sheffield.
60. Ubel PA, Dekay M, Baron J, Asch DA. Public preferences for efficiency and racial equity in kidney transplant allocation decisions. *Transplant Proc* 1996; **28**(5): 2997–3002.
61. Block MAG, Sandiford P, Ruiz JA, Rovira J. Beyond health gain: the range of health system benefits expressed by social groups in Mexico and Central America. *Social Sci Med* 2001; **52**(10): 1537–1550.
62. Culyer AJ, Wagstaff A. Equity and equality in health and health-care. *J Health Econ* 1993; **12**(4): 431–457.
63. Wagstaff A. QALYs and the equity-efficiency trade-off. In *Cost–Benefit Analysis*, Layard A, Glaister S (eds). Cambridge University Press: Cambridge (reprinted from *J Health Econ* 1991; **10**: 21–41, with corrections).
64. Johannesson M. Should we aggregate relative or absolute changes in QALYs? *Health Econ* 2001; **10**(7): 573–577.
65. Lindholm L, Rosen M. On the measurement of the nation's equity adjusted health. *Health Econ* 1998; **7**(7): 621–628.
66. Lindholm L, Rosen M, Emmelin M. An epidemiological approach towards measuring the trade-off between equity and efficiency in health policy. *Health Policy* 1996; **35**(3): 205–216.
67. Emmelin MAC, Lindholm LA, Stenlund HC, Dahlgren LG. Pol-ethical considerations in public health – the views of Swedish health care politicians. *Eur J Public Health* 1999; **9**(2): 124–130.
68. Lindholm LA, Emmelin MA, Rosen ME. Health maximization rejected—the view of Swedish politicians. *Eur J Public Health* 1997; **7**(4): 405–410.
69. Lindholm L, Rosen M, Emmelin M. How many lives is equity worth? A proposal for equity adjusted years of life saved. *J Epidemiol Community Health* 1998; **52**(12): 808–811.
70. Andersson F, Lyttkens CH. Preferences for equity in health behind a veil of ignorance. *Health Econ* 1999; **8**(5): 369–378.
71. Johannesson M, Gerdtham UG. A note on the estimation of the equity-efficiency trade-off for QALYs. *J Health Econ* 1996; **15**: 359–368.
72. Dolan P, Robinson A. The measurement of preferences over the distribution of benefits: the importance of the reference point. *Eur Econ Rev* 2001; **45**(9): 1697–1709.
73. Cuadras-Morato X, Pinto-Prades JL, Abellan-Perpinan JM. Equity considerations in health care: the relevance of claims. *Health Econ* 2001; **10**(3): 187–205.
74. Olsen JA. A note on eliciting distributive preferences for health. *J Health Econ* 2000; **19**(4): 541–550.
75. Choudhry N, Slaughter P, Sykora K, Naylor D. Distributional dilemmas in health policy: large benefits for a few or smaller benefits for many? *J Health Serv Res Policy* 1997; **2**: 212–216.
76. Olsen JA. Persons vs years—2 ways of eliciting implicit weights. *Health Econ* 1994; **3**(1): 39–46.
77. Rodriguez-Miguez E, Pinto-Prades JL. Measuring the social importance of concentration or dispersion of individual health benefits. *Health Econ* 2002; **11**: 43–53.
78. Ubel PA, Dekay ML, Baron J, Asch DA. Cost-effectiveness analysis in a setting of budget constraints—is it equitable? *New Engl J Med* 1996; **334**(18): 1174–1177.
79. Ubel PA, Baron J, Nash B, Asch DA. Are preferences for equity over efficiency in health care allocation 'all or nothing'? *Med Care* 2000; **38**(4): 366–373.
80. Tsuchiya A, Dolan P, Shaw R. Measuring people's preferences regarding ageism in health: some methodological issues and some fresh evidence. *Social Sci Med* 2003; **57**(4): 687–696.