Sheffield Economic Research Paper Series

SERP Number: 2009010

ISSN 1749-8368

Richard Edlin, Aki Tsuchyia* and Paul Dolan

Measuring the societal value of lifetime health

May 2009

Department of Economics
University of Sheffield
9 Mappin Street
Sheffield
S1 4DT
United Kingdom
www.shef.ac.uk/economics

Abstract:

This paper considers two societal concerns in addition to health maximisation: first, concerns for the societal value of lifetime health *for* an individual; and second, concerns for the value of lifetime health *across* individuals. Health-related social welfare functions (HRSWFs) have addressed only the second concern. We propose a model that expresses the former in a metric – the adult healthy-year equivalent (AHYE) – that can be incorporated into standard HRSWFs. An empirical study based on this formulation shows that both factors matter: health losses in childhood are weighted more heavily than losses in adulthood and respondents wish to reduce inequalities in AHYEs.

Key words: D63

JEL: Health; Social welfare function; Equity

Acknowledgments:
This paper is an output of the project "The relative societal value of health gains to different beneficiaries", funded by the NCCRM and NICE (RM03/JH11). We wish to acknowledge the other members of the project team (Chris Armitage, Amy Hukin, John Brazier, Rachel Ibbottson, Stirling Bryan, Dick Eiser and Jan Abel Olsen). We are grateful to the Centre for Research Evaluation at Sheffield Hallam University who collected the data, and all the study participants. The usual disclaimer applies.

1. Aims and Background:

Publicly funded health care systems aim to maximise population health (e.g. as measured in terms of Quality Adjusted Life Years, QALYs), and to account for inequalities in health. An analytical tool that can be used to balance the possibly conflicting objectives is the health related social welfare function (HRSWF), defined as a function of population health. Dolan et al (2005) suggest that the public are sensitive to issues beyond the expected health benefit of an intervention when asked about how society should allocate healthcare resources. In particular, the general public takes account of: age or the timing of illness within lifetime health (Busschbach et al, 1993; Mooney et al, 1995; Rodriguez et al, 2000); the severity of ill health and potential for health benefits (Nord, 1993, 1995, 1999; Nord et al, 1999; Ubel, 1999); individual characteristics and life circumstances, including individual responsibility (Charny et al, 1989; Dolan et al, 1999; Ratcliffe, 2000; Ubel et al, 2001; Dolan and Tsuchiya, 2009); and differences in the lifetime health of groups (Johannesson and Gerdtham, 1996; Lindholm et al, 1996; Lindholm and Måns Rosén, 1998; Andersson and Lyttkens, 1999).

These factors can be grouped into two different kinds. Factors of the first kind affect the societal value of lifetime health *within* a person. Two lifetime health profiles made up with the same amount of total QALYs may be valued differently by society if the distribution of ill health across the lifetime is not the same: for example, a QALY loss resulting from severe ill health during childhood may be regarded differently from the same QALY loss during adulthood; or permanent mild disability may be regarded differently from a short-lasting severe condition with the same QALY loss. In other words, these factors affect the levels of lifetime health that appear as arguments in the HRSWF. Factors of the second kind affect how society aggregates lifetime health *across* people. So, the same amount of health across people may add up to different levels of social welfare depending on how inequality averse is the HRSWF.

In this paper, we concentrate on two factors (quality of life and the timing) that appear within an "adult healthy year equivalent" (AHYE) and one factor (inequality aversion) that appears as standard SWF parameter. The AHYE captures the societal value of a health profile in terms of an equivalent number of years in full health to an adult. AHYEs are a more general form of QALYs: only when every quality of life 'time slice' is weighed equally will AHYEs reduce to QALYs.

Suppose there are two interventions: Intervention 1 gives a 10 year old 1 QALY over 5 years, and Intervention 2 gives an 80 year old 2 QALYs over 5 years.

Both are expected to live for another 5 years with or without treatment and die. Since (as a matter of positive description) Intervention 2 generates twice as much health measured in QALYs, standard QALY analysis will say Intervention 2 is twice as effective as Intervention 1. There are two separate things that can be considered. First, assume now that the societal value of 1 QALY of health varies by age of the patient such that a QALY to an adolescent is twice as valuable as a QALY to an adult. This would mean that a QALY to a teenager is equivalent to 2 AHYEs. If so, then the two interventions will be equally effective in terms of AHYEs. Second, assume instead that there is inequality aversion over lifetime health. To the extent that the 10 year old patient is expected to have a lower level of lifetime health compared to the 80-year old, the 1 QALY gained by the former may be valued more highly than the 2 QALYs gained by the latter. In reality, either or both of these considerations may hold. At the conceptual level, these two are independent of each other, but at the empirical level, it may not be straightforward to distinguish between them.

2. The Adult Healthy Year Equivalent and the SWF

Whereas the QALY is measured as a sum of health over time, the AHYE is derived as a sum of *weighted* health over time. That is, the number of AHYEs to a group *i*, or v_i, is given by:

$$v_i = \sum_{t=1}^{\infty} V(h_{it}, t),$$

where V is a weighting function based on health-related quality of life (h_{it}) of group *i* at time *t* and timing (t), where V increases in health ($\frac{dV}{dh_{it}} > 0$). Here, health states worse than dead can be considered so long as each group receives positive lifetime health ($v_i > 0$). We further assume that v_i is multiplicatively separable into health and timing components so that:

$$v_i = \sum_{t=1}^{\infty} f(h_{it}) T(t), \qquad \frac{df}{dh_{it}} > 0, \; T(t) > 0.$$

This equation allows us to asses the lifetime health of individuals in terms of AHYEs if we can estimate a relationship for *f* and *T*. In the study reported here, dichotomous variables were used to represent both relationships. Given such a relationship we could compare the lifetime health of individuals, and hence consider the trade-offs that society is willing to make between individuals in a lifetime health-based SWF.

The AHYE used here employs dichotomous variables for timing (representing health experienced up to or after 18 years of age), and severity/quality of life at 0 (equivalent to dead), 0.25 or 1.00 (full health). AHYEs are defined as weighted sum over five variables, and for simplicity these are combined into a vector $\mathbf{y}_i$ = $(y_{b,1.00,i}, y_{b,0.25,i}, y_{b,0,i}, y_{B,1.00,i}, y_{B,0.25,i})'$, where $y_{b,H,i}$ is the time spent in health state H as a child by group i, and $y_{B,H,i}$ the time spent in health state H as an adult by group i. The weights on these parameters for state H are given by b_H and B_H during childhood and adulthood, respectively. For comparability with the QALY, the adult dead and full health states are given are weighted at 0 and 1 respectively ($B_{1.00} = 1$ and $B_0 = 0$). In general the number of AHYEs can be written as:

$$v_i(\mathbf{y}_i) = b_{1.00} y_{b,1.00,i} + b_{0.25} y_{b,0.25,i} + b_0 y_{b,0,i} + y_{B,1.00,i} + B_{0.25} y_{B,0.25,i}$$

(Eq. 1)

Where multiplicative separability holds, the number of AHYEs becomes:

$$v_i(\mathbf{y}_i) = b_{1.00} y_{b,1.00,i} + (b_{1.00} B_{0.25}) y_{b,0.25,i} + y_{B,1.00,i} + B_{0.25} y_{B,0.25,i}$$

(Eq. 2)

In order to model the sensitivity in social preferences to lifetime health differences, we assume that those preferences can be represented by a constant elasticity of substitution SWF over AHYE-defined lifetime health. This general type of SWF has been used before but with life years (Williams, 1997; Dolan et al, 2002), healthy life years (Lindholm and Måns Rosén, 1998) or QALYs (Williams, 1997) as its metric. In a two-group case where each group is equally sized and made up of homogeneous individuals (in health, at least), the SWF is:

$$W = \left[0.5 v_1^{-r} + 0.5 v_2^{-r}\right]^{-\frac{1}{r}}, \qquad r \in [-1, \infty) \setminus 0$$

Where W is social welfare and r reflects the overall strength of inequality aversion.

Here, the inequality aversion parameter (r) reflects the degree to which society values a more equal distribution of health vis-à-vis maximising health. As citizens, people might be concerned only about average lifetime health ($r = -1$), only concerned about the health of the worst off group ($r \to \infty$) or, more realistically, concerned about both average health and inequalities ($r > -1$). Since the SWF is homothetic, the size of r relates to sensitivity to the relative differences in lifetime health. The SWF found here seeks to estimate both the AHYE and inequality aversion parameters simultaneously, and does so by finding a best fit solution in which multiplicative separability holds.

3. Framework of analysis

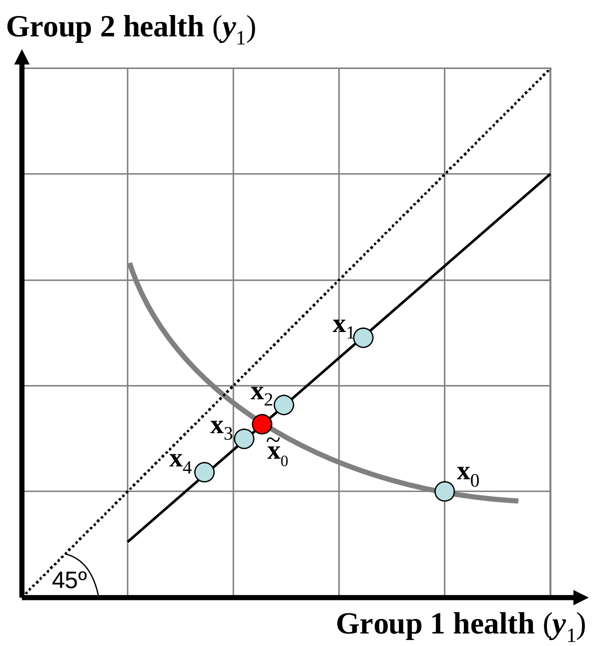
In brief, our method: 1) uses pairwise social preferences over states of the world representing the lifetime health for two homogeneous groups; 2) analyses these preferences to find states with equal social welfare using Thurstone scores; 3) takes points with equal social welfare to allow us to fit parameters to a non-linear SWF to find a best-fit model satisfying the multiplicative separability assumption.; 4) parameterises uncertainty using bootstrapping. We do not claim that our methods are the only possible of resolving parameterising the SWB but our aim was to establish the tractability of this issue and provide results for a given, and actually quite generalisable, set of methods.

3.1. Identifying points with equal social welfare

The SWF for each state of the world (**x**) is defined using the number of AHYEs each group receive, which are themselves an as-yet-undefined function of the lifetime health profiles of two equally sized and homogenous groups; for simplicity, we denote this as $\mathbf{x} = \{\mathbf{y}_1; \mathbf{y}_2\}$. (Whilst only two groups are used in this study, there is no restriction on the number of groups that can be incorporated.) “Choice sets” are defined by comparing pairs of states (sets of lifetime health profiles) against each other, with each choice set having one consistent “study state” $\mathbf{x}_0$ against a series of four “reference states” $\mathbf{x}_1$ to $\mathbf{x}_4$.

The reference states are constructed to be collinear (see Figure 1) with higher index numbers referring to worse states (e.g. $\mathbf{x}_1$ f $\mathbf{x}_2$ f $\mathbf{x}_3$ f $\mathbf{x}_4$) so long as all years spent in states better than dead are valued positively (or not too negatively) at a societal level. We use this ordering to infer preferences between the reference states and reduce the burden on respondents relative to the case in which all states would be compared to each other.

Figure 1: Reference and study states.



Our method is based on Case V Thurstone scores (Thurstone, 1927a, 1927b) and provides cardinal scale values for the social welfare of each state at the aggregate level across respondents. This method uses comparisons of all states against all other states, and where extreme preferences are found for one state over another, these values are typically censored to fall within a set range, since unanimity will cause indeterminacy. The typical response to this issue is either to omit such data or restrict proportions to fall within a permitted range (Guildord, 1954; Edwards, 1957) and we use a range of 2%-98% in our main analysis.

We use data for the all possible comparisons of the five states; thus 25 comparisons in total. 16 preferences are inferred using the relationships between the reference states: "better" states receive $r(\mathbf{x}_i, \mathbf{x}_j) = 0.98$, "worse" states receive $r(\mathbf{x}_i, \mathbf{x}_j) = 0.02$, and $r(\mathbf{x}_i, \mathbf{x}_i) = 0.50$ where a state is compared against itself. Of the remaining nine comparisons, we have four comparisons of the study state against reference states, another four for the reference states against the study state (since $r(\mathbf{x}_i, \mathbf{x}_j) = 1 - r(\mathbf{x}_j, \mathbf{x}_i)$), and one comparison of the study state against itself. Therefore, for each choice set we only need to elicit the four pairwise choices of the study state against the reference states.

The Thurstone scores assume an underlying random utility model in which $\mathbf{x}_i$ is preferred to $\mathbf{x}_j$ where the assessment of social welfare for the former exceeds that for the latter ($W(\mathbf{x}_i) > W(\mathbf{x}_j)$). The scores assume that each of these social welfare assessments is independently drawn from normal distributions with a common variance. The larger is the mean social welfare assessment for a state ($\overline{W_i}$), the more often we would expect $\mathbf{x}_i$ to be preferred against the other states. Let $r(\mathbf{x}_i, \mathbf{x}_j)$ be the average proportion of the time that $\mathbf{x}_i$ is preferred to $\mathbf{x}_j$ (indifference scored as a "half" preference to each option):

$$r(\mathbf{x}_i, \mathbf{x}_j) = \int_{-z(\mathbf{x}_i, \mathbf{x}_j)}^{\infty} \frac{1}{\sqrt{2p}} e^{-\frac{t^2}{2}} dt$$

Here, each $z(\mathbf{x}_i, \mathbf{x}_j)$ value is a function of the difference in social welfare between the states and the average of all the $z(\mathbf{x}_i, \mathbf{x}_j)$ values for $\mathbf{x}_i$ gives a cardinally measurable estimate of $\overline{W_i}$. With estimates for the study state and each of the reference states, we know where the study state falls in an ordered list of all the reference states. Further, since the social welfare estimates are cardinally measurable, we can estimate a "new" state ($\tilde{\mathbf{x}}_0$) with the same social welfare value ($\overline{W_0}$) as the study state:

$$\tilde{\mathbf{x}}_0 = \mathbf{x}_3 + \frac{\overline{W_0} - \overline{W_3}}{\overline{W_2} - \overline{W_3}} \mathbf{x}_2$$

3.2. Finding the SWF parameters

The process of deriving parameter estimates in the SWF uses these societally-equivalent states ($\mathbf{x}_0, \tilde{\mathbf{x}}_0$). These equivalent states should have the same estimated social welfare. Given a set of candidate parameters for the SWF, we interpret any difference in estimated social welfares between states as an error, and the sum of squared errors as an estimate of goodness of fit. By varying the SWF parameters we seek to minimise these errors and find a best fit.

As above, an AHYE-based SWF allows social preferences to be interpreted in different ways. In the introduction, we gave the example in which a QALY gain to a 10 year old was deemed to be twice as valuable as a QALY gain to an 80 year old. Here, if the QALY and AHYE gains coincide ($b_{1.00}$=1, $B_{0.25}$=0.25), this general preference suggests some weighting to reducing inequalities ($\hat{r} > -1$). However, if childhood health was deemed to be twice as valuable as adult health ($b_{1.00}$=2, $B_{0.25}$=0.25), we could explain this preference with a sum-ranking function ($r = -1$). For intermediate cases (1 <$b_{1.00}$<2), we would expect some weight on reducing inequalities but not as much as in the case where no differential timing of ill health occurs ($\hat{r} > r > -1$).

The relationships between parameters cause a specific computational problem. Suppose we have start with the r that would best represent inequality aversion in a QALY-based SWF (i.e. the $b_{1.00}$=1, $B_{0.25}$=0.25 case). Increasing r further suggests that society is more inequality averse, and this will be consistent with the data only if there is less relative difference between the scenarios when using AHYEs than when using QALYs. With less inequality in SWF terms, there is less potential for differences between the social welfare estimates of indifferent states, and hence lower total sums of squared errors for a best fit solution (conditioned on r) as inequality aversion increases. Minimising the sum of square errors across all parameters (including r) does not lead to convergence. Instead, at each stage we estimate best fit solutions for the AHYE parameters ($b_{1.00}, b_{0.25}$, and $B_{0.25}$) conditional on r without requiring multiplicative separability (using Equation 1). (b_0 is not estimated as all individuals survive into adulthood in our questions – given multiplicative separability is imposed, $b_0 = b_{1.00} \times B_0$ =0.)

Formally, we find best fit estimates for parameter values conditioned on inequality aversion and estimate the deviation from multiplicative separability within these estimates ($b_{1.00} \times B_{0.25} - b_{0.25}$). By varying inequality aversion, we can find the best fit solution satisfying multiplicative separability. This provides estimates for both parameter values and inequality aversion parameters – and hence a version of the SWF that separates out the societal measurement of lifetime health from preferences for equality across these societal measurements.

4. Methods

The overall research project assessed the relative health gains to different beneficiaries using a SWF framework slightly more general than the one used here. The project considered both other types of weighting and additional levels that are not reported here given space constraints and our stated aim of finding feasibility. The main project report is available, as is a general summary of the SWF findings (Dolan et al, 2008a, 2008b).

Recruitment:

Interviewing was conducted by the Centre for Research and Evaluation at Sheffield Hallam University within England and Wales. Seventeen areas were isolated and sampled to obtain a mixture of gender, age and education. Interviewers visited these areas and knocked on doors to obtain participants who had been informed previously by letter about the study. The sampling frame aimed to recruit a minimum number of people in categories defined by gender, age, and education.

Materials:

The main study carried out face-to-face interviews that took place at the respondent's home. This comprised a self-completion "beliefs" questionnaire (not reported here); a brief introduction to NICE, the need for decision making and priority setting, and the concept of quality of life; followed by the main pairwise choice task. The interview elicited responses for 64 questions in 16 choice sets, of which four choice sets are reported here.

The early choice sets in the project dealt with choices in which all life years were lived in 100% health until death. By presenting respondents with questions that did not involve illness, the aim was to provide respondents some general familiarity with the task and question format. Following 16 choices (four sets) of this type, the questions considered here were introduced. In these questions, individuals live either in poor (25%) health or full (100%) health.

Table 1 gives the reference states for the four choice sets considered here. These reference states in these questions are collinear; in addition, there is the same difference in AHYE terms between $\mathbf{x}_1$ and $\mathbf{x}_2$ as between $\mathbf{x}_2$ and $\mathbf{x}_4$, and the same difference in AHYE terms between $\mathbf{x}_2$ and $\mathbf{x}_3$ as between $\mathbf{x}_3$ and $\mathbf{x}_4$. The study states that are compared against the reference states appear in the lower half of the table. In all four choice sets, the study states involve more inequality in health (in QALY terms) than in the reference states.

Table 1: Reference States ($\mathbf{x}_1$ to $\mathbf{x}_4$) and Study States ($\mathbf{x}_0$)

Reference States	Group 1 Health	Group 2 Health
$\mathbf{x}_1$	62 years full health 16 years poor health	60 years full health 8 years poor health
$\mathbf{x}_2$	60 years full health 8 years poor health	56 years full health 8 years poor health
$\mathbf{x}_3$	59 years full health 4 years poor health	54 years full health 8 years poor health
$\mathbf{x}_4$	58 years full health	52 years full health 8 years poor health

Study States ($\mathbf{x}_0$)	Group 1 Health	Group 2 Health
Choice Set 1	66 years full health 8 years poor health	50 years full health 16 years poor health
Choice Set 2	66 years full health 16 years poor health	4 years poor health 54 years full health 4 years poor health
Choice Set 3	72 years full health 16 years poor health	48 years full health 16 years poor health
Choice Set 4	8 years poor health 72 years full health 8 years poor health	48 years full health 16 years poor health

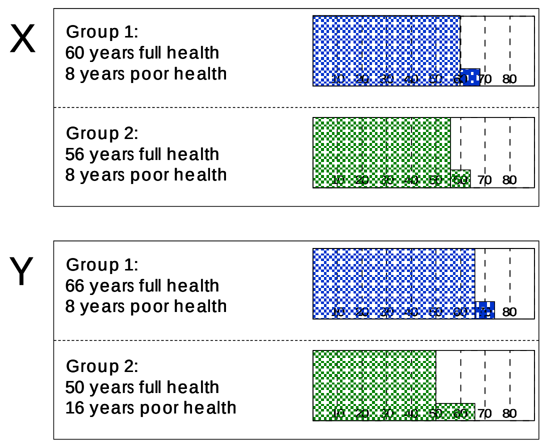
In the description of the health states, poor health was identified as a 0.250 (or 25%) quality of life. Four years in health of 0.250 was explained to the respondent as being worth 1 year in health of 1.000 health to the individual experiencing it, so that the 25% applies as an individual (rather than a societal) judgement. These questions are used to find weights given for childhood versus adult health, 25% health versus 100% health, and a value of *r*.

The choices were presented in a mixed text/graphics format, with the information also read out to respondents. Respondents were initially told that the consequences of health decisions could affect the distribution of health and at each choice were asked “Which scenario would you prefer NICE to bring about?” Responses could indicate a preference for either option or indifference (“I don’t mind if it’s X or Y”).

Analysis:

In Figure 2, the reference state $\mathbf{x}_2$ (labelled "X") is compared against the study state $\mathbf{x}_0$ (labelled "Y"). In Scenario X, Group 1 receives 60 years in full health followed by 8 years in poor (25%) health, whilst Group 2 receives 56 years in full health followed by 8 years in poor (25%) health. This is compared to Scenario Y in which Group 1 receives six more years in full health, and where Group 2 receives 6 years less in full health and 8 more years in poor health. In QALY terms, Group 1 obtains 6 more QALYs and Group 2 obtains 4 fewer QALYs by choosing Y over X. (In AHYE terms, this trade-off may vary but will still probably favour Y on efficiency grounds and X on equality grounds.)

Figure 2: Sample Choice - reference state $\mathbf{x}_2$ against $\mathbf{x}_0$ in Choice Set 1



In order to identify trade-offs, Thurstone-equivalent states ($\tilde{\mathbf{x}}_0$) were found for each of the four choice sets and parameters for AHYEs and inequality aversion (*r*) derived using Microsoft Excel. The trade-offs between the states were computed in both AHYE and QALY terms.

As these methods above provides only point estimates, we use bootstrapping (sampling with replacement) to analyse uncertainties. Assuming that the data are representative of the underlying uncertainty, bootstrapping allows the construction of additional samples of the same size as the original sample. By re-running the analysis on these samples it is possible to estimate a distribution for each parameter. This distribution allows us to gauge the extent of uncertainty in the point estimate of each parameter. In this way, we can address the non-methodological uncertainties of our estimates. As the number of values necessary to provide a convergent estimate for uncertainty is unknown, the analysis here uses n=5000 bootstrapped observations in order to find convergence.

5. Results

Within the main study, 582 interviews were conducted by nine interviewers, of which 559 had complete data across all 16 choice sets. It was not possible to compute a response rate, since the number of doors knocked on and the number of individuals asked to participate was not recorded. Whilst we selected a range of geographical locations, our emphasis was to obtain a sample across the range of gender, age, and education and not necessarily to aim for a representative sample of the general population.

Table 2 presents the background of our sample, and it is clear that through our sampling frame we over-sampled the retired and those aged over 60, and under-sampled those with an education to a high school level only relative to the 2001 census. The sample also contains a smaller number of individuals in employment and in ethnic minorities. The interview took an average of 55 minutes to complete.

Table 2: Background of the sample

Sample size		Social QALY sample (%)	2001 Census (%)
Gender:	Female	55	52
Age:	40-59	32	33
	60+	32	27
Ethnicity	White	95	92[a]
Employment status:	Self-employed	7	8
	Other Employed	39	52
	Retired	29	14
Education:	School only	47	78[a]
	HE/FE	53	22[a]
House ownership:	Owned/mortgage	71	71b[c]

[a] Ages 16-74 only

[b] 2000 data. Office of National Statistics.

Table 3 shows the proportion of the 559 respondents with complete data that prefer the study state ($\mathbf{x}_0$) to the reference states ($\mathbf{x}_1$ to $\mathbf{x}_4$). As expected, the study state is preferred more often as we move from the best reference state ($\mathbf{x}_1$) to the worst reference state ($\mathbf{x}_4$). There appears to be a relatively large spread of individual preferences, however, with 35% of preferences preferring the study state to the best reference state, and nearly 50% preferring the study state to the worst reference state.

Table 3: Preference for study states over reference states ($r(\mathbf{x}_0, \mathbf{x}_{l+1})$)

	$\mathbf{x}_0$ preferred to $\mathbf{x}_1$	$\mathbf{x}_0$ preferred to $\mathbf{x}_2$	$\mathbf{x}_0$ preferred to $\mathbf{x}_3$	$\mathbf{x}_0$ preferred to $\mathbf{x}_4$
Choice Set 1	37%	38%	42%	48%
Choice Set 2	43%	46%	47%	52%
Choice Set 3	36%	38%	43%	47%
Choice Set 4	39%	43%	46%	50%

Cells show $r(\mathbf{x}_0, \mathbf{x}_i)$.

Individual indifference is scored as a 50% preference.

Table 4 shows both the study states ($\mathbf{x}_0$) and their estimated equivalent states ($\tilde{\mathbf{x}}_0$) derived from the Thurstone scores in each choice set. Within each of the equivalent states, ill-health always occurs at the end of life so, for instance, the equivalent state in Choice Set 1 involves 59.23 years of full health, followed by 4.92 years in severe health. The tables show the difference between the study and equivalent states 1) as QALY differences and 2) as AHYE differences (using Table 5 estimates). Moving from the study state to the equivalent state in all four choice sets would involve a greater trade-off from the healthier Group 1 to the less healthy Group 2. By definition, social welfare is unaffected by all four moves, and on both metrics there appears to be inequality aversion.

Table 4: Study and equivalent points, plus trade-offs between them

Choice Set 1

	Group 1 Health	Group 2 Health
$\mathbf{x}_0$	66 years in 100% health 8 years in 25% health	50 years in 100% health 16 years in 25% health
$\tilde{\mathbf{x}}_0$	59.23 years in 100% health 4.92 years in 25% health	54.46 years in 100% health 8 years in 25% health
QALY differences	-7.54 QALYs	+2.46 QALYs
AHYE differences	-7.60 AHYEs	+2.31 AHYEs

Choice Set 2

	Group 1 Health	Group 2 Health
$\mathbf{x}_0$	66 years in 100% health 16 years in 25% health	4 years in 25% health 54 years in 100% health 4 years in 25% health
$\tilde{\mathbf{x}}_0$	59.40 years in 100% health 5.61 years in 25% health	54.81 years in 100% health 8 years in 25% health
QALY differences	-9.19 QALYs	+0.81 QALYs
AHYE differences	-9.38 AHYEs	+3.23 QALYs

Choice Set 3

	Group 1 Health	Group 2 Health
$\mathbf{x}_0$	72 years in 100% health 16 years in 25% health	48 years in 100% health 16 years in 25% health
$\tilde{\mathbf{x}}_0$	59.23 years in 100% health 4.92 years in 25% health	54.46 years in 100% health 8 years in 25% health
QALY differences	-15.54 QALYs	+4.46 QALYs
AHYE differences	-15.74 AHYEs	+4.31 QALYs

Choice Set 4

	Group 1 Health	Group 2 Health
$\mathbf{x}_0$	8 years in 25% health 72 years in 100% health 8 years in 25% health	48 years in 100% health 16 years in 25% health
$\tilde{\mathbf{x}}_0$	59.34 years in 100% health 5.34 years in 25% health	54.67 years in 100% health 8 years in 25% health
QALY differences	-15.33 QALYs	+4.67 QALYs
AHYE differences	-10.67 AHYEs	+4.53 QALYs

Estimates for the AHYE and r parameters are obtained using these study and equivalent states and using the methods above. In testing, it appeared that only a single solution satisfying multiplicative separability ($k(r) = 0$) appears to exist for each sample. Table 5 presents our estimates for the AHYE parameters with the uncertainty figures obtained by bootstrapping.

Table 5: AHYE parameters and r

		Study Estimates	Standard Deviation	95% Confidence Interval
Adult, 1.00 health	$B_{1.00}$	1.000	-	-
Adult, 0.25 health	$B_{0.25}$	0.268	0.012	(0.244, 0.292)
Child, 1.00 health	$b_{1.00}$	1.828	0.031	(1.768, 1.888)
Child, 0.25 health	$b_{0.25}$	0.490	0.027	(0.439, 0.542)

Estimates based on 5000 bootstrapped observations

In the AHYE, 25% health as an adult ($B_{0.25}$) is worth 0.268 of the value attached to a year of full health, as against 0.250 in the QALY. This is not significantly different, however (95% CI: 0.244, 0.292). Under standard QALY calculations 3 years of 25% health to an adult who otherwise would have died and one year of improvement from 25% health to full health for an adult are both worth 0.75QALYs. Using AHYEs, the former will be 0.804 AHYEs and the later is worth 0.732 AHYEs, suggesting that the AHYE might justify only a premium of 10% for severity (as poor quality of life) over the standard QALY calculations

In contrast, the AHYE gives a significantly higher weight to the health that we experience as children than does the QALY. Here, a year of full health received as a child ($b_{1.00}$) receives an 83% higher weight (1.828) than full health received as an adult (1.000). Likewise (by multiplicative separability), a year of 25% health as a child receives the same premium over a year of 25% health as an adult (0.490 versus 0.268).

6. Discussion

In this paper, we consider two concerns regarding the societal value of health that have been conflated in previous work. The first is a social concern for how an individual's lifetime health is comprised: the same number of QALYs may have a different social value if when those QALYs are experienced differs e.g. ill health in childhood may be weighted more highly than similar ill health in adulthood, holding everything else constant. The second is a social concern for the distribution of health across people. The former issue affects the arguments of the social welfare function, whilst the latter issue affects the function itself through changes in its parameters.

Our analysis is novel in that we measure social preferences over the former in terms of adult healthy-year equivalents. The closest alternative to our approach was in Bleichrodt et al (2004), where an application of the rank-dependent model found an approximately linear function for the value of a QALY between the ages of 10 and 40, and suggesting that those in worse health (according to ranking) received greater weight in the social value of a QALY.

Inevitably, our approach has required certain assumptions in order to obtain data. Whilst we do not believe that these assumptions are unreasonable, several could be relaxed e.g. we rely on assumptions for preferences between the reference states that could be tested; we consider states that occupy the same "neighbourhood" of a constant elasticity of substitution SWF; this functional form is problematic when extended to consider those expected to receive very little lifetime health. However, these are issues that can be addressed in further analysis of the current data and future studies. Our experimental design was also influenced by the need to be relatively robust to a range of possible preferences, rather than efficiency in identifying specific parameter values; again, this might be addressed in subsequent studies.

Existing literature has argued that the values placed on health benefits differ according to other characteristics beyond just the length and quality of life included in the QALY. According to such a view, the QALY might be expanded to include additional dimensions or – as in the current study – replaced by an alternative metric. It also suggests that the period of life up until 18 years of age contributes a higher proportion of our lifetime health than would be explained simply summing the number of QALYs received. There is also weak evidence that a QALY gained by giving a person four more years in 0.25 health may be worth more than giving the same person one more year in full health. Per QALY, this suggests that treatments moving people from extremely poor health to moderate health may be more valuable than those moving people from moderate to good health, or good to very good health. As expected, we also find that society places a premium on those treatments that reduce inequalities in lifetime health across groups.

By distinguishing between different types of possible reasoning, this study has aimed to provide novel insight into how social preferences can be more accurately represented within a SWF. More widely, our study suggests that attempts to place differential weights on certain QALYs should be treated with caution. Our framework suggests that a societal preference could be interpreted in different ways. For example, inequality aversion could be thought to suggest prioritising those groups in society who receive less lifetime health, whilst strong non-QALY reasoning in what constitutes health could be thought to suggest prioritising treatments for particular diseases (e.g. those that reduce childhood health, or that have more severe health consequences). As the interpretation of societal preferences will affect the policy recommendations that can be drawn from them, we require robust methods to analyse such preferences. To this end, we believe that this paper may provide a useful step forwards both in encouraging debate and in presenting a novel empirical approach.

REFERENCES

Andersson F, Lyttkens CH. Preferences for equity in health behind a veil of ignorance. Health Economics 1999; 8(5):369-378.

Bleichrodt H, Diecidue E, Quiggin J. Equity weights in the allocation of health care: the rank-dependent QALY model. Journal of Health Economics 2004; 23(1):157-171.

Busschbach JJV, Hessing DJ, De Charro FT. The utility of health at different stages in life: A quantitative approach. Social Science & Medicine 1993; 37(2):153-158.

Charny MC, Lewis PA, Farrow SC. Choosing who shall not be treated in the NHS. Social Science & Medicine 1989; 28(12):1331-1338.

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. British Medical Journal 1999; 318(7188):916-919.

Dolan P, Edlin R, Tsuchiya A. The relative societal value of health gains to different beneficiaries: Report to National Collaborating Centre for Research Methodology. http://www pcpoh bham ac uk/publichealth/methodology/docs/publications/JH11_Social_Value_QALY_Final_Report_Paul_Dolan_et_al_2008 pdf [Jan 2009]; 2008a.

Dolan P, Edlin R, Tsuchiya A. The relative societal value of health gains to different beneficiaries. Health Economics and Decision Science Discussion Papers University of Sheffield 2008; 2008b

Dolan P, Shaw R, Tsuchiya A, Williams A. QALY Maximisation and People's Preferences: A Methodological Review of the Literature. Health Economics 2005; 14(2):197-208.

Dolan P, 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 Papers 2002; 02-02.

Dolan P, Tsuchiya A. The social welfare function and individual responsibility: Some theoretical issues and empirical evidence. Journal of Health Economics 2009; 28: 210–220

Edwards AL. Techniques of Attitude Scale Construction. Appleton-Century-Crofts: New York; 1957.

Guilford JP. Psychometric Methods. McGraw-Hill: New York; 1954.

Johannesson M, Gerdtham U-G. A note on the estimation of the equity-efficiency trade-off for QALYs. Journal of Health Economics 1996; 15:359-368.

Lindholm L, Måns Rosén. On the measurement of the nation's equity adjusted health. Health Economics 1998; 7(7):621-628.

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.

Mooney G, Jan S, Wiseman V. Examining preferences for allocating health care gains. Health Care Analysis 1995; 3(3):261-265.

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.

Nord E. The Person-trade-off Approach to Valuing Health Care Programs. Med Decis Making 1995; 15(3):201-208.

Nord E. Cost-Value Analysis in Health Care: Making Sense out of QALYs. Cambridge, UK: Cambridge University Press; 1999.

Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. Health Economics 1999; 8(1):25-39.

Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs egalitarianism: An Australian survey of health issues. Social Science & Medicine 1995; 41(10):1429-1437.

Ratcliffe J. Public Preferences for the Allocation of Donor Liver Grafts for Transplantation. Health Economics 2000; 9(2):137-148.

Rodriguez E, Pinto JL. The social value of health programmes: is age a relevant factor? Health Economics 2000; 9(7):611-621.

Thurstone LL. A Law of Comparative Judgement. Psychological Review 1927a; 34:273-286.

Thurstone LL. The Method of Paired Comparisons for Social Values. Journal of Abnormal and Social Psychology 1927b; 21:384-400.

Ubel PA. How stable are people's preferences for giving priority to severely ill patients? Social Science & Medicine 1999; 49(7):895-903.

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 Transplantation 2001; 7(7):600-607.

Williams A. Intergenerational equity: an exploration of the 'fair innings' argument. Health Economics 1997; 6(2):117-132.