Paper presented at HESG/Aberdeen, Aug 2008; NOT FOR CITATION

The societal value of lifetime health: Measurement and inequality aversion.*

Richard Edlin[1], Paul Dolan[2], Aki Tsuchiya[3]

1. Background:

Within health economics, social welfare functions are often defined purely in health terms, with non-health outcomes disregarded [1-3]. Within such functions, the quality adjusted life year (QALY) compresses information about both quality and length of life into a single measure that can be interpreted in utility terms. The QALY values health as the summation of health-related quality of life 'utilities', where dead and full health are scaled to 0 and 1. The QALY is consistent with utility maximisation and welfarism (and hence also maximisation subject to a utilitarian social welfare function - SWF) if the QALY represents valid cardinal utility [4] and the value of health is consistent across individuals. Where the QALY is not interpreted as representing utility, such a function still represents sum ranking of health or a focus on the average health of the population, which is consistent with non-welfarism.

As a description of societal preferences, a utilitarian/sum-ranking SWF fails to accurately reflect the views of those in society on at least two counts. When asked about how society should allocate healthcare resources – that is, in the decision frame of a health-related SWF (HRSWF) – there is evidence that the public are sensitive to inequalities in health when their views are elicited [5-8]. There is also evidence that non-QALY arguments also matter within a HRSWF [9]. For example, in addition to efficiency and inequalities in QALY terms, the general public takes account of: age or the timing of illness within lifetime health [10-13]; the severity of ill health [14-16]; and other characteristics including individual responsibility.

---

* 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 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] Corresponding author: Leeds Institute of Health Sciences, University of Leeds; <r.p.edlin@leeds.ac.uk>
[2] Tanaka Business School, Imperial College, London
[3] School of Health and Related Research and Dept of Economics, University of Sheffield.

However, previous attempts at considering and capturing societal preferences typically consider *either* sensitivity to QALY inequalities *or* the introduction of non-QALY information, but not both at the same time. Williams' fair innings approach [17] probably provides the most robust and broadest theoretical treatment of sensitivity to QALY inequalities but has been criticised for overlooking issues with the appropriateness of the QALY as the sole measure of benefit [18]. This approach could be characterised as examining "equity" by making arbitrary assumptions about benefits. In contrast, Nord's cost-value analysis [19, 20] pays attention to the introduction of non-QALY information but does not account for inequality when interpreting preference data. This approach could be characterised as examining "benefit" by making arbitrary assumptions about equity.

Thus, while both approaches provide important critiques to a sum-ranking SWF, they cannot in themselves allow us to identify an appropriate SWF that reflects both QALY inequalities and the relevance of non-QALY information. The difficulty in doing so is that the two types of reasoning are inherently confounded in elicited preferences, and each of the above approaches are reasonable responses the problem. However, it may be possible to estimate both inequality aversion and non-QALY preferences together, without the need to make arbitrary assumptions about either. Basically, inequality aversion is a matter of the functional form of the HRSWF, and the incorporation of non-QALY information is about the arguments of the HRSWF.

2. The analytical framework

Suppose that public preferences can be represented by a constant elasticity of substitution SWF (Dolan, 1988) that assesses the health of two homogeneous and equally sized groups.

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

where: W is social welfare,

v_i is the lifetime health of Group *i*,

and r reflects the overall strength of inequality aversion.

Here, v_i can incorporate more information about the lifetime health of individuals than provided by the QALY. 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 in the v_i, the size of r relates to sensitivity to the relative differences in lifetime health.

We use a functional form for v_i that equals:

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

where V is a weighting function based on health-related quality of life (h_{it}) of group i at time t and timing (t) that increases in health ($\frac{dV}{dh_{it}} > 0$). If v_i is multiplicatively separable into health and timing components then:

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

Within the study, we used a dichotomous variable for timing representing whether health is experienced up to or after 18 years of age, and quality of life at 0.25 (or 25% health) and 1.00 (100% health). A quality of life level at 0.00 (0% health, or dead) is also used for computational reasons, although this is not valued. In contrast, the QALY takes the form $\sum h_{it}$ and thus includes no timing dimension or social weights. Expanding v_i we can write this as a function of six "lifetime health parameters":

$$v_i = FHC \cdot y_{FHC} + SHC \cdot y_{SHC} + DC \cdot y_{DC} + FHA \cdot y_{FHA} + SHA \cdot y_{SHA} + DA \cdot y_{DA} \quad \text{(Eq. 1)}$$

where:

FHC is the value of a year in full health whilst aged < 18,
SHC is the value of a year in 25% health whilst aged < 18,
DC is the value of being dead whilst aged < 18,
FHA is the value of a year in full health whilst aged ≥ 18,
SHA is the value of a year in 25% health whilst aged ≥ 18, and
DA is the value of being dead whilst dead aged ≥ 18,

and the y_{XXX} variables give the number of years spent in each health/time combination. For comparability with the QALY, *FHA* = 1 and *DA* = 0. If we can also say that health and timing are multiplicatively separable then:

$$v_i = y_{FHA} + FHC \cdot y_{FHC} + SHA \cdot y_{SHA} + (FHC \times SHA) \cdot y_{SHC}. \quad \text{(Eq.2)}$$

In order to measure health here, we use “adult healthy-year equivalents” (AHYEs) in preference to the QALY, where the former include societal weights and the latter do not. In other words, the number of AHYEs is a measure of benefit that includes non-QALY information. Our survey is designed around identifying a series of states of the world that appear to lie on the same social welfare contour. Since these should be valued equally in the SWF, any estimated difference in the level of social welfare between the points can be interpreted as an error. By varying the parameters of the social welfare function we aim to find a fit using both the inequality aversion parameter, *r*, and the lifetime health parameters defined above (and hence, AHYE).

It is not possible to simply minimise the sum of squared errors to find a best-fit solution in a simple one-step algorithm. Where the value attached to health as a child (*FHC*) increases, the amount of lifetime health will typically increase for all groups and, crucially, the relative size of health differences becomes smaller. For example, consider a scenario with two groups in which those in one group live for 40 years in full health before dying and those in the other lives for 80 years in full health before dying. Obviously, those in the first group obtain 40 QALYs and those in the latter 80 QALYs. But what if *FHC* = 10? Here, the first group receives 202 AHYEs (10x18 + 22) and latter 242 AHYEs (10x18 + 62). Twice as many QALYs are obtained by the latter group as the former, but only about 20% more AHYEs.

Suppose that this unequal case was deemed to be equivalent to another scenario in which both groups live for 50 years in full health. Assume further that, in both cases we are willing to give up 30 healthy adult years (=30 QALYs =30 AHYEs) from the better-off group in order to get 10 healthy adult years (=10 QALYs = 10 AHYEs) for the worse-off group. So in both cases we are willing to trade-off health at a rate 3:1 on average. In the QALY case we are willing to do this in order to get equality only when there is a 100% difference in health; in the AHYE case we are willing to do this where there is only a 20% difference in health. Hence, we appear to be more sensitive to differences in relative inequalities (so have a higher *r*) in the AHYE case.

More generally, if questions relate to trade-offs in adult health then as *FHC* rises so too will the value of *r*. Further, since the relative size of differences between equivalent points fall (in AHYE space), so too will their differences in their estimated social welfare levels. If the error can be reduced by simply increasing *FHC* or *r* (and solving for the remaining parameters), then minimising the sum of squared errors will not estimate a social welfare function as the estimates will fail to converge.

Instead, we use Equation 1 (with $FHC = 1$, $DA = 0$) and find parameters consistent with Equation 2. If we do not use options that consider death to a child (i.e. y_{DC} =0), DC can be left undefined. (Where Equation 2 holds $DC = FHC \times DA = 0$.) If we take *r* as given and minimise squared errors with respect to *FHC*, *SHA*, and *SHC* then solutions can be found. Define $FHC(r)$, $SHA(r)$ *and* $SHC(r)$ as the functions for the parameters in *r*. As *r* increased, we found that SHA remained approximately constant whilst FHC and SHA increased approximately linearly with a generally consistent difference between them. Now, define $k(r) =$ $FHC(r)$ × $SHA(r)$ - $SHC(r)$ as the measure of the difference between the value of *SHC* against the result that could be derived with multiplicative separability. Where $k(r) = 0$, Equation 2 holds. (Within testing prior to obtaining actual data there appears to be a single solution in *r*. This was confirmed in the actual data.)

The survey in our study identified states of the world on the same social welfare contour by asking series of pairwise choices that compare a “study state” (x_5) against four “reference states” (x_1 to x_4). Each set of four pairwise choices is referred to as a “Choice Set”, and is used to identify a pair of indifferent points. The reference states are constructed to be collinear in the y_{XXX} variables (see Figure 1) with $x_1 \succ x_2 \succ x_3 \succ x_4$ so long as years spent in states better than dead at an individual level are also valued positively (or not too negatively) at a societal level. Given preferences for x_5 versus x_1 to x_4, we can find a consistent rank order (e.g. $x_1 \succ x_5 \succ x_2 \succ x_3 \succ x_4$) where the individual preferences allow.

However, we assume a random utility model in which such judgements are made with error so that x_i is only preferred to x_j if the assessment of social welfare $W(x_i)$ exceeds $W(x_j)$. These figures are drawn independently from normal distributions with means $\overline{W_i}$ and $\overline{W_j}$ and variance σ^2. Now the larger is the welfare difference between the states $\overline{W_i}$-$\overline{W_j}$, the more often x_i will be preferred. Let $\rho(x_i, x_j)$ is the proportion of the time that x_i is preferred to x_j, with indifference split equally between the two states. Trivially, $\rho(x_i, x_i)$ =0.5. For:

$$\rho(x_i, x_j) = \int_{-z(x_i, x_j)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Each $z(x_i, x_j)$ value is a function of the difference in social welfare between the two states. The average of the z values across all $\rho(x_i, x_j)$ gives a measure of the aggregate preference for x_i and since each state is compared against all other states, these aggregate preferences are compared in a consistent way across states. This method is based on Case V Thurstone scores [21, 22] and provides cardinal scale values for the social welfare of each state at the aggregate level across respondents. Within this method, extreme values are typically censored to fall within a set range, since unanimity will cause indeterminate scores. The typical response to this issue is either to omit such data or restrict proportions to fall within a permitted range [23, 24] and we use 2%-98% range within our main analysis.

Values for the study state can be found using the direct valuations of x_5 against four reference states x_1 to x_4, plus $\rho(x_5, x_5)$ =0.5. For the four reference states $\rho(x_i, x_5)$ = 1- $\rho(x_5, x_i)$, $\rho(x_i, x_i)$ =0.5, and the preferences against the other three states are found using the expected ordering $x_1 \succ x_2 \succ x_3 \succ x_4$ (so $\rho(x_i, x_j)$ =0.02 or 0.98). With estimates for $\overline{W_i}$, we define a sixth, 'equivalent', state (x_6) as:

$$x_6 = x_3 + \frac{\overline{W_5} - \overline{W_3}}{\overline{W_2} - \overline{W_3}} x_2$$

Here $\frac{\overline{W_5} - \overline{W_3}}{\overline{W_2} - \overline{W_3}}$ indicates the relative preference of x_5 against x_2 and x_3. Since the estimated social welfare values are measured on a cardinal scale, x_5 and x_6 should have the same social welfare (so $\rho(x_5, x_6)$ =0.5).

So given a set of questions, we first find sets of socially equivalent points and these are then used to populate a social welfare function. However, this gives only a point estimate for the parameters of the social welfare function. This would be of relatively little value to decision makers (as a framework) even in an ideal world where all relevant methodological questions were answered, as it does not provide any indication of the inherent uncertainties in the analysis. To resolve this we use bootstrapping (sampling with replacement). 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 an estimate of uncertainty in the point estimate of each parameter. In this way, we can address the non-methodological uncertainties of our estimates.

4. Methods

Our study attempted to assess the relative health gains to different beneficiaries using the HRSWF framework above. The methods here report our main study findings in full, with an additional study (using a similar methodology) reported in summary form. Further details on the additional study are available on request.

Recruitment:

Interviewing was conducted by the Centre for Research and Evaluation at Sheffield Hallam University using a UK-wide sample. 17 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. Participants had earlier been informed about the study (by letter). The sampling frame aimed to recruit 30 people in each combination of gender, age (18-39, 40-59, 60+), and education (high school level or below, further/higher education and above).

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 only the eight choice sets that are reported here. The other choices considered a slightly different SWF that used an additional parameter.

Choice Sets 1-4 (CS1-CS4) involve choices where all life years are lived in 100% health, after which all those in the groups die. They are used to test the accuracy of SWF that enumerated in QALY terms. This also provided an opportunity for respondents to gain familiarity with the question format. Choice Sets 5-8 (CS5-CS8) typically involve a period of ill health at 25% quality of life. 4 years in 25% health was explained to the respondent as being worth 1 year in 100% health to the individual experiencing it. These questions are used to find weights given for childhood versus adult health, 25% health versus 100% health, and a value of r. Figure 2 gives an example of one of the pairwise choices (CS5). Here, 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.

Note that the questions above do not discriminate between adult and childhood health. Textual, graphical, and auditory prompts were used for each question and respondents 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").

CS1 and CS2 use the same reference states but differ in their study state. CS3 and CS4 use the same states as CS1 and CS2 but halve the amount of health received in both the study and reference states. Table 1 presents the study and reference states for CS1, CS2 and CS5-CS8.

<u>Analysis:</u>

The individual-level responses for CS1-CS8 are analysed in terms of whether they allow a rank-consistent ordering to be formed between x_1 to x_5 using the assumed order between x_1 to x_4. To investigate individual-level rationality, individual-level responses to each choice set were analysed according to whether they were consistent with a rank ordering between x_1 to x_5. Summary variables were constructed that identified the implied rank ordering where this was possible and "non-rank" where not.

Since the CES function is homothetic, the level of *r* giving indifference between the study and equivalent state should be the same between CS1 and CS3, and between CS2 and CS4. The parameters for the AHYE-based HRSWF are found using CS5-CS8 and the methods above.

As the methods above provides only point estimates, we use bootstrapping (sampling with replacement) to analyse uncertainties. 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. Subsequent analyses were based on the number necessary to find convergence here. Sensitivity testing was also conducted to assess the impact of a change in the range of the allowable pairwise data from (0.02, 0.98) to (0.01, 0.99) and (0.05, 0.95).

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. Our data was obtained from 17 areas throughout England and Wales. Whilst we selected a range of geographical locations, our emphasis was to obtain a sample across the range of gender, age, and education.

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. 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 3 shows the proportion of the 559 respondents with complete data that prefer the study state (x_5) to the reference states (x_1 to x_4). As expected, the study state is preferred less often (a lower percentage) for better reference states at an aggregate level.

The majority of individual-level responses are consistent with a personal rank ordering on every choice set. However, there is some evidence of individual learning. For choice set 1, 79% of individual-level responses allowed a rank ordering, compared with 84-86% across the remaining questions involving full health only (CS2-CS4). Where periods of ill-health were introduced (CS5), the complexity of task increased. This was reflected in the lower individual-level rank consistency in choice set 5 (57%), although this increased in subsequent choice sets (65%, 73%, 74% for CS6-CS8). This suggests that whilst individuals found the task challenging it was feasible.

Testing total QALYs as an outcome dimension

Each option in CS2 and CS4 was designed to contain precisely half the health (in QALY terms) of the equivalent option in CS1 and CS3. If we assume that lifetime health equals the number of QALYs obtained within a CES SWF, then we would expect similar results for CS1 and CS3, and CS2 and CS4. Of the individuals giving rank-consistent preferences to CS1 and CS3, 51% (n = 201/393) choose in exactly the same way across the two choices, whilst 93% (n = 367/393) changed in no more than one choice. Of the individuals giving rank-consistent preference to CS2 and CS4, 47% (n=204/431) gave exactly the same response to both questions and 85% (n=367/431) changed no more than one choice. However, chi-squared tests of the summary variables show a strong and significant difference between the results of both pairs of choice sets ($p < 0.001$).

The Thurstone equivalent state lies along the same line as x_1 through to x_4. Table 4 gives the pairs of study states (x_5) and the equivalent states (x_6) for CS1 to CS4. Within CS1 to CS4, the aggregate preferences suggest a trade-off between total health and reducing inequalities. In CS1, for example, respondents are willing to sacrifice 5.84 QALYs (70.00 – 64.16 QALYs) to the better off group in order to obtain 2.16 QALYs (58.16 – 56.00 QALYs) for the worse off group. This suggests an implicit marginal rate of substitution between the health of the worst off to the best off of 2.7; that is, the health of the worst is worth 2.7 (= 5.84 / 2.16) that of the health of the best off. For the other choice sets, this figure varies from 1.42 (CS4) to 1.94 (CS3).

Table 5 reports the inequality aversion parameters for CS1 to CS4 under the assumption that the QALY describes the way that society judges lifetime health. For example, the trade-off for CS1 suggests an inequality aversion parameter of $r = 5.24$ within a SWF based only on that point. The central estimates for the r parameters appear to differ across the four choice sets. Consider a case where Group 1 has a life expectancy of 70 years in full health and Group 2 has a life expectancy of 60 years in full health. The inequality aversion estimate from CS1 would suggest that this is equivalent to case where both groups live 63.83 years in full health. Here, Group 1 loses 6.17 years and Group 2 gains 3.83 years, suggesting that across these improvements Group 2's health is worth 61% more than Group 1's health (6.17/3.83-1). For CS2-CS4, the comparable figures are 21%, 38% and 13%, respectively.

However, the differences found between CS1 and CS3, and between CS2 and CS4 do appear to suggest a general violation of the CES-SWF in the case considered here. Here, either the QALY is not a cardinal measure of (socially-assessed) utility to an individual, the CES SWF is inappropriate, or both. Given both the weight of prior evidence and the broad categorisation of preferences allowable under a CES SWF, the first of these options appears plausible.

<u>Deriving a measure of social benefit within the SWF</u>

Moving away from the QALY model, we allow weights for both the severity of the health state and the timing of ill-health (adult versus child). These are found using CS5 to CS8, in which the options presented involve periods of both ill-health.

Table 6 shows both the study states (x_5) and their social welfare equivalents (x_6) derived from the results of CS5 to CS8. Within each of the equivalent states, ill-health always occurs at the end of life so, for instance, the equivalent state in CS5 involves 59.23 years of full health, followed by 4.92 years in severe health. The trade-offs defined in these states are complex, as they involve periods in ill-health, periods in good health, health as children and health as adults.

Estimates for the HRSWF parameters are presented in Table 7, along with the assumptions of a standard QALY-based CEA. The standard CEA assumes no inequality aversion in its objective function (r = -1), whilst the SWF found here has significantly higher inequality aversion. When defining lifetime health (the AHYE), years spent as a child (FHC > 1, SHC > 0.25) obtain a significantly higher weight than predicted by the standard QALY model. Overall, the AHYEs gives 96% more weight to the first 25% health for children relative to that given by the QALY. In contrast, years spent in 25% health as an adult obtain an insignificantly higher weight (SHA = 0.268) than predicted by the standard model.

Convergence of the bootstrapped estimates for the standard deviation of parameter estimates were found very quickly, with n=100 sufficient in most cases. The analysis of later portions of the study (including some sensitivity analysis and analysis of the additional study) used n=200 bootstrapped iterations. Sensitivity testing also revealed that changing the cut-off values when computing Thurstone scores had no significant effect on parameter estimates.

<u>Additional study results</u>

The additional study expanded on the main-study social welfare function by replicating the choice sets CS5-CS8 (as A1-A4) but used 50% health states (for half the duration) in place of 25% health states. The best-fit solution here (using *FHC* and *r* from the main study) suggests that 50% health is weighted as equivalent to 0.615 of an adult healthy year – this suggests that the health between dead and 50% health is worth 60% more ($\frac{0.615}{0.5}/\frac{0.385}{0.5}=1.60$) than the health between 50% health and full health. The uncertainty in this figure is found through bootstrapping (n=200), and it appears that the value given to 50% health is significantly greater than 0.50 (95% CI, 0.533-0.698).

Choice Set A5 considers ill-health (at 50%) between years 10 and 18; for the inequality aversion parameter and age/timing outside these ages, the values from the main study apply. Between 10 and 18 years, the value that minimises squared utility errors over choice sets A1-5 assumes that a year of full health whilst a child is worth 1.073 times as much a year of full health as an adult (95% CI: 0.928-1.218). That is, it appears that health after the age of 10 is viewed very much like health to an adult. (In light of this finding, the main study analysis was re-run with an age cut-off of 10 years, and the only noticeable impact was an insignificant reduction in inequality aversion (r=5.68) which is likely to be due to a drop in the general level of AHYEs achieved within each option.)

Choice Set A6 considers ill-health between years 30 and 38, and was used to construct a weighting for the early period of adulthood (18 to 40 years). As with A5, timings outside these ages and inequality aversion parameters were taken from the main study. The numéraire period for health (*FHA* = 1.000) in which a year in full health equals 1 AHYE is now the period from 40 years onwards. Between 18 and 40 years, the value that minimises errors over Choice Sets A1-A4 plus A6 assumes that a year in full health is worth 0.989 AHYEs (95%CI: 0.888-1.091). Again, this suggests that those aged 30-38 (and by extension 18 to 40 year olds) are not treated differently from older adults.

Finally, A7 used both 25% health and 50% health in the same question by using the study state from CS5 – i.e. the same as A1 with periods of ill-health at 25%. The weight for SHA was selected here in order to solve for equality in utilities between the Thurstone-equivalent and study state in Choice Set A7. The best-fit figure here suggests that the first 25% of health is worth 0.231 of a full health life year (95% CI 0.212-0.249). This estimate is significantly lower than the estimate suggested in the main study for 25% health.

6. Discussion

We believe our methods and analyses are unique in estimating both the social valuation of lifetime health reflecting non-QALY information and the value given to more equal distributions of lifetime health together. However, we do not claim that considering both the definition of benefit and inequality aversion in a SWF provides a complete account of societal preferences regarding health care resource allocation (and are confident that it does not).

The closest alternative we could find to our approach was provided by an application of the rank-dependent model that 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 [25]. Our analysis in contrast used societal preferences and considered the relative size of differences between groups rather than rank order. Further, had this research aimed to find an inequality measure rather than a SWF, our choice of tools may have been different. The CES function has been criticised as restrictive by those advocating inequality-based measures [26], as has social welfare more generally [27]. An inequality measure does not consider a role for efficiency, it so is of limited usefulness in our current context.

This paper suggests that individuals treat the quality of life of children and adults differently. Although there is uncertainty as to where the demarcation between “childhood” and “adulthood” lies, there seems to be almost twice the value given to health experienced in the former. This research suggests that the QALY model is a reasonable approximation of what individuals care about at a societal level for adults, although there is some evidence that the quality of life weights are not necessarily used at a societal level. In particular, there is some evidence that obtaining a low level of health (i.e. improvements from “dead”) may be more valuable (per QALY gain) than improving the health of those already in good health. However, it may be more valuable still to move someone still in a relatively poor health state to one that is considerably better, but still not “healthy”. Further research is warranted to provide greater understanding of the relationship between social value and improvements in quality of life.

This study should be interpreted as an attempt to explore the methodology of eliciting preference rather than providing anything close to a finished product. At present, the analysis relies on a CES functional form that can not be tested with our existing data given the inter-relationships between benefits and inequality. There is also aggregate intransitivity in the SWF, which is of concern. The study state in CS8 involves similar health to that in CS7 but where some of the ill-health has been transferred to the start of life. The higher weight given to health whilst very young suggests that CS8 *should* have lower social welfare since less lifetime health is received by Group 1. However, the equivalent state involves more health in CS8 than in CS7 – suggesting a higher social welfare, although this difference is not significant.

Whilst the current paper does not provide a “super-QALY”, it shows that research combining disparate factors in societal decision making is feasible and contributes to an understanding of what matters when addressing health care resource allocation from a societal perspective.

REFERENCES

1. Dolan P. The measurement of individual utility and social welfare. *Journal of Health Economics* 1998; 17(1): 39–52.
2. Dolan P, Robinson A. The measurement of preferences over the distribution of benefits: The importance of the reference point. *European Economic Review* 2001; 45(9):1697-1709.
3. Abásolo I, Tsuchiya A. Exploring Social Welfare Functions and violation of monotonicity: An example from inequalities in health. *Journal of Health Economics* 2004; 23(2): 313-329.
4. Johannesson M. The relationship between cost-effectiveness analysis and cost-benefit analysis. *Social Science and Medicine* 1995; 41(4):483–489.
5. Johannesson M, Gerdtham UG. A note on the estimation of the equity-efficiency trade-off for QALYs. *Journal of Health Economics* 1996; 15: 359–368.
6. 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.
7. Lindholm L, Rosen M. On the measurement of the nation's equity adjusted health. *Health Economics* 1998; 7(7): 621–628.
8. Andersson F, Lyttkens CH. Preferences for equity in health behind a veil of ignorance. *Health Economics* 1999; 8(5): 369–378.
9. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature, *Health Economics* 2005; 14:197-208.
10. Busschbach JJV, Hessing DJ, Decharro FT. The utility of health at different stages in life – a quantitative approach. *Social Science and Medicine* 1993; 37(2):153–158.
11. Mooney G, Jan S, Wiseman V. Examining preferences for allocating health-care gains. *Health Care Anal* 1995; 3(3): 261–265.
12. Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs. egalitarianism – an Australian survey of health issues. *Social Science and Medicine* 1995; 41(10): 1429–1437.
13. Rodriguez E, Pinto JL. The social value of health programmes: is age a relevant factor? *Health Economics* 2000; 9(7): 611–621.
14. Nord E. The trade-off between severity of illness and treatment effect in cost-value analysis of healthcare. *Health Policy* 1993; 24(3): 227–238.
15. Nord E. The person-trade-off approach to valuing health-care programs. *Medical Decision Making* 1995; 15(3): 201–208.
16. Ubel PA. How stable are people's preferences for giving priority to severely ill patients? *Social Science and Medicine* 1999; 49(7): 895–903.
17. Williams A. Intergenerational equity: an exploration of the 'fair innings' argument, Health Economics 1997; 6:117-132.
18. Nord E. The Desirability of a Condition Versus The Well Being and Worth of a Person. *Health Economics* 2001:579-581.
19. Nord E. *Cost-Value Analysis in Health Care: Making Sense out of QALYs*. 1999. Cambridge (UK): Cambridge University Press.
20. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* 1999; 8:25-39.
21. Thurstone LL. A law of comparative judgment. *Psychological Review*, 1927:34:273-286.

22. Thurstone LL . The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology* 1927:21:384-400.
23. Guilford JP. *Psychometric methods*. 1954. New York: McGraw-Hill.
24. Edwards AL *Techniques of attitude scale construction*. 1957. New York: Appleton-Century- Crofts.
25. Bleichrodt H, Bleichrodt H, Diecidue E, Quiggin J. Equity weights in the allocation of health care: the rank-dependent QALY model. *Journal of Health Economics* 2004; 23:157-71.
26. Dutta I (2007). Health inequality and non-monotonicity of the health related social welfare function. *Journal of Health Economics* 26(2) 414-31.
27. Bommier A, Stecklov G (2002). Defining health inequality: why Rawls succeeds where social welfare theory fails. *Journal of Health Economics* 21(3):497-513.

Figure 1: Reference and study states.

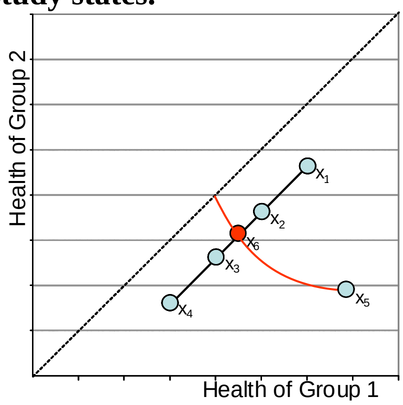

Figure 2: Sample Choice from CS5

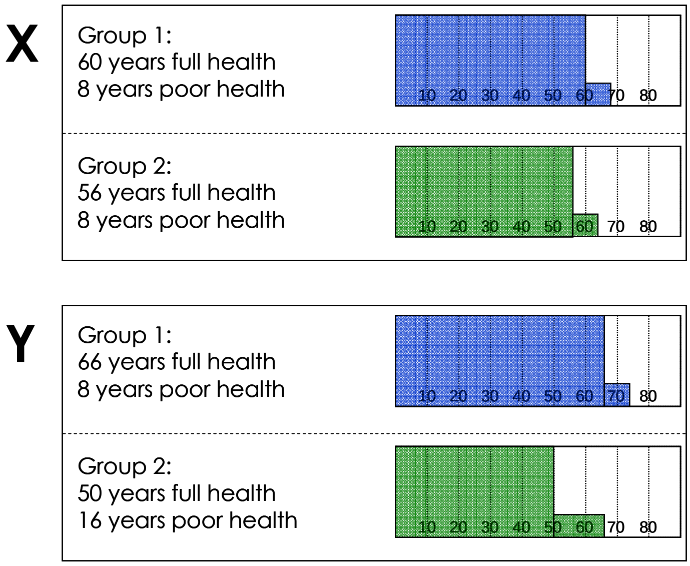

Table 1: Choice and Reference States and Equivalents

Choice Set & States	Group 1 Health	Group 2 Health
X_1	70 years in 100% health	64 years in 100% health
X_2	66 years in 100% health	60 years in 100% health
X_3	62 years in 100% health	56 years in 100% health
X_4	58 years in 100% health	52 years in 100% health
X_5 (CS1)	70 years in 100% health	56 years in 100% health
X_5 (CS2)	74 years in 100% health	52 years in 100% health

Choice Set & States	Group 1 Health	Group 2 Health
X_1	62 years full health 16 years poor health	60 years full health 8 years poor health
X_2	60 years full health 8 years poor health	56 years full health 8 years poor health
X_3	59 years full health 4 years poor health	54 years full health 8 years poor health
X_4	58 years full health	52 years full health 8 years poor health
X_5 (CS5)	66 years full health 8 years poor health	50 years full health 16 years poor health
X_5 (CS6)	66 years full health 16 years poor health	4 years poor health 54 years full health 4 years poor health
X_5 (CS6)	72 years full health 16 years poor health	48 years full health 16 years poor health
X_5 (CS6)	8 years poor health 72 years full health 8 years poor health	48 years full health 16 years poor health

Table 2: Background of the sample

Sample size		Social QALY sample (%)	2001 Census (%)
Gender:	Female	55	52
Age:	18-29	20[a]	19
	30-39	16	20
	40-49	17	17
	50-59	15	16
	60-69	17	12
	70+	15	15
Ethnicity	White	95	92[b]
Employment status:	Self-employed	7	8
	Other Employed	39	52
	Retired	29	14
Education:	School only	47	78b
	HE/FE	53	22[b]
House ownership:	Owned/mortgage	71	71[c]
Disabled?	Yes	14	18[d]

[s] Includes 14 aged below 18.
[b] Ages 16-74 only
[c] 2000 data. Office of National Statistics.
[d] Limiting long-term illness .

Table 3: Proportions preferring study state by reference state and choice set

Choice Set	vs. Reference State			
	x_1 (best)	x_2	x_3	x_4 (worst)
CS1	13%	37%	77%	77%
CS2	18%	33%	64%	78%
CS3	11%	39%	84%	86%
CS4	24%	40%	73%	83%
CS5	37%	38%	42%	48%
CS6	43%	46%	47%	52%
CS7	36%	38%	43%	47%
CS8	39%	43%	46%	50%

Table 4: Study States and Equivalents: CS1-CS4

Choice Set & States		Group 1 Health	Group 2 Health	Trade-offs
CS1	Study State	70 years in 100% health	56 years in 100% health	2.70 QALYs (Group 1) per QALY (Group 2)
	Equivalent Difference	64.16 years in 100% health - 5.84 QALYs	58.16 years in 100% health + 2.16 QALYs	
CS2	Study State	74 years in 100% health	52 years in 100% health	3.64 QALYs (Group 1) per QALY (Group 2)
	Equivalent Difference	63.80 years in 100% health - 10.2 QALYs	57.80 years in 100% health + 2.80 QALYs	
CS3	Study State	35 years in 100% health	28 years in 100% health	1.94 QALYs (Group 1) per QALY (Group 2)
	Equivalent Difference	32.36 years in 100% health - 2.64 QALYs	29.36 years in 100% health + 1.36 QALYs	
CS4	Study State	37 years in 100% health	26 years in 100% health	2.03 QALYs (Group 1) per QALY (Group 2)
	Equivalent Difference	32.31 years in 100% health -4.69 QALYs	29.31 years in 100% health + 2.31 QALYs	

Results from samples of 5000 bootstrapped observations

Table 5: Inequality parameter (*r*) estimates by choice set

Choice Set	Mean	Minimum	Maximum	Std Dev	95% CI
CS1	5.24	3.64	6.96	0.50	(4.266, 6.216)
CS2	1.51	0.78	2.19	0.19	(1.139, 1.871)
CS3	3.16	1.68	5.01	0.45	(2.289, 4.041)
CS4	0.55	-0.05	6.96	0.18	(0.207, 0.895)

Estimates based on 5000 bootstrapped observations

Table 6: Study States and Equivalents: CS5-CS8

Choice Set & States		Group 1 Health	Group 2 Health
CS5	Study State	66 years in 100% health 8 years in 25% health	50 years in 100% health 16 years in 25% health
	Equivalent	59.23 years in 100% health 4.92 years in 25% health	54.46 years in 100% health 8 years in 25% health
CS6	Study State	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
	Equivalent	59.40 years in 100% health 5.61 years in 25% health	54.81 years in 100% health 8 years in 25% health
CS7	Study State	72 years in 100% health 16 years in 25% health	48 years in 100% health 16 years in 25% health
	Equivalent	59.23 years in 100% health 4.92 years in 25% health	54.46 years in 100% health 8 years in 25% health
CS8	Study State	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
	Equivalent	59.34 years in 100% health 5.34 years in 25% health	54.67 years in 100% health 8 years in 25% health

Results from samples of 5000 bootstrapped observations

Table 7: SWF parameters; standard CEA assumptions and study estimates

Choice Set	Label	CEA	Study Estimates	Standard Deviation	95% CI
Inequality aversion parameter	r	-1.00	6.32	0.29	(5.76, 6.88)
Lifetime health judgements					
Value of 100% health as an adult	FHA	1.000	1.000	-	-
Value of 100% health as a child	FHC	1.000	1.828	0.031	(1.768, 1.888)
Value of 25% health as an adult	SHA	0.250	0.268	0.012	(0.244, 0.292)
Value of 25% health to a child	SHC	0.250	0.490	0.027	(0.439, 0.542)

Estimates based on 5000 bootstrapped observations