

A person trade-off study to estimate age-related weights for health gains in economic evaluation

Petrou, Stavros ; Kandala, Ngianga-Bakwin ; Robinson, Angela; Baker, Rachel

Published in:
PharmacoEconomics

DOI:
[10.1007/s40273-013-0085-y](https://doi.org/10.1007/s40273-013-0085-y)

Publication date:
2013

Document Version
Author accepted manuscript

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):
Petrou, S, Kandala, N-B, Robinson, A & Baker, R 2013, 'A person trade-off study to estimate age-related weights for health gains in economic evaluation', *PharmacoEconomics*, vol. 31, no. 10, pp. 893-907 .
<https://doi.org/10.1007/s40273-013-0085-y>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

A PERSON TRADE-OFF STUDY TO ESTIMATE AGE-RELATED WEIGHTS FOR HEALTH GAINS IN ECONOMIC EVALUATION

Running Head: Age-related weights in health care decision-making

Authors: Stavros Petrou,^a Ngianga-Bakwin Kandala,^a Angela Robinson^b and Rachel Baker^c

^a Warwick Medical School, University of Warwick, Coventry, UK.

^b Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK.

^c Yunus Centre for Social Business and Health, Glasgow Caledonian University, Glasgow,
UK.

Contact for correspondence:

Stavros Petrou

Warwick Medical School

University of Warwick

Coventry

CV4 7AL

UK

Tel: 02476 151124

FAX: 02476 151586

E-mail: S.Petrou@warwick.ac.uk

Abstract

Background: An increasing body of literature is exploring whether the age of the recipient of health care should be a criterion in how health care resources are allocated. The existing literature is constrained both by the relatively small number of age comparison groups within preference elicitation studies, and by a paucity of methodological robustness tests for order and framing effects and the reliability and transitivity of preferences that would strengthen confidence in the results. This paper reports the results of a study aimed at estimating granulated age-related weights for health gains across the age spectrum that can potentially inform health care decision-making.

Methods: A sample of 2500 participants recruited from the health care consumer panels of a social research company completed a person trade-off (or ‘matching’) study designed to estimate age-related weights for 5 year and 10 year life extensions. The results are presented in terms of matrices for alternative age comparisons across the age spectrum.

Results: The results revealed a general, although not invariable, tendency to give more weight to health gains, expressed in terms of life extensions, to younger age groups. In over 85% of age comparisons, the person trade-off exercises revealed a preference for life extensions by the younger of the two age groups compared. This pattern held regardless of the method of aggregating responses across study participants. Moreover, the relative weight placed on life extensions by the younger of the two age groups was generally, although not invariably, found to increase as the age difference between the comparator age groups increased. Further analyses revealed that the highest mean relative weight placed on life extensions was estimated for 30 year olds when the ratio of means method was used to aggregate person trade-off responses across study participants. The highest mean relative weight placed on life extensions was estimated for 10 year olds for 5 year life extensions and

for 30 year olds for 10 year life extensions, when the median of individual ratios method was used to aggregate person trade-off responses across study participants. Methodological tests framed around alternative referents in the person trade-off questions and the stability of preferences had no discernible effects on the study results.

Conclusion: This study has produced new evidence on age-related weights for health gains that can potentially inform health care decision-making.

Key points for decision makers

- This study provides evidence that individuals tend to give more weight to health gains, expressed in terms of life extensions, to younger age groups.
- The highest mean relative weight placed on life extensions was evident for older children and young adults.
- The results of this study should inform deliberations on whether health gains should be weighted according to age in health care decision-making. Nevertheless, they should be subject to re-retesting. Moreover, future research should aim to estimate distributional weights for health gains according to combinations of characteristics of potential relevance to decision-makers, for example, age in combination with gender.

1. Introduction

Economic evaluation has increasingly been used to inform the health care decision-making processes of government agencies throughout the industrialised world [1-3]. Decision-makers commonly seek to maximise health benefits given constrained health care resources. A dominant maximand proposed for health economic evaluation remains the quality-adjusted life year (QALY), a preference-based measure of health outcome that combines length of life and health-related quality of life in a single metric [4]. In economic evaluation, the incremental cost-effectiveness of a health intervention or programme is commonly compared to a cost-effectiveness threshold, a pre-determined value for money benchmark for the unit of health gain of interest (e.g. a QALY gain). A health intervention or programme with an incremental cost-effectiveness ratio (ICER) that falls below the threshold is deemed by decision-makers to offer good value for money and is generally recommended on cost-effectiveness grounds. In contrast, a health intervention or programme with an ICER that falls above the threshold is deemed by decision-makers not to offer good value for money and is generally rejected on cost-effectiveness grounds. In England and Wales, the National Institute for Health and Care Excellence (NICE) applies a threshold of £20,000-£30,000 per QALY gained [5, 6]. In other European countries, there are few public statements about value for money benchmarks for an additional QALY, although these might rest between \$10,000 and \$50,000 [7]. In the US, the threshold has been claimed to rest at \$50,000 per QALY gained, although the scientific basis for this threshold appears to be nebulous with no evidence of its consistent use as a tool for informing the regulatory and reimbursement decisions of federal programmes [8]. Moreover, the bases for these thresholds largely remain unclear, or at least under-specified, particularly with regards to how they might reflect shadow prices of budget constraints consistent with the health maximisation principle [9].

A methodological concern in the health economics literature, which has gained real policy traction, surrounds whether the health gain of interest (e.g. a QALY gain) should be given equal value in health care decision-making across different population groups. Decision-making bodies have commonly returned to the position that an additional unit of health should be of equal value regardless of the socio-demographic characteristics of the recipient or their pre- or post-treatment levels of health. Nevertheless, many decision makers take the characteristics of the beneficiaries of interventions into account, at least in their deliberations [6]. Indeed, research studies have been funded with the explicit remit of identifying relative weights that might be attached to health gains derived by different beneficiaries in the event of policy makers' support for differential weighting [10, 11]. Empirical evidence in the literature suggests that some people may be willing to give greater priority to certain population groups in health care decision-making, such as those in the poorest health states [12], those with more urgent conditions [13], those from lower social classes [14] and those with dependents [15], and less priority to those deemed responsible for their ill-health [16]. Overviews of the empirical evidence on relative weights that might be attached to health gains derived by different beneficiaries can be found elsewhere [10, 11].

In this context, an increasing body of literature is exploring whether the age of the recipient of health care might be considered a relevant criterion in how health care resources are allocated. A number of revealed preference studies suggest that the value individuals place on reducing health risks or achieving health gains may be higher for children than for adults [17-20]. However, many of the values estimated in the revealed preference literature are based on choices made by parents on the part of their families and consequently are likely to be driven by the effects of self-interest [21]. Perhaps more pertinent evidence is provided by stated preference studies that ask study participants to act as the social decision maker motivated by some notion of what is good for the society of which they are part. A number of stated preference studies have focussed on potential age-related weights for health gains

that might be applied within an economic evaluation framework [22-30]. This literature is constrained both by the relatively small number of age comparison groups within studies, with the number of ages assessed varying between two [29] and five [23, 26-28], and by a paucity of methodological robustness tests for order and framing effects and the reliability and transitivity of preferences that would strengthen confidence in the results. Furthermore, few attempts have been made to disentangle different forms of ageism within individuals' preference structures; for example, 'health maximisation ageism' [28] or 'utilitarian ageism' [31], where preference is given to the treatment of younger patients simply because they are expected to live longer after treatment; 'productivity ageism' [28], where preferences are determined by the recipients' productivity potential; or 'fair innings ageism' [28] (or 'egalitarian ageism' [31]), where preferences are driven by an aversion to inequality in age of death.

This paper builds on the current literature by presenting the results of a person trade-off study that calculated age-related weights for health gains across the entire lifespan at a more granulated level than has hitherto been achieved. In the process, it addresses a number of methodological issues surrounding the use of the person trade-off approach in this context.

2. Methods

2.1 Sample and data collection

We recruited a sample of 2500 participants from the health care consumer panels of a social research company (Synovate Healthcare) during February and March 2011. Members of the health care consumer panels received individualised e-mails inviting them to participate in an internet survey. Individuals who responded positively to the initial invitation subsequently

received an on-line link to the survey, which they were asked to complete. The survey had been developed according to good practice principles for on-line research [32], and subsequently piloted amongst a convenience sample of staff and students at the University of Warwick. Straightforward unambiguous non-open-ended questions, without requirement for free text entry, were used throughout the survey. The study sample was restricted to residents of the United Kingdom, aged at least 18 years of age. Each participant was compensated to the value of £3 for their time.

2.2 Research design

The research design built upon the methods of the UK SVQ Study [10], which included a person trade-off study (or ‘matching’ study based on nomenclature in the transport literature) aimed at estimating the relative weights to be attached to health gains according to characteristics of recipients of those gains. The main distinguishing design features between the person trade-off component of the UK SVQ Study and our study were (i) health gains in the former were delineated in terms of QALYs (expressed in terms of age, severity levels and QALY components) whereas health gains in our study were delineated in terms of life years gained; and (ii) the age-related beneficiaries of health gains in the former were grouped into four broad age groups (0-20, 20-40, 40-60, 60-80), whereas in our study they were described at a more granulated level across the age spectrum.

In our study, different stages in the life cycle were represented in two principal ways; first, by 19 ages that increased incrementally by five years between birth and 90 years of age (newborns, 5, 10, 15, 20...90); and, second, by 10 ages that increased incrementally by ten years between birth and 90 years of age (newborns, 10, 20, 30...90). The former generated an age matrix of 171 possible age comparisons (age matrix 1; e.g. 5 v 35, 20 v 65, etc.), whilst

the latter generated an age matrix of 45 possible age comparisons (age matrix 2; e.g. 10 v 30, 10 v 70, etc.). It was considered that a further granulation of the life cycle into yearly intervals would not have generated additional meaningful information. Study participants were asked to imagine themselves as the social decision-maker for a resource-constrained health system. They were each asked to complete eight person trade-off questions that considered pairs of health programmes (A and B) that targeted different age groups. For each person trade-off question, they were initially asked to choose one of two programmes targeted at different age groups of equal size, assuming that (i) the programmes would extend the lives of recipients for a fixed period in full health after which they would die, (ii) the programmes were of equal cost, and (iii) that resources were only available to fund one programme. In keeping with the methods of the UK SVQ Study [10], the initial choice between the two programmes was forced; one of the two programmes had to be selected. Participants were subsequently asked to indicate how many people would have to be treated by the programme targeted at their more preferred age group (relative to a fixed number of people treated by the programme targeted at their less preferred age group) for the two programmes to be of equal social value. A ‘ping-pong’ method that oscillated up and down the scale was used to converge on an indifference point between the more preferred and less preferred age groups. The values of the upper and lower ends in each of the ping pong iterations mirrored those adopted by the person trade-off component of the UK SVQ Study [10], with a maximum of six iterations possible within each person trade-off question. Previous research indicates that the ‘ping-pong’ method may be less prone to anchor point bias compared to the alternative titration variant where the values of one option are steadily varied in the same direction.[33] For each study participant, six of the person trade-off questions were selected at random using computer-generated random numbers from age matrix 1; the fixed period of health gain assumed for these questions was an extra 5 years of

life. A further two person trade-off questions were selected at random using computer-generated random numbers from age matrix 2; the fixed period of health gain assumed for these questions was an extra 10 years of life. An illustrative example of a person trade-off question is provided in On-Line Appendix I.

As a warm up to the person trade-off questions, study participants were also asked to rank ten hypothetical age groups (newborns, 10, 20, 30...90) in the order in which they would choose to give an extra 5 years of life. In addition, study participants described their own health status at the point of completion of the survey using the EuroQol EQ-5D measure [34] and the categorical self-reported health status (very good, good, fair, bad, very bad) and long-standing illness (yes, no) measures routinely incorporated into the annual Health Survey for England . Socio-demographic data provided by the study participants included their age, gender, marital status, ethnicity, employment status, annual income, education status, area of residence, and whether they had children and, if so, how many.

The survey questions were ordered as follows: socio-demographic measures, self-reported health measures, ranking exercises and, finally, person trade-off questions.

2.3 Methodological considerations within research design

The use of the person trade-off technique as a preference elicitation tool for informing health care resource allocation raises a number of methodological issues. We addressed two of these methodological issues in our study. First, the reference number of individuals included in the initial iteration of each person trade-off question may influence responses, i.e. respondents may be concerned about the absolute difference between the numbers of individuals in each age group as well as the relative difference between them. Previous studies have framed person trade-off questions with a reference number of individuals ranging from 1 [35] to

1000 [36]. In order to assess whether our estimated age-related weights for health gains were influenced by the framing of the person trade-off questions, the reference number of individuals in each age group was set at 100 for one half of the study participants ($n=1250$) and at 1000 for the other half of study participants ($n=1250$). Study participants were randomly allocated to alternative referents using computer-generated random numbers. Second, little evidence currently exists on whether the person trade-off technique generates stable responses, given unchanged preferences, over time [37-39]. If the technique is to be promoted as an operational tool that can inform health care resource allocation, evidence is required on its test-retest reliability. In order to assess the stability of the preferences of our study sample, all study participants were asked upon completion of the first survey whether they would complete a second (retest) survey. The first 500 individuals (equivalent to one fifth of the sample) who responded positively to this request were asked to complete identical questions, in the same order and using the same on-line format, approximately one week after the first survey. These individuals did not have recourse to their previous responses. Each participant was compensated to the value of a further £3 for their input into the retest survey.

2.4 Data analysis

A detailed statistical analysis plan was followed. The socio-demographic and health characteristics of the entire study sample and the sample who participated in the test-retest reliability exercise were summarised using descriptive statistics. The results of the ranking exercise were presented in terms of the number and proportion of study participants who gave a particular rank to each of the ten hypothetical age groups (newborns, 10, 20, 30...90). Aggregate rankings were summarised as Borda scores whereby the first through tenth ranked ages were given scores of ten through 1, and the scores summed across the respondents for each age group

[28, 40]. The age group with the highest total Borda score was ranked first, the age group with second highest total Borda score was ranked second, and so forth. In order to assess whether the results of the ranking exercise was influenced by the age and parental status of study participants, total Borda scores were also calculated for three age sub-groups of study participants: those less than 40 years of age ($n=1367$), those aged 40-49 years ($n=791$), and those at least 60 years of age ($n=342$); and for two parental status sub-groups of study participants: parents ($n=1152$) and non-parents ($n=1348$).

There is no single preferred measure for aggregating person trade-off ratios across individuals. Taking the arithmetic mean of individual ratios suffers from an asymmetric property and produces inconsistent results [10]. Following the methodology of the UK SVQ study, we adopted two methods for aggregating person trade-off responses for each age comparison across study participants [10]. First, we adopted the ratio of means method whereby, at an individual level, the preferred age group was assigned a value of 1 and the less preferred age group a value equal to the number of people in the preferred age group divided by the number of people in the less preferred age group. Means of values for each age group were estimated across study participants and the ratio of those means calculated. Second, we adopted the median of individual ratios method whereby we estimated a ratio for each study participant and, subsequently, estimated the median value across study participants. See On-Line Appendix II for worked examples of the ratio of means method and the median of individual ratios method and illustrations of their respective symmetrical properties. Both methods were used to estimate relative age weights for 5 year life extensions for all age comparisons delineated by age matrix 1 and for 10 year life extensions for all age comparisons delineated by age matrix 2.

We used an F-test to test for constancy in the relative social value of life extensions for different ages. We also tested for consistency of values via chaining tests performed on 20 random non-adjacent pairings covered by both age matrices; consistency requires that the value

(v) placed on life extensions covered by consecutive sets of adjacent pairings should be consistent with values placed on life extensions covered by non-adjacent pairings, such that $V_{avb} * V_{bvc} = V_{avc}$ (where, in this context, a, b and c represent alternative ages within the age matrices).

In 40 random age comparisons (20 drawn from age matrix 1 and 20 drawn from age matrix 2), the Student's *t* test procedure was used to compare relative weights for health gains by the reference number of individuals in the person trade-off questions (100 v 1000). Similarly, in 20 random age comparisons, the Student's *t* test procedure (assuming unequal variance) was used to test individual-level mean ratio differences by period of life extension (5 years v 10 years). The latter analyses were restricted to individuals for whom both 5 year and 10 year life extension person trade-off questions were answered for each random age comparison.

For the 500 study participants who undertook the second (retest) survey, the test–retest reliability of person trade-off responses was assessed by calculating the intraclass correlation coefficient (ICC) [41]. ICC values of between 0 and 0.2 are considered to represent poor agreement, between 0.3 and 0.4 fair agreement, between 0.5 and 0.6 moderate agreement, between 0.7 and 0.8 strong agreement, and >0.8 excellent agreement [42]. Separate analyses were conducted for all 500 study participants, 250 participants for whom the reference number of individuals in the person trade-off questions was 100, and 250 participants for whom the reference number was 1000. The F-test was used to test the null hypothesis that the sample ICC is 0.

All analyses were performed using STATA software (StataCorp. 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP). Differences were considered statistically significant if *p*-values were less than 0.05.

3. Results

3.1 Study sample

The survey took an average of 21 minutes to complete (median duration: 15 minutes). The baseline characteristics of the study participants are described in table I. The mean (standard deviation (SD)) age of the study sample was 40.5(14.3) years. The majority of the sample was female (69.0%), married (51.2%), of white ethnic origin (92.1%), in paid employment or self-employed (63.0%) and had a highest educational qualification below degree or higher national certificate or higher national diploma level (60.7%). The mean EQ-5D utility score for the study participants, generated by applying the York A1 tariff to each set of responses to the EQ-5D descriptive system [43], was 0.796 (SD: 0.286). The numbers (%) of study participants who described their own health as very good, good, fair, bad or very bad were 567(22.7%), 1260(50.4%), 562(22.5%), 101(4.0%) and 10(0.4%), respectively, whilst the number (%) of study participants who reported a long-standing illness was 911(36.4%).

3.2 Ranking exercises

Table II shows how the study participants ranked the ten hypothetical age groups in the ranking exercise. Over half (54.4%) of the study sample ranked newborns first in the order in which they chose to give an extra 5 years of life. Broadly speaking first placed rankings decreased monotonically with increasing age of the recipient of the programme, although a higher proportion of study participants gave a first ranking to 90 year olds compared to 80 year olds (1.9% v 0.4%). The total Borda score was higher for 10 year olds than for newborns (20,079 v 19,277), but decreased monotonically thereafter with increasing age of the

recipient, declining to 5592 for 90 olds. The same pattern for the total Borda scores was observed when the study sample was divided into three sub-groups that differed by age and into two sub-groups that differed by parental status, of study participants (figure I).

3.3 Person trade-off exercises

Table III summarises the person trade-off generated relative age weights for 5 year life extensions for each of the 171 possible age comparisons covered by age matrix 1. The left hand number in each cell (signifying the relative weight attached to a life extension for individuals of the age described by the row heading relative to individuals of the age described by the column heading) exceeded 1 in 155 (90.6%) age comparisons for the ratio of means method (values in upper half of each cell) and in 158 (92.4%) age comparisons for the median of individual ratios method (values in lower half of each cell); for these age comparisons, a preference was given to life extensions by the younger age group. A general (although not invariable) pattern of increasing magnitude of left hand numbers as one progresses to the right of the age grid can be observed. This suggests that the relative weight placed on life extensions by the younger of two age groups generally increases as the age difference between the age groups increases.

The person trade-off generated relative age weights for 10 year life extensions for each of the 45 possible age comparisons covered by age matrix 2 are summarised in table IV. As with the relative age weights for 5 year life extensions, a general pattern of preference given to life extensions by the younger of the two age groups in each age comparison emerged; this was the case in 39 (86.7%) age comparisons for the ratio of means method and in 40 (88.9%) age comparisons for the median of individual ratios method. Moreover, a general pattern of

increasing relative weights placed on 10 year life extensions by the younger of the two age groups as the age difference between the age groups increased also emerged.

The F-test rejected the assumption of constancy in the relative social value of life extensions for different ages ($p>0.05$). For the ratio of means method, 30 years of age was set as the referent in table V (online appendix III) for both 5 and 10 year life extensions as it emerged with the highest mean relative age weight across age comparisons in both age matrices; table III where there were 18 possible age comparisons for each age group and table IV where there were 9 possible age comparisons for each age group. Relative weights for life extensions by other age groups can be inferred in this summary table. For example, using the ratio of means methods, a weight of 0.91 can be inferred for 5 year life extensions by 15 year olds relative to the referent. For the median of individual ratios method, 10 years of age was set as the referent for 5 year life extensions as it emerged with the highest mean relative age weight across age comparisons in table III; whilst for 10 year life extensions, 30 years of age was set as the referent as it emerged with the highest mean relative age weight across age comparisons in table IV.

3.4 Methodological tests

The results of the chaining tests performed on 20 random non-adjacent pairings across age matrices 1 and 2 revealed that in 14 cases the values implied by chaining were more extreme than those derived directly (online appendix IV). However, in 12 of the 20 chaining tests performed, the values implied by chaining fell within 10% of those derived directly. In all 40 random age comparisons, the Student's *t* test procedure revealed no significant differences in relative weights for health gains by the reference number of individuals in the person trade-off questions (100 v 1000) (online appendix V). Table VI summarises the individual-level

mean ratio differences by period of life extension (5 years v 10 years) for those individuals who answered both 5 year and 10 year life extension person trade-off questions for the same age groups. No significant differences were observed in 18 of the 20 random age comparisons.

The baseline characteristics of the 500 study participants who undertook the second (retest) survey broadly reflected those of the larger sample. The retest survey took an average of 14 minutes to complete (median duration: 15 minutes). The results of the test-retest reliability exercises are summarised in table VII. The results are further stratified into two sub-groups; the first for whom the reference number of individuals in each age group was set at 100 ($n=250$) and the second for whom it was set at 1000 ($n=250$). For each group, the mean relative weight generated by each ordered person trade-off question increased in value between test and retest. For subjects with a referent of 100 individuals, the ICCs revealed strong agreement (0.7-0.8) between values at test and retest for three of the eight person trade-off questions, and excellent agreement (>0.8) for five of the eight person trade-off questions. For subjects with a referent of 1000 individuals, the ICCs revealed strong agreement for four of the eight person trade-off questions, and excellent agreement for the remaining four. For all test-retest analyses, the F-test rejected the null hypothesis that the sample ICC is 0 ($p<0.001$).

4. Discussion

In this paper, we report the results of a person trade-off (or ‘matching’) study aimed at estimating granulated age-related weights for health gains across the age spectrum that can potentially inform health care decision-making. The results revealed a general, although not invariable, tendency to give more weight to health gains, expressed in terms of life

extensions, to younger age groups. In over 85% of age comparisons, the person trade-off exercises revealed a preference for life extensions by the younger of the two age groups compared. This pattern held regardless of the method of aggregating responses across study participants. Moreover, the relative weight placed on life extensions by the younger of the two age groups was generally, although not invariably, found to increase as the age difference between the comparator age groups increased. Further analyses revealed that the highest mean relative weight placed on life extensions was estimated for 30 year olds when the ratio of means method was used to aggregate responses across study participants. The highest mean relative weight placed on life extensions was estimated for 10 year olds for 5 year life extensions and for 30 year olds for 10 year life extensions, when the median of individual ratios method was used to aggregate responses across study participants.

Comparison of our results with those of other studies are constrained by a number of factors, including considerable diversity of methods across studies, such as the preference elicitation technique, differences in the age profiles of the recipients of health gain, differences in the framing of health gains, and differences in the characteristics of study participants. Revealed preference studies have estimated monetary valuations for the life-saving benefits of interventions, policies and regulations based on individuals' real world choices between alternatives that differ in mortality risk and monetary consequences. A number of these revealed preference studies suggest that the value individuals place on reducing health risks or achieving health gains, often expressed in terms of a monetary value of statistical life, is higher for children than for adults [17-20]. In all of these studies, values were only derived for between two and four age groups, each delineated in broad terms. Moreover, separating out altruistic preferences for the health of one's own children or family members is problematic for many of these revealed preference studies.

Perhaps more pertinent evidence to our research question is provided by other stated preference studies that asked participants faced with a resource-constrained health system to take on the role of citizens rather than consumers. The studies varied considerably in their stated preference elicitation technique, age profiles of the recipients of health gain, framing of health gains, and size and composition of study samples, which limits comparisons between studies. Notably, the number of ages assessed varied between two [29] and five [23, 26-28]. In nine of 14 studies, preference was generally expressed for health gains by the youngest age group assessed [22, 24, 27-31, 44, 45]; in the remaining five studies [10, 23, 25, 26, 46] preference was generally expressed for health gains by the second youngest age group assessed. It should be noted, however, that the results of these stated preference studies generally contrast with the deliberations of NICE's Citizen's Council, which considered that age should not be a criterion in how health care resources are allocated [47].

Our study was designed to address methodological concerns raised by the use of the person trade-off technique as a preference elicitation tool. In the first survey of the full study sample, computer-generated random numbers were used to determine the order of person trade-off questions, thereby minimising the role of order effects in the preference elicitations [48]. Study participants were also randomly allocated to alternative reference number of individuals using computer-generated random numbers in order to account for possible framing effects. Adopting alternative referents in the person trade-off questions had no discernible effects on the study results. Moreover, tests of reliability of preferences revealed either strong or excellent agreement between values at test and retest in a large study sub-sample. The results of the methodological tests performed generate confidence in the results of the person trade-off exercises. Nevertheless, it should be noted that other methodological tests, such as tests of cardinal transitivity [36], remained outside of the scope of our study.

A number of factors, over and above those common to stated preference studies that ask hypothetical questions, need to be considered when interpreting our results. First, we cannot claim that our study sample is representative of the population of the United Kingdom. Our study sample was, for example, predominantly female and reported a lower mean EQ-5D utility score and a higher rate of long-standing illness than reported in the Health Survey for England, the annual cross-sectional national survey commissioned by the Department of Health with the view to monitoring trends in the English population's health. Nevertheless, our study sample was broadly distributed across a range of socio-demographic, geographic and health variables and this is, to our knowledge, the largest preference elicitation study aimed at estimating age-related weights for health gains. Second, and related, the study sample did not include any children, primarily because it was restricted to the membership of the health care consumer panels of the recruiting social research company. One viewpoint is that society does not generally view children as autonomous legal, social and economic agents and, consequently, judges the values of children to be of limited relevance to social decision making. Our own viewpoint is that there is scope for incorporating the values of children with the necessary cognitive competences into social decision making deliberations in this area. Third, the measure of health gain for which we derived age-related weights was expressed in terms of years of life in full health. A valid concern that is that survey respondents could have interpreted the concept of full health differently for different stages of the age spectrum. Moreover, measuring health gains in terms of QALYs would arguably have been more informative to decision-makers. Nevertheless, our pilot research had demonstrated that disaggregating health gains in terms of life extensions and health-related quality of life improvements would have been cognitively challenging given the on-line format of the survey. Fourth, by setting the duration of health gain in the person trade-off questions at either 5 or 10 years, we controlled for 'health maximisation ageism' [28] or 'utilitarian

ageism' [31], the preference for the treatment of younger patients simply because they are expected to live longer after treatment. However, we are unable to assess whether the general preference for health gains for younger age groups over older age groups was primarily driven by 'productivity ageism'[28], where preferences are determined by the recipients' productivity potential, or 'fair innings ageism' [28] (or 'egalitarian ageism' [31]), where preferences are driven by an aversion to inequality in age of death. Disentangling these factors, and indeed other factors, within the preference structures of the study participants would have required detailed qualitative research. Fifth, although our study was based on a large sample of 2500 participants, the relatively small numbers of observations for each age comparison within age matrix 1 (e.g. 5 v 35, 20 v 65, etc.) and age matrix 2 (e.g. 10 v 30, 10 v 70, etc.) limited the potential for regression analyses of the study data. We present in online appendix VI results of logistic regression analyses for ten random age comparisons. The probability of choosing the programme that saves lives among the older individuals was estimated in each logistic regression. The explanatory variables included the number of lives saved among younger individuals, the age of the lives of older individuals saved (entered as a dummy variable), and the age and sex of the respondent. A negative coefficient on the number of lives saved among younger individuals indicates that the programme that saves lives among older individuals is less attractive [44]. The sample sizes for each logistic regression numbered less than 100, but we present this appendix as an exemplar of how data of our type might be further analysed in larger studies.

5. Conclusion

In conclusion, this study has produced new evidence that should inform deliberations on whether health gains should be weighted according to age in health care decision-making. At

this stage, we consider it important that the results should be treated with caution, and subject to re-testing. In addition, future research should extend these analyses by estimating distributional weights for health gains according to combinations of characteristics of potential relevance to decision-makers, for example, age in combination with gender.

Acknowledgements

We are grateful to colleagues in our respective departments for their comments on the paper. Funding from the UK Medical Research Council for this study is gratefully acknowledged. The views herein expressed are those of the authors and not necessarily those of the funding body.

There are no conflicts of interest. SP designed the study, undertook some of the analyses and wrote the paper. NK conducted several of the analyses. AR and RB contributed to the study design and interpretation of results. SP acts as the guarantor.

References

1. Canadian Agency for Drugs and Technology in Health (CADTH). Guidelines for the economic evaluation of health technologies: Canada (3rd ed.). CADTH: Ottawa, Canada 2006.
2. National Institute for Health and Clinical Excellence (NICE). NICE guide to the methods of technology appraisal. NICE: London, UK 2008.
3. Pharmaceutical Benefits Advisory Committee (PBAC). Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (version 4.3). Canberra, Australia 2008.
4. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care*. 1989;5(4):559-75.
5. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ*. 2004;13(5):437-52.
6. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. *BMJ*. 2004;329(7459):224-7.
7. Barbieri M, Drummond M, Willke R, Chancellor J, Jolain B, Towse A. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalizability. *Value Health*. 2005;8(1):10-23.
8. Grosse SD. Assessing cost-effectiveness in healthcare: history of the \$50,000 per QALY threshold. *Expert Rev Pharmacoecon Outcomes Res*. 2008;8(2):165-78.
9. Culyer A, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R et al. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *J Health Serv Res Policy*. 2007;12(1):56-8.

10. Baker R, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G et al. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. *Health Technol Assess.* 2010;14(27):1-162.
11. Dolan P, Edlin R, Tsuchiya A, on behalf of the NICE Social QALY Team. The relative societal value of health gains to different beneficiaries: A summary. *Health Economics and Decision Science Discussion Paper Series No. 08/12.* ScHARR, University of Sheffield: Sheffield, UK 2008.
12. Gyrð-Hansen D. Investigating the social value of health changes. *J Health Econ.* 2004;23(6):1101-16.
13. Shmueli A. Investing in early detection versus intensive treatment for breast cancer: a study of the Israeli public priorities. *Health Expect.* 1999;2(3):179-84.
14. Abasolo I, Tsuchiya A. Exploring social welfare functions and violation of monotonicity: an example from inequalities in health. *J Health Econ.* 2004;23(2):313-29; discussion 32-4.
15. Browning CJ, Thomas SA. Community values and preferences in transplantation organ allocation decisions. *Soc Sci Med.* 2001;52(6):853-61.
16. Ubel PA, Baron J, Asch DA. Social responsibility, personal responsibility, and prognosis in public judgments about transplant allocation. *Bioethics.* 1999;13(1):57-68.
17. Agee M, Crocker T. Smoking parents' valuations of own and children's health. Paper presented at the Association of Environmental and Resource Economists Workshop; June 13-15; Bar Harbor, ME, USA 2001.
18. Blomquist GC, Miller TR, Levy DT. Values of risk reduction implied by motorist use of protection equipment - New evidence from different populations. *J Transp Econ Policy.* 1996;30(1):55-66.
19. Leung J, Guria J. Value of statistical life: Adults versus children. *Accident Anal Prev.* 2006;38(6):1208-17.

20. Nastis S, Crocker T. Pregnant mother's valuation of own and of child health. Paper presented at the US EPA workshop on Valuing Environmental Health Risks to Children; 2003 October 20-21; Washington, D.C., USA.
21. Dolan P, Olsen JA, Menzel P, Richardson J. An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Econ.* 2003;12(7):545-51.
22. Charny MC, Lewis PA, Farrow SC. Choosing who shall not be treated in the NHS. *Soc Sci Med.* 1989;28(12):1331-8.
23. Busschbach JJV, Hessing DJ, Decharro FT. The utility of health at different stages in life - a quantitative approach. *Soc Sci Med.* 1993;37(2):153-8.
24. Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs egalitarianism: an Australian survey of health issues. *Soc Sci Med.* 1995;41(10):1429-37.
25. Rodriguez E, Pinto JL. The social value of health programmes: is age a relevant factor? *Health Econ.* 2000;9(7):611-21.
26. Jelsma J, Shumba D, Kristian H, De Weerd W, De Cock P. Preferences of urban Zimbabweans for health and life lived at different ages. *Bull World Health Organ.* 2002;80(3):204-9.
27. Tsuchiya A. The value of health at different ages. York, UK: Centre for Health Economics, University of York, Discussion paper 184, 2001.
28. Tsuchiya A, Dolan P, Shaw R. Measuring people's preferences regarding ageism in health: some methodological issues and some fresh evidence. *Soc Sci Med.* 2003;57(4):687-96.
29. Eisenberg D, Freed GL, Davis MM, Singer D, Prosser LA. Valuing health at different ages: evidence from a nationally representative survey in the US. *Appl Health Econ Health Policy.* 2011;9(3):149-56.

30. Prosser LA, Payne K, Rusinak D, Shi P, Uyeki T, Messonnier M. Valuing health across the lifespan: health state preferences for seasonal influenza illnesses in patients of different ages. *Value Health*. 2011;14(1):135-43.
31. 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):103-11.
32. Fielding N, Lee RM, Blank G. *Online research methods*. London, UK: Sage; 2008.
33. Brazier J, Ratcliffe, J., Salomon, J., Tsuchiya, A. *Measuring and valuing health benefits for economic evaluation*. Oxford: Oxford University Press; 2007.
34. Brooks R. EuroQol: The current state of play. *Health Policy*. 1996;37(1):53-72.
35. Nord E. The validity of a visual analogue scale in determining social utility weights for health states. *Int J Health Plan M*. 1991;6(3):234-42.
36. Dolan P, Tsuchiya A. The person trade-off method and the transitivity principle: an example from preferences over age weighting. *Health Econ*. 2003;12(6):505-10.
37. Nord E. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy*. 1993;24(3):227-38.
38. Pinto-Prades JL, Abellan-Perpignan JM. Measuring the health of populations: the veil of ignorance approach. *Health Econ*. 2005;14(1):69-82.
39. Robinson S. Test-retest reliability of health state valuation techniques: the time trade off and person trade off. *Health Econ*. 2011;20(11):1379-91.
40. Marchant T. Cardinality and the Borda score. *Eur J Oper Res*. 1998;108(2):464-72.
41. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30-46.
42. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.

43. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. *Health Econ.* 1996;5(2):141-54.
44. Johannesson M, Johannsson PO. Is the valuation of a QALY gained independent of age? Some empirical evidence. *J Health Econ.* 1997;16(5):589-99.
45. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. *Health Econ.* 2000;9(2):137-48.
46. Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs - How the public discounts time and age. *J Risk Uncertainty.* 1994;8(3):243-65.
47. Wailoo A, Tsuchiya A. Briefing paper for methods review workshop on QALY weighting. London, UK 2011.
48. Ubel PA, Richardson J, Baron J. Exploring the role of order effects in person trade-off elicitation. *Health Policy.* 2002;61(2):189-99.

Table I: Baseline characteristics of the study participants (N=2,500)

Variable	Summary statistics
Age (years); Mean(SD)	40.5(14.3)
Gender; N(%)	
Male	775(31.0)
Women	1725(69.0)
Marital status; N(%)	
Single (never married)	878(35.1)
Married, living with husband/wife	1280(51.2)
Civil partner	62(2.5)
Married and separated	66(2.6)
Divorced	167(6.7)
Widowed	47(1.9)
Ethnicity; N(%)	
White	2302(92.1)
Mixed	29(1.2)
Asian or Asian British	81(3.2)
Black or Black British	39(1.6)
Chinese/other/prefer not to say	49(2.0)
Employment; N(%)	
Paid employment or self-employed	1575(63.0)
Unpaid work/unemployed	119(4.8)
Full-time education	124(5.0)
Permanently unable to work	91(3.6)
Retired	301(12.0)
Looking after home or family/other	290(11.6)
Annul income (£); N(%)	
<10,000	270(10.8)
10,000-20,000	509(20.4)
20,000-30,000	531(21.2)
30,000-40,000	430(17.2)
40,000-50,000	255(10.2)
>50,000	281(11.2)
Rather not say	224(9.0)
Education (highest qualification); N(%)	
Degree, teaching/nursing qualification and higher	510(20.4)
HNC/HND or equivalent	257(10.3)
A-level/AS-level/GCSE	1517(60.7)
Other	216(8.6)
Have children; N(%)	
Yes	1152(46.1)
No	1348(53.9)
Number of children; Mean(SD)	1.9(0.93)
EQ-5D utility score: Mean(SD) [‡]	0.796(0.286)
Self-reported general health: N(%)	
Very good	567(22.7)
Good	1260(50.4)

Fair	562(22.5)
Bad	101(4.0)
Very bad	10(0.4)
Long-standing illness: N(%)	
Yes	911(36.4)
No	1589(63.6)
Area of residence: N(%)	
London	274(11.0)
South-East	425(17.0)
South-West	224(9.0)
Eastern	154(6.2)
East Midlands	186(7.4)
West Midlands	191(7.6)
Yorkshire	202(8.1)
North-West	283 (11.3)
North-East	102(4.1)
Wales	147(5.9)
Scotland	266(10.6)
North Ireland	46(1.8)

‡ Estimated using the York A1 tariff set [43].

Table II: Distribution of age rank order assigned by study participants

Age	Rank 1 N(%)	Rank 2 N(%)	Rank 3 N(%)	Rank 4 N(%)	Rank 5 N(%)	Rank 6 N(%)	Rank 7 N(%)	Rank 8 N(%)	Rank 9 N(%)	Rank 10 N(%)	Borda score [‡]
Newborn	1359(54.4)	160(6.4)	102(4.1)	120(4.8)	97(3.9)	142(5.7)	153(6.1)	105(4.2)	110(4.4)	152(6.1)	19277
10	432(17.3)	1210(48.5)	175(7.0)	162(6.5)	179(7.2)	134(5.4)	84(3.4)	35(1.4)	61(2.4)	28(1.1)	20079
20	279(11.2)	390(15.6)	982(39.3)	156(6.2)	147(5.9)	116(4.6)	84(3.4)	92(3.7)	64(2.6)	190(7.6)	17640
30	202(8.1)	291(11.6)	383(15.3)	930(37.2)	80(3.2)	99(4.0)	141(5.6)	97(3.6)	212(8.5)	65(2.6)	16532
40	87(3.5)	160(6.4)	259(10.4)	360(14.4)	1033(41.3)	106(4.2)	114(4.6)	238(9.5)	95(3.8)	48(1.9)	15038
50	43(1.7)	54(2.2)	142(5.7)	233(9.3)	388(15.5)	1200(48.0)	242(9.7)	78(3.1)	59(2.4)	61(2.4)	13392
60	27(1.1)	46(1.8)	100(4.0)	157(6.3)	234(9.4)	490(19.6)	1334(53.4)	55(2.2)	32(1.3)	25(1.0)	12027
70	15(0.6)	36(1.4)	105(4.2)	142(5.7)	226(9.0)	122(4.9)	269(10.8)	1551(62.0)	18(0.7)	16(0.6)	10055
80	9(0.4)	51(2.0)	121(4.8)	196(7.8)	73(2.3)	44(1.8)	59(2.4)	203(8.1)	1732(69.3)	12(0.5)	7868
90	47(1.9)	102(4.1)	131(5.2)	44(1.8)	43(1.7)	47(1.9)	20(0.8)	46(1.8)	117(4.7)	1903(76.1)	5592

Modal rank is presented in bold.

[‡] Aggregate rankings summarised as Borda scores whereby the first through tenth ranked ages were given scores of ten through 1, and the scores summed across the respondents (n=2500) for each age group. Borda scores are calculated by treating ranks as if they were expressed on an interval scale.

Table III: Age Matrix 1: Relative age weights for 5 year life extensions: ratio of means method* and median of individual ratios method**

	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90
0		0.63(1.58) 0.55(1.82)	0.66(1.52) 0.55(1.54)	0.66(1.51) 0.65(1.54)	0.82(1.22) 0.75(1.33)	0.60(1.68) 0.31(3.25)	0.69(1.44) 0.60(1.68)	0.60(1.65) 0.35(2.86)	0.74(1.36) 0.65(1.54)	0.79(1.27) 0.66(1.52)	0.69(1.44) 0.45(2.22)	0.64(1.57) 0.35(2.86)	1.30(0.77) 1.54(0.65)	1.16(0.86) 1.03(0.97)	1.83(0.55) 3.64(0.28)	1.63(0.61) 2.86(0.35)	1.75(0.57) 3.64(0.28)	2.52(0.40) 13.33(0.08)	4.23(0.24) 33.3(0.03)
5			1.01(0.99) 1.00(1.00)	1.15(0.85) 1.01(0.99)	1.05(0.95) 1.01(0.99)	0.97(1.03) 0.99(1.01)	1.08(0.93) 1.05(0.95)	1.10(0.91) 1.01(0.99)	1.23(0.81) 1.54(0.65)	1.12(0.89) 1.21(0.83)	1.07(0.93) 1.20(0.83)	1.30(0.77) 1.54(0.65)	1.61(0.62) 2.86(0.35)	1.91(0.52) 2.86(0.35)	2.73(0.37) 4.44(0.23)	2.73(0.37) 8.00(0.13)	3.37(0.30) 8.00(0.13)	3.12(0.32) 13.3(0.08)	5.60(0.18) 100(0.01)
10				1.18(0.85) 1.01(0.99)	1.20(0.83) 1.12(0.90)	1.08(0.92) 1.01(0.99)	0.98(1.03) 1.01(0.99)	1.16(0.86) 1.14(0.88)	1.29(0.78) 1.54(0.65)	1.28(0.78) 1.54(0.65)	1.45(0.69) 1.54(0.65)	1.41(0.71) 1.54(0.65)	1.64(0.61) 2.86(0.35)	1.77(0.57) 2.86(0.35)	2.21(0.45) 4.44(0.23)	3.63(0.28) 8.00(0.13)	3.71(0.27) 13.3(0.08)	5.05(0.20) 33.3(0.03)	5.73(0.17) 100(0.01)
15					1.05(0.95) 1.01(0.99)	0.93(1.08) 0.99(1.01)	1.14(0.87) 1.21(0.83)	1.14(0.88) 1.09(0.92)	1.09(0.92) 1.03(0.97)	1.19(0.84) 1.33(0.75)	1.47(0.68) 1.54(0.65)	1.89(0.53) 2.50(0.40)	2.13(0.47) 2.50(0.40)	2.18(0.46) 2.86(0.35)	2.46(0.41) 3.64(0.28)	3.33(0.30) 8.00(0.13)	3.49(0.29) 13.3(0.08)	4.52(0.22) 13.3(0.08)	6.25(0.16) 33.3(0.03)
20						0.99(1.01) 1.01(0.99)	1.12(0.89) 1.01(0.99)	1.04(0.96) 1.01(0.99)	1.10(0.91) 1.14(0.88)	1.19(0.84) 1.21(0.83)	1.47(0.68) 1.33(0.75)	1.69(0.59) 1.82(0.55)	1.95(0.51) 2.86(0.35)	1.94(0.52) 2.86(0.35)	2.92(0.34) 3.64(0.28)	4.07(0.25) 13.3(0.08)	3.65(0.27) 8.00(0.13)	3.83(0.26) 8.00(0.13)	5.18(0.19) 33.3(0.03)
25							0.99(1.01) 1.01(0.99)	1.14(0.88) 1.03(0.97)	1.16(0.86) 1.09(0.92)	1.25(0.80) 1.21(0.83)	1.71(0.59) 2.22(0.45)	1.49(0.67) 1.82(0.55)	1.77(0.56) 2.22(0.45)	2.39(0.42) 2.86(0.35)	2.94(0.34) 4.44(0.23)	3.16(0.32) 5.71(0.18)	4.02(0.25) 8.00(0.13)	4.78(0.21) 33.3(0.03)	5.21(0.19) 33.3(0.03)
30								1.08(0.92) 1.03(0.97)	1.23(0.81) 1.09(0.92)	1.31(0.76) 1.21(0.83)	2.04(0.49) 2.22(0.45)	1.60(0.63) 1.82(0.55)	2.27(0.44) 2.86(0.35)	2.21(0.45) 2.86(0.35)	3.71(0.27) 8.00(0.13)	3.70(0.27) 5.71(0.18)	3.80(0.26) 5.71(0.18)	5.02(0.20) 33.3(0.03)	6.88(0.15) 33.3(0.03)
35									1.15(0.87) 1.03(0.97)	1.26(0.79) 1.18(0.85)	1.78(0.56) 2.22(0.45)	1.45(0.69) 1.54(0.65)	2.62(0.38) 2.86(0.35)	2.41(0.41) 2.86(0.35)	3.00(0.33) 5.71(0.18)	3.80(0.26) 5.71(0.18)	3.81(0.26) 8.00(0.13)	4.81(0.21) 13.3(0.08)	5.60(0.18) 13.3(0.08)
40										1.17(0.86) 1.03(0.97)	1.53(0.65) 2.22(0.45)	1.38(0.73) 1.27(0.79)	1.71(0.58) 1.82(0.55)	2.35(0.43) 2.86(0.35)	3.24(0.31) 5.71(0.18)	3.10(0.32) 8.00(0.13)	4.22(0.24) 8.00(0.13)	3.92(0.26) 13.3(0.08)	5.88(0.17) 33.3(0.03)
45											1.29(0.77) 1.14(0.88)	1.50(0.67) 1.33(0.75)	1.84(0.54) 1.82(0.55)	2.04(0.49) 2.86(0.35)	2.86(0.39) 2.86(0.35)	2.96(0.34) 4.44(0.23)	4.18(0.24) 8.00(0.13)	4.50(0.22) 8.00(0.13)	6.11(0.16) 33.3(0.03)
50												1.28(0.78) 1.14(0.88)	1.50(0.67) 1.33(0.75)	1.56(0.64) 1.54(0.65)	2.48(0.40) 2.86(0.35)	2.34(0.43) 3.64(0.28)	4.32(0.23) 10.0(0.10)	4.11(0.24) 8.00(0.13)	6.88(0.15) 100(0.01)
55													1.21(0.82) 1.14(0.88)	1.69(0.59) 1.82(0.55)	2.69(0.37) 3.64(0.28)	2.86(0.35) 3.64(0.28)	6.30(0.16) 13.3(0.08)	4.08(0.24) 8.00(0.13)	5.59(0.18) 33.3(0.03)
60														1.29(0.77) 1.21(0.83)	1.70(0.59) 1.82(0.55)	2.44(0.41) 2.86(0.35)	3.17(0.32) 4.44(0.23)	3.76(0.27) 8.00(0.13)	6.25(0.16) 33.3(0.03)
65															1.63(0.62) 1.54(0.65)	1.93(0.52) 2.22(0.45)	3.09(0.32) 4.44(0.23)	3.48(0.29) 5.71(0.18)	3.96(0.25) 13.3(0.08)
70																1.01(0.52) 1.82(0.55)	1.35(0.74) 2.00(0.50)	2.92(0.34) 3.64(0.28)	0.58(1.71) 13.3(0.08)
75																	1.66(0.60) 1.54(0.65)	2.41(0.41) 2.86(0.35)	3.44(0.29) 8.00(0.13)
80																		1.88(0.53) 1.82(0.55)	2.48(0.40) 2.86(0.35)
85																			2.28(0.44) 2.86(0.35)
90																			

0 denotes newborns. * Results for ratio of means methods are presented in upper half of each cell. ** Results for median of individual ratios method are presented in lower half of each cell.

Left hand number in each cell reports relative weight for life extension by age group described by row headings, whilst right hand number reports relative weight for life extension by age group described by column headings; e.g., when comparing 20 year olds with 40 year olds using the ratio of means method, the relative weight for 5 year life extensions for 20 year olds is 1.10, whilst the relative weight for 5 year life extensions for 40 year olds is 0.91.

Table IV: Age Matrix 2: Relative age weights for 10 year life extensions: ratio of means method* and median of individual ratios method**

	0	10	20	30	40	50	60	70	80	90
0		0.68(1.48) 0.55(1.82)	0.61(1.64) 0.45(2.22)	0.90(1.11) 0.91(1.10)	0.96(1.04) 1.01(0.99)	0.82(1.21) 0.65(1.54)	1.34(0.74) 2.86(0.35)	2.23(0.43) 5.71(0.18)	2.85(0.35) 10.0(0.10)	4.26(0.23) 33.3(0.03)
10			1.05(0.95) 1.01(0.99)	0.95(1.06) 1.00(1.00)	1.28(0.78) 1.33(0.75)	1.59(0.63) 2.22(0.45)	2.40(0.42) 3.64(0.28)	3.52(0.28) 8.00(0.13)	4.93(0.20) 13.33(0.08)	4.23(0.24) 33.3(0.03)
20				1.03(0.97) 1.03(0.97)	1.31(0.77) 1.21(0.83)	1.30(0.77) 1.54(0.65)	2.12(0.47) 2.86(0.35)	3.11(0.32) 5.71(0.18)	4.61(0.22) 13.33(0.08)	7.47(0.13) 33.3(0.03)
30					1.25(0.80) 1.14(0.88)	1.50(0.66) 1.54(0.65)	2.62(0.38) 0.28(3.64)	2.99(0.33) 5.71(0.18)	4.67(0.21) 13.33(0.08)	7.71(0.13) 100(0.01)
40						1.43(0.70) 1.33(0.75)	2.16(0.46) 2.86(0.35)	3.20(0.31) 4.44(0.23)	4.36(0.23) 13.33(0.08)	5.68(0.18) 33.3(0.03)
50							1.63(0.61) 1.82(0.55)	2.89(0.35) 3.64(0.28)	3.02(0.33) 5.71(0.18)	5.67(0.18) 33.3(0.03)
60								1.90(0.53) 2.22(0.45)	3.45(0.29) 5.71(0.18)	4.65(0.21) 33.3(0.03)
70									2.28(0.44) 2.86(0.35)	4.08(0.24) 13.33(0.08)
80										2.90(0.35) 3.64(0.28)
90										

0 denotes newborns. * Results for ratio of means method are presented in upper half of each cell. ** Results for median of individual ratios method are presented in lower half of each cell. Left hand number in each cell reports relative weight for life extension by age group described by row heading, whilst right hand number reports relative weight for life extension by age group described by column heading; e.g., when comparing 20 year olds with 40 year olds using the ratio of means method, the relative weight for 10 year life extensions for 20 year olds is 1.31, whilst the relative weight for 10 year life extensions for 40 year olds is 0.77.

Table V: Relative age weights by period of life extension and analytical method

Age	5 year life extensions		10 year life extensions	
	Ratio of means method [±]	Median of individual ratios method [†]	Ratio of means method [±]	Median of individual ratios method [†]
Newborn	0.54	0.37	0.62	0.45
5	0.81	0.85	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
10	0.89	1.00	0.88	0.52
15	0.91	0.51	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
20	0.88	0.47	0.97	0.49
25	0.91	0.58	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
30	1.00	0.59	1.00	1.00
35	0.95	0.36	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
40	0.90	0.48	0.89	0.47
45	0.90	0.41	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
50	0.86	0.78	0.79	0.40
55	0.89	0.43	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
60	0.78	0.36	0.73	0.39
65	0.70	0.23	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
70	0.54	0.20	0.65	0.20
75	0.60	0.16	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
80	0.57	0.13	0.60	0.13
85	0.56	0.13	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
90	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\

[±] For the ratio of means method for the 5 and 10 year life extensions, 30 years of age was set as the referent as it emerged with the highest mean relative age weight across age comparisons in tables III and IV.

[†] For the median of individual ratios method for 5 year life extensions, 10 years of age was set as the referent as it emerged with the highest mean relative age weight across age comparisons in table III; whilst for 10 year life extensions, 30 years of age was set as the referent as it emerged with the highest implied relative age weight across age comparisons in table IV.

Table VI: Individual-level mean ratio differences by period of life extension (5 v 10 years)

Age group comparison	5 year life extensions	10 year life extensions	<i>p</i>-value*
20 v 40	6.84	2.89	0.16
40 v 50	3.33	3.36	0.99
60 v 80	0.32	0.32	0.99
10 v 60	6.66	4.47	0.48
40 v 80	0.80	2.36	0.31
0 v 40	24.2	11.1	0.02
40 v 70	1.57	2.25	0.71
40 v 60	2.16	1.55	0.75
0 v 80	7.54	4.04	0.31
30 v 40	5.13	2.76	0.39
10 v 70	5.82	0.74	0.03
20 v 30	2.40	3.57	0.54
10 v 40	9.24	4.49	0.16
30 v 70	0.70	1.52	0.47
10 v 20	5.34	10.17	0.18
0 v 30	24.20	15.21	0.10
10 v 50	3.66	5.15	0.49
10 v 90	5.38	4.35	0.75
50 v 90	2.05	1.38	0.75
10 v 80	4.00	0.23	0.06

* Student's *t* test procedure with unequal variance was used to test individual-level mean ratio differences by period of life extension. Analyses restricted to individuals for whom both 5 year and 10 year life extension person trade-off questions were answered.

Table VII: Summary statistics for PTO weights and ICCs for PTO questions: test-retest analysis

PTO question	N	Mean (SD) test	Mean (SD) retest	ICC	p-value*
<i>All subjects (n=500)</i>					
1 st	500	178(276)	184(274)	0.82	<0.001
2 nd	500	201(297)	209(292)	0.88	<0.001
3 rd	500	219(301)	221(297)	0.50	<0.001
4 th	500	217(303)	227(311)	0.90	<0.001
5 th	500	201(292)	212(305)	0.93	<0.001
6 th	500	195(285)	208(297)	0.88	<0.001
7 th	500	162(265)	174(267)	0.84	<0.001
8 th	500	182(278)	184(275)	0.85	<0.001
<i>Subjects with referent of 100 individuals (n=250)</i>					
1 st	250	33(32)	39(35)	0.75	<0.001
2 nd	250	34(34)	38(33)	0.82	<0.001
3 rd	250	39(34)	42(35)	0.79	<0.001
4 th	250	39(35)	42(36)	0.86	<0.001
5 th	250	38(34)	40(34)	0.88	<0.001
6 th	250	36(35)	40(35)	0.86	<0.001
7 th	250	31(32)	33(33)	0.79	<0.001
8 th	250	36(34)	37(34)	0.81	<0.001
<i>Subjects with referent of 1000 individuals (n=250)</i>					
1 st	250	323(331)	330(327)	0.74	<0.001
2 nd	250	369(345)	380(334)	0.80	<0.001
3 rd	250	399(340)	401(334)	0.77	<0.001
4 th	250	395(346)	413(351)	0.83	<0.001
5 th	250	364(342)	384(355)	0.89	<0.001
6 th	250	352(334)	377(344)	0.82	<0.001
7 th	250	294(324)	316(319)	0.77	<0.001
8 th	250	329(333)	331(327)	0.78	<0.001

ICC denotes intraclass correlation coefficient.

* F-test used to test the null hypothesis that the sample ICC is 0.

Figure I: Borda scores for age rankings by age group and parental status of study participants

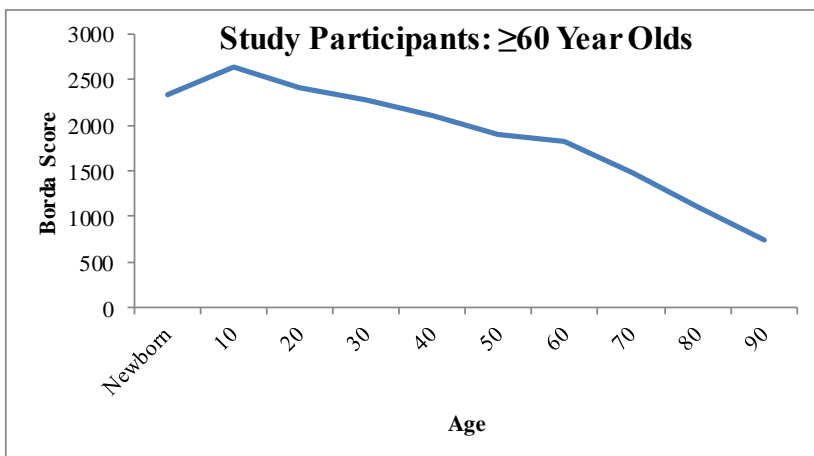
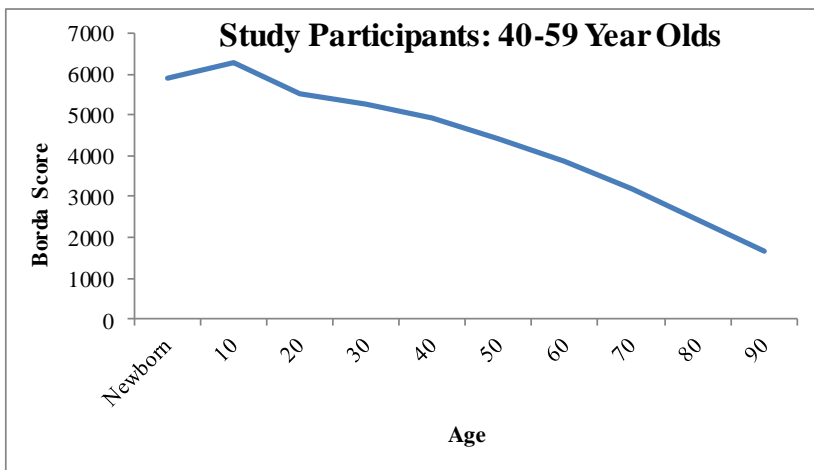
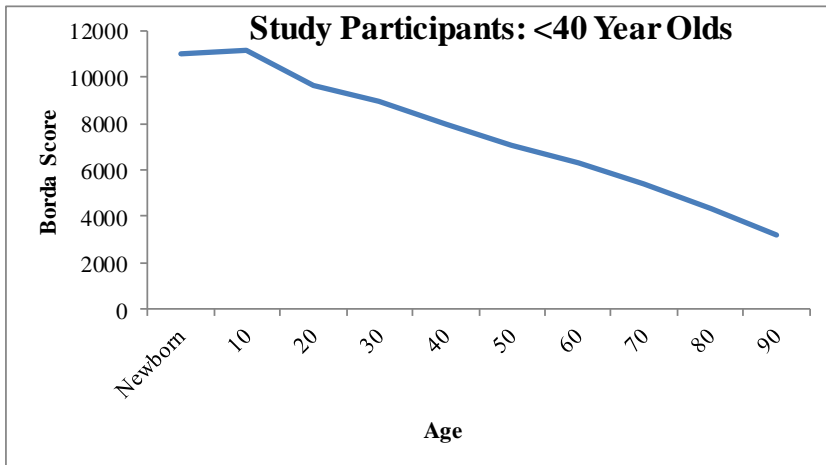
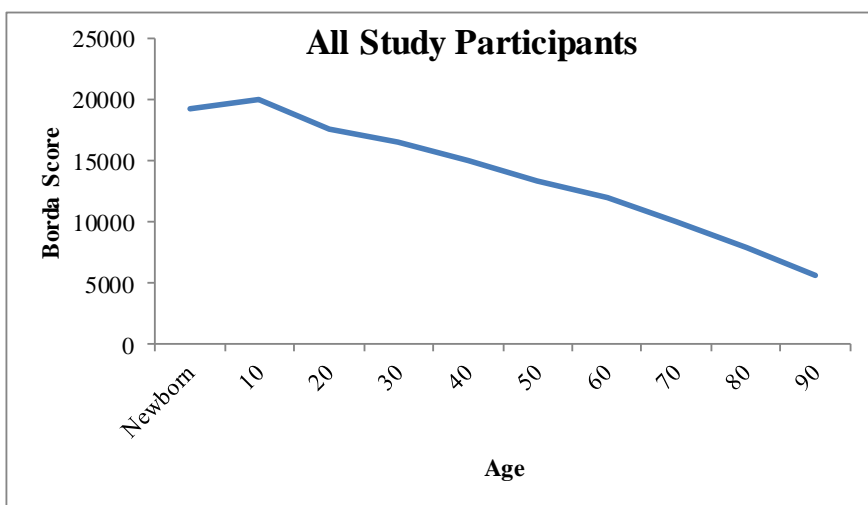
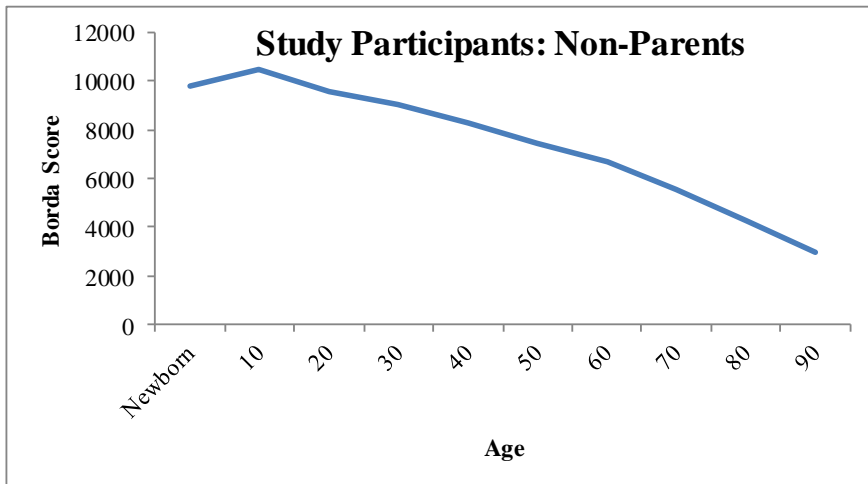
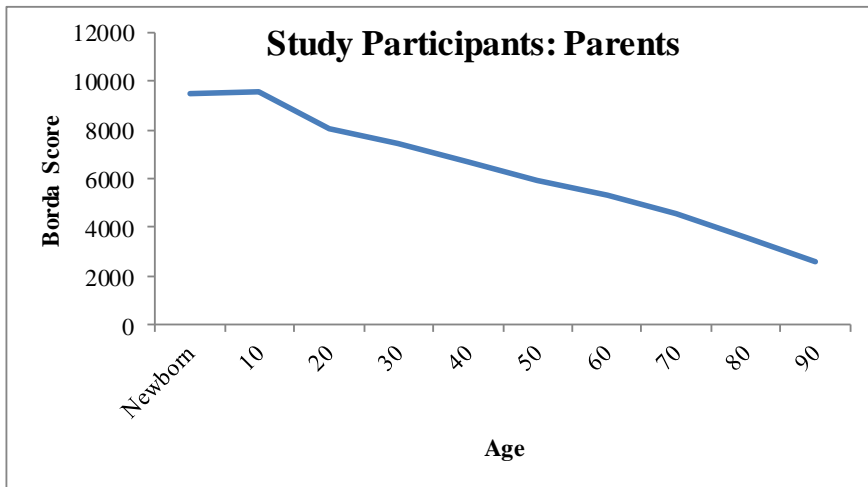


Figure I: Borda scores for age rankings by age group and parental status of study participants (continued)



On-line Appendix I: Example of Person Trade-Off Question

Now we are going to take you through an example of a prioritisation exercise we would like you to complete. Imagine you are in a position where you have to choose between two health care programmes. Both programmes would extend the lives of patients who would otherwise die immediately if they were not treated. Both programmes cost the same amount, but you cannot provide both.

Programme A would be targeted at patients who are 20 years old. It would allow them to live in full health for an extra 5 years after which they die.

Programme B would be targeted at patients who are 55 years old. It would allow them to live in full health for an extra 5 years after which they die.

Now would you give priority to:

EITHER

Programme A which would extend the lives of **100** 20 year olds for an extra 5 years after which they die

OR

Programme B which would extend the lives of **100** 55 year olds for an extra 5 years after which they die

• Programme A

• Programme B

(If A is selected)

What if fewer than 100 20 years old could be treated for that amount of money?

Now would you give priority to:

Programme A which would extend the lives of **60** 20 year olds for an extra 5 years after which they die

OR

Programme B which would extend the lives of **100** 55 year olds for an extra 5 years after which they die

• Programme A

• Programme B

(If B is selected)

What if fewer than 100 55 years old could be treated for that amount of money?

Now would you give priority to:

Programme A which would extend the lives of **100** 20 year olds for an extra 5 years after which they die

OR

Programme B which would extend the lives of **60** 55 year olds for an extra 5 years after which they die

• Programme A

• Programme B

In both cases a ping pong exercise ensued until the point of indifference was reached.

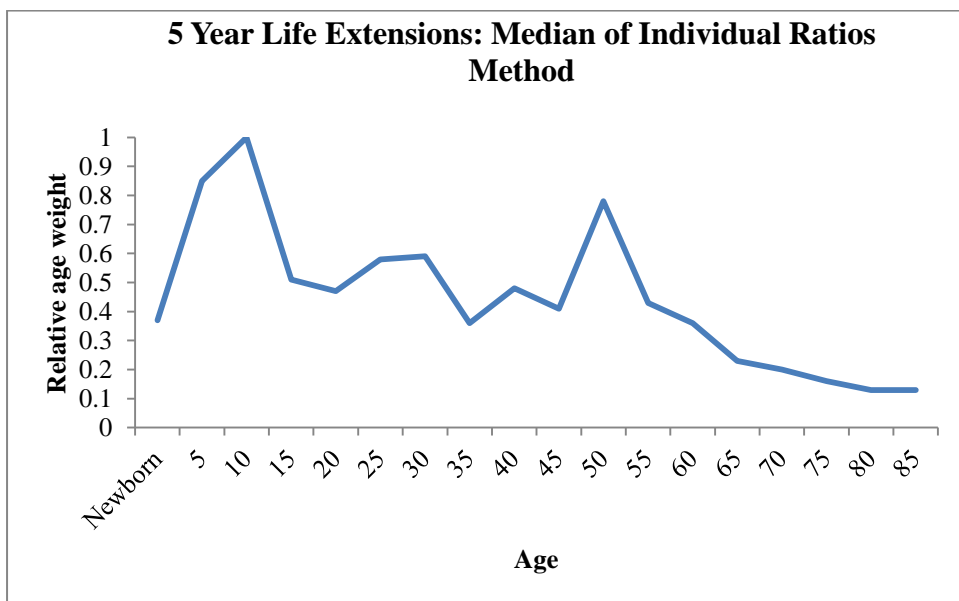
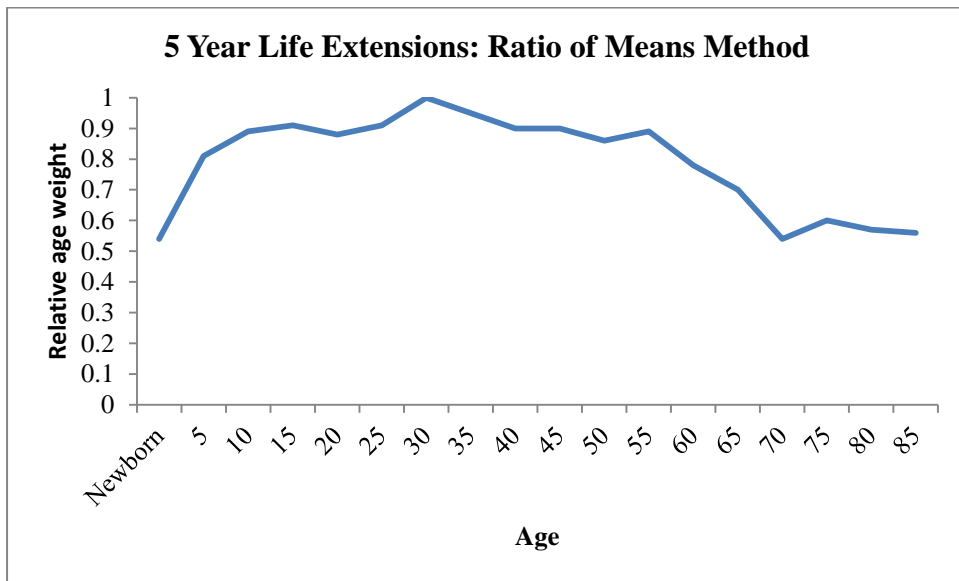
On-line Appendix II: Examples of aggregation of person trade-off ratios

Respondent	No. of individuals at indifference point		Ratio of means method		Median of individual ratios method	
	Age A	Age B	Age A	Age B	Ratio A/B	Ratio B/A
1	100	50	0.5	1	2	0.5
2	100	80	0.8	1	1.25	0.8
3	100	60	0.6	1	1.67	0.6
4	100	50	0.5	1	2	0.5
5	100	80	0.8	1	1.25	0.8
6	50	100	1	0.5	0.5	2
7	90	100	1	0.9	0.9	1.11
8	70	100	1	0.7	0.7	1.43
9	50	100	1	0.5	0.5	2

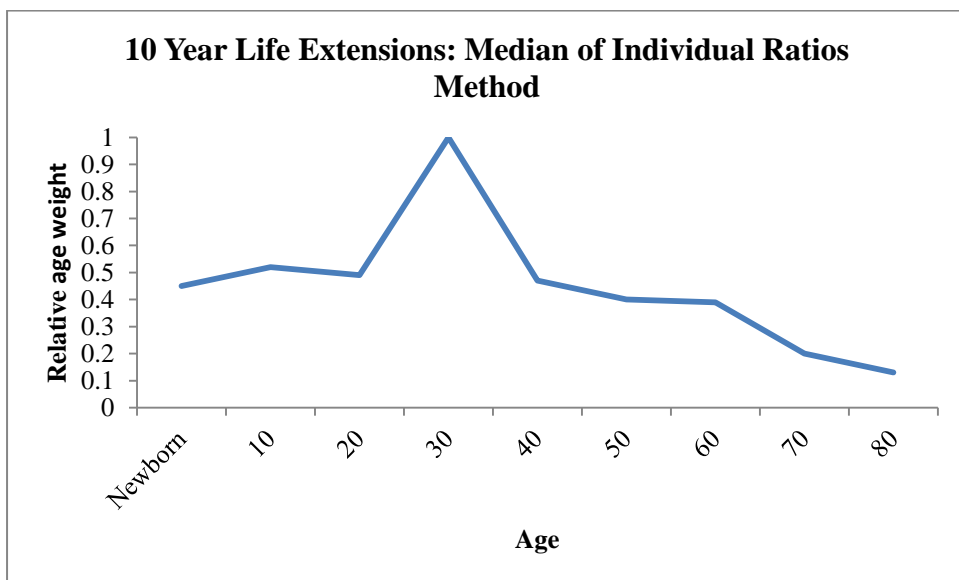
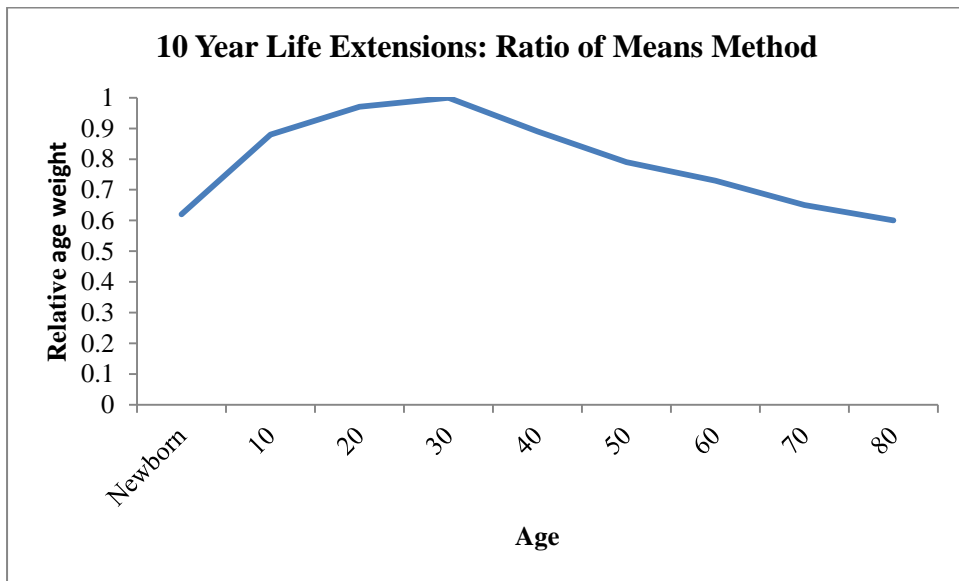
Ratio of means method: mean value for age A is 0.800, whilst the mean value for age B is 0.844. The implied weight of age B relative to age A is $0.844/0.800$, i.e. 1.055. Alternatively, implied weight of age A relative to age B is $0.800/0.844$, i.e. 0.948. The measure is symmetrical, i.e. $1/1.055$ equals 0.948.

Median of individual ratios method: the median of individual ratios for age A relative to age B is 1.25. Alternatively, the median of individual ratios for age B relative to age A is 0.8. The measure is symmetrical, i.e. $1/1.25$ equals 0.8.

On-line Appendix III: Relative age weights by period of life extension and analytical method



On-line Appendix III: Relative age weights by period of life extension and analytical method (continued)



Online Appendix IV: Results of chaining tests for relative age weights

Age group comparison	Direct result	Chained result
0 v 10	0.66	0.64
5 v 15	1.15	1.19
10 v 20	1.20	1.24
25 v 35	1.14	1.07
30 v 50	1.50	1.79
40 v 60	2.16	2.33
50 v 70	2.89	3.10
20 v 40	1.31	1.29
50 v 80	3.02	7.06
45 v 55	1.50	1.65
55 v 65	1.69	1.56
60 v 80	3.45	4.33
0 v 20	0.61	0.71
65 v 75	1.93	1.65
35 v 45	1.26	1.34
70 v 90	4.08	6.53
30 v 40	1.23	1.24
50 v 60	1.50	1.55
65 v 75	1.93	1.65
75 v 85	2.41	3.12

Online Appendix V: Comparison of relative weights for health gains by reference number of individuals (100 v 1000) and period of life extension (5 v 10 years)*

5 year life extension				10 year life extension			
Age group comparison	Referent= 100	Referent= 1000	<i>p</i> -value [‡]	Age group comparison	Referent= 100	Referent= 1000	<i>p</i> -value [‡]
30 v 85	0.26	2.20	0.34	20 v 40	1.78	3.70	0.34
40 v 50	3.18	3.51	0.93	40 v 50	3.57	3.01	0.86
50 v 60	0.77	0.70	0.60	60 v 80	0.29	0.36	0.45
60 v 90	0.17	0.15	0.74	10 v 60	5.54	2.90	0.48
65 v 80	0.33	2.86	0.33	40 v 80	4.26	0.41	0.18
30 v 40	6.34	3.80	0.60	0 v 40	8.51	13.64	0.37
0 v 60	14.20	8.86	0.40	40 v 70	2.35	2.17	0.95
20 v 60	6.78	0.52	0.14	40 v 60	0.43	2.53	0.29
5 v 60	7.41	8.26	0.88	0 v 80	3.98	4.12	0.97
0 v 40	24.39	24.07	0.97	30 v 40	1.50	3.62	0.31
25 v 55	5.69	0.86	0.08	10 v 70	1.04	0.40	0.32
65 v 90	7.60	0.17	0.08	20 v 30	1.74	5.18	0.15
20 v 50	2.20	0.87	0.33	10 v 40	2.45	7.08	0.20
25 v 75	1.84	2.14	0.89	30 v 70	2.71	0.31	0.26
45 v 60	0.53	2.81	0.27	10 v 20	10.70	9.55	0.82
5 v 85	6.93	5.11	0.64	0 v 30	11.51	20.24	0.15
65 v 75	0.71	1.99	0.40	10 v 50	7.70	2.56	0.10
60 v 80	0.37	0.27	0.24	10 v 90	5.38	3.66	0.69
10 v 25	8.51	6.25	0.64	50 v 90	0.87	1.85	0.58
0 v 30	21.28	27.79	0.49	10 v 80	0.22	0.23	0.89

*Relative age weights estimated using the ratio of means method.

[‡] Estimated using the Student's *t* test procedure.

Online Appendix VI: Results of logistic regression analyses of choosing the programme that saves older individuals

20 v 40 Year Olds (Referent=100)*	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 20 year olds	-0.0938	0.3409	0.783	-0.7620	0.5742
Age of older individuals saved	0.4838	0.8898	0.587	-1.2602	2.2280
Age of respondent	0.0244	0.0239	0.307	-0.0224	0.0713
Male gender	-1.1222	0.9686	0.247	-3.0208	0.7762
Constant	-0.9158	1.4512	0.528	-3.7602	1.9285

* The reference number of individuals in each person trade-off question was 100.

20 v 40 Year Olds (Referent=1000)*	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 20 year olds	-0.9246	0.7600	0.224	-2.4143	0.5651
Age of older individuals saved	1.4702	3.1777	0.644	-4.7580	7.6985
Age of respondent	0.0283	0.0271	0.297	-0.0248	0.0815
Male gender	0.4267	0.8341	0.609	-1.2081	2.0615
Constant	1.7835	2.3067	0.439	-2.7377	6.3047

* The reference number of individuals in each person trade-off question was 1000.

40 v 50 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 40 year olds	-0.2568	0.4633	0.579	-1.1649	0.6512
Age of older individuals saved	0.9008	1.4168	0.525	-1.8762	3.6778
Age of respondent	0.03713	0.0323	0.250	-0.0261	0.1004
Male gender	-0.4016	1.0296	0.697	-2.4196	1.6164
Constant	-3.0038	2.1633	0.165	-7.2438	1.2361

40 v 50 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 40 year olds	1.6121	1.3603	0.236	-1.0541	4.2784
Age of older individuals saved	-6.7438	5.3987	0.212	-17.3251	3.8373
Age of respondent	0.0443	0.0344	0.198	-0.0232	0.1118
Male gender	1.4974	1.3918	0.282	-1.2305	4.2255
Constant	-8.9435	4.7490	0.060	-18.2515	0.3643

10 v 60 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	0.3667	0.3434	0.286	-0.3063	1.0398
Age of older individuals saved	-1.1581	1.3748	0.400	-3.8527	1.5365
Age of respondent	0.0598	0.0305	0.050	0.0000	0.1198
Male gender	0.9962	0.8331	0.232	-0.6366	2.6291
Constant	-4.4146	1.7867	0.013	-7.9167	-0.9125

10 v 60 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	0.4043	0.6845	0.555	-0.9372	1.7459
Age of older individuals saved	-0.1650	2.2008	0.940	-4.4786	4.1485
Age of respondent	0.1341	0.0531	0.012	0.0300	0.2382
Male gender	-0.6914	1.0137	0.495	-2.6783	1.2954
Constant	-8.0276	3.5913	0.025	-15.0666	-0.9886

0 v 40 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among newborns	-0.5613	0.3361	0.095	-1.2203	0.0975
Age of older individuals saved	0.8232	1.2764	0.519	-1.6786	3.3250
Age of respondent	0.1114	0.0448	0.013	0.0234	0.1993
Male gender	-0.6892	1.0581	0.515	-2.7632	1.3846
Constant	-2.3366	1.5251	0.126	-5.3259	0.6526

0 v 40 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among newborns	0.6885	0.6898	0.318	-0.6634	2.0405
Age of older individuals saved	-2.1709	2.6892	0.420	-7.4418	3.0998
Age of respondent	0.0852	0.0411	0.038	0.0045	0.1659
Male gender	0.4758	0.8706	0.585	-1.2306	2.1823
Constant	-4.7010	2.4417	0.585	-9.4867	0.0846

0 v 80 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among newborns	1.1815	0.5313	0.026	0.1400	2.2230
Age of older individuals saved	-3.3821	2.0313	0.096	-7.3634	0.5992
Age of respondent	0.1003	0.0482	0.037	0.0058	0.1948
Male gender	0.7069	1.0845	0.515	-1.4187	2.8326
Constant	-6.9125	2.8835	0.017	-12.5641	-1.2608

0 v 80 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among newborns	1.6157	0.9839	0.101	-0.3128	3.5443
Age of older individuals saved	-4.5825	3.9543	0.247	-12.3329	3.1679
Age of respondent	0.0012	0.0375	0.973	-0.0723	0.0748
Male gender	0.4702	0.0375	0.732	-2.2198	3.1604
Constant	-5.9782	2.8195	0.034	-11.5045	-0.4520

30 v 40 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 30 year olds	-0.4319	0.4433	0.330	-1.3008	0.4370
Age of older individuals saved	0.7434	1.4366	0.605	-2.0722	3.5591
Age of respondent	0.0772	0.0367	0.035	0.0052	0.1491
Male gender	-0.7388	1.041	0.478	-2.7803	1.3025
Constant	-3.3175	2.2833	0.146	-7.7927	1.1576

30 v 40 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 30 year olds	0.7684	1.2344	0.534	-1.6509	3.1878
Age of older individuals saved	-2.9125	5.3616	0.587	-13.4211	7.5960
Age of respondent	-0.0382	0.0348	0.272	-0.1065	0.0300
Male gender	2.2472	1.0959	0.040	0.0992	4.3952
Constant	-2.5882	3.5426	0.465	-9.5318	4.3552

10 v 70 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	0.2213	0.4017	0.582	-0.5661	1.0088
Age of older individuals saved	0.4814	1.3174	0.715	-2.1006	3.0635
Age of respondent	0.0236	0.0393	0.547	-0.0533	0.1006
Male gender	0.8762	0.9055	0.333	-0.8986	2.651
Constant	-3.4344	2.0027	0.086	-7.3597	0.4907

10 v 70 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	1.0557	0.5740	0.066	-0.0693	2.1808
Age of older individuals saved	-2.4761	2.0893	0.236	-6.5711	1.6189
Age of respondent	0.0668	0.0315	0.034	0.0050	0.1286
Male gender	-1.1570	0.9772	0.236	-3.0723	0.7582
Constant	-7.0172	2.4879	0.005	-11.8936	-2.1408

10 v 40 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	-0.2904	0.3015	0.336	-0.8814	-3.2546
Age of older individuals saved	0.4886	0.9505	0.607	-1.3744	2.3517
Age of respondent	0.0169	0.0336	0.614	-0.0489	0.0828
Male gender	-0.2918	0.7513	0.698	-1.7644	1.1807
Constant	-0.4165	1.4480	0.774	-3.2546	2.4216

10 v 40 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	-0.4964	0.5354	0.354	-1.5459	0.5530
Age of older individuals saved	2.1301	1.9508	0.275	-1.6935	5.9538
Age of respondent	0.0376	0.0251	0.134	-0.0116	0.0870
Male gender	-0.0386	0.7318	0.958	-1.4730	1.3958
Constant	-1.1942	1.8854	0.526	-4.8897	2.5011

10 v 20 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	-0.2545	0.4093	0.534	-1.0568	0.5476
Age of older individuals saved	2.1639	1.3725	0.115	-0.5261	4.8540
Age of respondent	-0.0228	0.0324	0.482	-0.0865	0.0408
Male gender	1.1692	1.0570	0.269	-0.9024	3.2409
Constant	0.3858	1.9252	0.841	-3.3874	4.1592

10 v 20 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among 10 year olds	-1.0540	0.7359	0.152	-2.4964	0.3884
Age of older individuals saved	4.8290	2.9214	0.098	-0.8968	10.5549
Age of respondent	0.0531	0.0363	0.143	-0.0180	0.1244
Male gender	0.6609	0.9230	0.474	-1.1481	2.4701
Constant	-0.7444	2.4959	0.765	-5.6364	4.1474

0 v 30 Year Olds (Referent=100)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among newborns	-0.0346	0.2523	0.891	-0.5292	0.4599
Age of older individuals saved	0.9413	1.0114	0.352	-1.0410	2.9238
Age of respondent	0.0613	0.0318	0.054	-0.0010	.12384
Male gender	-0.3239	0.7107	0.649	-1.7169	1.0691
Constant	-1.4830	1.3883	0.285	-4.2042	1.2381

0 v 30 Year Olds (Referent=1000)	Coefficient	Standard error	P-value	95% Confidence interval	
In lives saved among newborns	-0.8638	0.7475	0.248	-2.3290	0.6013
Age of older individuals saved	2.6270	2.7908	0.347	-2.8429	8.0970
Age of respondent	-0.0138	0.0274	0.613	-0.0676	0.0398
Male gender	-0.6521	0.7619	0.392	-2.1454	0.8412
Constant	3.5883	2.3823	0.132	-1.0809	8.2577