Research Paper 12/04

Valuing Health at the End of Life:

A Stated Preference

Discrete Choice Experiment

Koonal Shah, Aki Tsuchiya,

Arne Risa Hole, Allan Wailoo

December 2012

A revised version of this paper has been published in Social Science & Medicine and is available for free download via Open Access.

The paper can be downloaded from: http://dx.doi.org/10.1016/j.socscimed.2014.11.022

Please cite as Shah, K.K, Tsuchiya, A. and Wailoo, A., 2015. Valuing health at the end of life: A stated preference discrete choice experiment. Social Science & Medicine, 124, pp.48-56.

Valuing Health at the End of Life:

A Stated Preference Discrete Choice Experiment

Koonal Shah[1,2], Aki Tsuchiya[2,3] , Arne Risa Hole[3] and Allan Wailoo[2]

[1]Office of Health Economics, London

[2]School of Health and Related Research, University of Sheffield

[3]Department of Economics, University of Sheffield

This OHE *Research Paper* also has been published as a NICE Decision Support Unit report. **Please cite as** Shah, K., Tsuchiya, A., Risa Hole, A. and Wailoo, A. (2012) *Valuing health at the end of life: A stated preference discrete choice experiment.* Report. Sheffield: NICE Decision Support Unit.

OHE Research Paper 12/04

December 2012

For further information, please contact

Koonal Shah

KShah@ohe.org

Office of Health Economics
Southside, 7th Floor
105 Victoria Street
London SW1E 6QT
United Kingdom
Tel: +44 20 7747 8850

©Office of Health Economics

About OHE Research Papers

OHE Research Papers are intended to provide information on and encourage discussion about a topic in advance of formal publication. Any views expressed are those of the authors and do not necessarily reflect the views or approval of the Office of Health Economics, its Editorial Board, or its sponsors. Once a version of the Research Paper's content is published in a peer-reviewed journal, that supersedes the Research Paper and readers are invited to cite the published version in preference to the original version.

About the Decision Support Unit

The Decision Support Unit (DSU) is a collaboration between the Universities of Sheffield, York and Leicester. We also have members at the University of Bristol, London School of Hygiene and Tropical Medicine and Brunel University. The DSU is commissioned by The National Institute for Health and Clinical Excellence (NICE) to provide a research and training resource to support the Institute's Technology Appraisal Programme. Please see the website for further information: www.nicedsu.org.uk

Acknowledgements

We are grateful for the contributions of John Brazier, Simon Dixon, Jenny Dunn, Clare Gardiner, Karen Gerard, Julia Honea, Anju Keetharuth, Rachel Knowles, Donna Rowen, Rob White and to all of the respondents who took part in the study.

CONTENTS

TABLES AND FIGURES

SUMMARY

- In 2009, NICE issued supplementary advice to its Appraisal Committees to be taken into account when appraising life extending, 'end of life' treatments. The advice indicates that under certain circumstances it may be appropriate to recommend the use of such treatments even if their base case cost effectiveness estimates exceed the range normally considered acceptable.

- However, the consultation carried out by NICE revealed concerns that there is little evidence to support the premise that society is prepared to fund end of life treatments that would not meet the cost effectiveness criteria used for other treatments. The study described here seeks to address this gap in the evidence.

- A discrete choice experiment (DCE) was used to elicit the preferences of a sample of members of the general public in England and Wales over a range of priority setting scenarios. Each choice task involved asking respondents which of two hypothetical patients they thought should be treated, assuming the health service has enough funds to treat one but not both of them. The patients were described in terms of their life expectancy and quality of life without treatment, and the life expectancy and quality of life gains achievable from treating them.

- In addition, the survey included two 'extension tasks' designed to examine the extent to which respondents' priority setting choices are influenced by information about how long the patients have known about their illness.

- The DCE was carried out using a web-based survey. A total of 3,969 respondents successfully completed the survey, each completing 10 DCE tasks plus two extension tasks. The sample is representative of the general population in terms of age and gender, and covers a range of social grades.

- The conditional logit model was used for modelling. The best fitting model analysed main effects plus three interactions: (i) life expectancy without treatment against life expectancy gain; (ii) life expectancy without treatment against quality of life gain; and (iii) life expectancy gain against quality of life gain.

- Using the model results, utility scores were calculated for all of the 110 possible profiles (combinations of attribute levels) in the full factorial design, as well as the predicted probability of choosing each profile from the full set of profiles. The highest ranked profiles (those with the greatest probability of being chosen) were those that involved substantial life expectancy and quality of life gains from treatment. There is a clear positive relationship between the size of the QALY gains from treatment in a given profile and the predicted probability of that profile being chosen. By comparison, whilst there is little evidence to suggest that profiles involving shorter life expectancy without treatment are more likely to be chosen than those involving longer life expectancy without treatment – the observed patterns are noisy.

- Overall, the results indicate that choices about which patient to treat are influenced more by the sizes of the health gains achievable from treatment than by patients' life expectancy or quality of life in absence of treatment. The extent to which patients are at their end of life does not appear to be the driving factor, although it should be noted that all of the scenarios in this study involve relatively poor prognoses (across all profiles, the patient who is 'best off' without treatment would still die within five years).

- Some respondents appear to support a QALY-maximisation type objective throughout, whilst a small minority always seek to treat those who are worse off without treatment. The majority of respondents, however, seem to advocate a mixture of the two approaches.

- Overall, the results call into question whether a policy of giving higher priority to end of life treatments than to other types of treatments is supported by the public, particularly if the health gains offered by the treatments being 'de-prioritised' are larger than those offered by the end of life treatments. The results also suggest that the focus on life extensions and absence of quality of life improvements in the current NICE end of life criteria may be consistent with public preferences.

- Results from the extension tasks show that including information about the amount of time patients have known about their prognosis has a clear impact on preferences. All other thing being equal, respondents are less likely to choose to treat a patient if they have known about their illness for two years than if they have only just found out about their illness. Further investigation of this factor is recommended.

INTRODUCTION

In January 2009, NICE issued supplementary advice to its Appraisal Committees to be taken into account when appraising life-extending, "end of life" treatments (NICE, 2009a). This advice constitutes an explicit departure from the reference case position that all equal-sized health gains are of equal social value, regardless of to whom they accrue and the context in which they are enjoyed (NICE, 2008). It indicates that if certain criteria are met, it may be appropriate to recommend the use of treatments for terminal illness that offer an extension to life even if their base case cost effectiveness estimates exceed the range normally considered acceptable (Rawlins and Culyer, 2004).

Some aspects of the supplementary advice were revised following a five week public consultation exercise (NICE, 2009b). The current criteria (NICE, 2009c) are reproduced below; if met, the Appraisal Committee is asked to consider the impact of giving greater weight to the health gains achieved in the later stages of disease.

1. The treatment is indicated for patients with a short life expectancy, normally less than 24 months
2. There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional three months, compared to current NHS treatment; and
3. The treatment is licensed, or otherwise indicated, for small patient populations

One way of understanding whether such a policy is appropriate and acceptable for society is to establish whether it is consistent with the preferences of members of the general public. In line with the NHS's policy objective of ensuring public involvement in health care priority setting activities (DH, 1997 and 2001), NICE's position on social value judgements is that "advice from NICE to the NHS should embody values that are generally held by the population of the NHS"(Rawlins and Culyer, 2004). Empirical studies of public preferences can provide meaningful information about these values as long as the methods used are scientifically defensible (Ryan, 2001). Richardson and McKie (2005), amongst others, have

argued that such studies should form part of an "empirical ethics" approach to allocating health care resources.

However, the consultation revealed concerns that there is little evidence to support the premise that society is prepared to fund life-extending end of life treatments that would not meet the cost effectiveness criteria used for other treatments (NICE, 2009b). NICE has acknowledged a need for further exploration of the issues. A recent review undertaken by the Department of Health (2010) also notes that "there is currently no robust evidence in the literature to support a particular magnitude of weighting" of health gains accruing to patients who are severely ill or at the end of life.

To this end, we conducted a small scale simple choice study (hereafter referred to as the "preference study") in mid-2011 with the aim of examining whether the policy of giving higher priority to life-extending end of life treatments (as specified by NICE) than to other types of treatment is consistent with the preferences of members of the general public. The results provide some weak evidence of public support for giving priority to patients with shorter remaining life expectancy, although we note that a sizeable minority of respondents expressed the opposite preference. The results also suggest that the current NICE policy may be insufficient in that it does not cover quality of life-improving end of life treatments, and is not concerned with whether the treatments under appraisal are indicated for patients whose disease progression has been sudden. Both of these factors appear to be influential in determining people's preferences regarding end of life treatments.

Shah et al (2012) provide a detailed report of the preference study described above. The paper concludes by recommending that a larger scale study is conducted to investigate people's preferences regarding end of life treatments more robustly. This report presents findings from that larger scale study.

METHODS

Using Discrete Choice Experiments to Explore Social Values

Discrete choice experiments (DCEs) are a stated preference technique which elicit people's preferences based on their stated preferences in hypothetical choices (Louviere, Hensher and Swait, 2000). DCEs are typically implemented in surveys comprising several "choice sets", each containing competing alternative "profiles" described using defined "attributes" and a range of attribute "levels". Respondents are asked to choose between these alternative profiles, and the resulting choices are analysed to estimate the contribution of each of the attributes to overall utility (Lancsar and Louviere, 2008).

Response data in DCEs are modelled within a random utility framework, which assumes that the utility (U_{nj}) that respondent *n* obtains from choosing alternative *j* can be separated into an explainable component (V_{nj}) and an unexplainable component (ε_{nj}):

$$U_{nj} = V_{nj} + \varepsilon_{nj}$$

The researcher does not observe ε_{nj} and treats the term as random. V_{nj} is the indirect utility function in which the attributes of the alternatives are arguments. The probability that the respondent chooses alternative *i* over alternative *j* is given by:

$$\begin{aligned} P_{ni} &= \Pr\left(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}\right) \quad \forall j \neq i \\ &= \Pr\left(\varepsilon_{ni} - \varepsilon_{nj} < V_{ni} - V_{nj}\right) \quad \forall j \neq i \end{aligned}$$

Assuming that the random terms are independently and identically distributed, the conditional logit model can be used to derive probability outcomes across a choice set(Louviere, Hensher and Swait, 2000). The predicted probability of alternative *i* being chosen from the complete set of alternatives (*j*=1,…,J) is given by:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} \quad j = 1, \ldots, \mathrm{J}$$

The number of studies using DCEs in health economics has grown rapidly in recent years (Ryan, Gerard and Amaya-Amaya, 2008), although most applications have been concerned with eliciting individual personal preferences from respondents who have been asked to

consider the choice context as it applies to themselves (Green and Gerard, 2008). An increasing trend, however, is to use DCEs to examine social preferences whereby respondents are asked to consider choices involving other people in society (see Green for a review of 11 such social preference DCEs (Green, 2007)). This is the context adopted in the study presented in this report.

Attributes and Levels

Table 1 lists the attributes and levels used in the study. "Life expectancy without treatment" and "life expectancy gain from treatment" were included as attributes as these form the basis for criteria 1 and 2 in the current NICE policy. The levels for these attributes were selected so as to examine whether there is a case for amending the cut-offs implied by the existing criteria. For life expectancy without treatment, the current cut-off of 24 months was included as one of the levels. In addition, two levels smaller and two levels larger than this cut-off (three months, 12 months; 36 months, 60 months) were included. An even larger level of 120 months (or 10 years) was considered but omitted due to concerns about displaying this amount of life using the computer-based diagrams (see below). Similarly, the current "life expectancy gain from treatment" cut-off of three months was included, as well as two smaller and two larger levels (1 month, 2 months; 6 months, 12 months). In addition, 0 months was included in order to examine preferences for end of life treatments that offer no life extension.

Table 1. Attributes and levels used in the study

Attribute	Unit	Levels
Life expectancy without treatment	months	3, 12, 24, 36, 60
Quality of life without treatment	%	50, 100
Life expectancy gain from treatment	months	0, 1, 2, 3, 6, 12
Quality of life gain from treatment	%	0, 25, 50

The inclusion of quality of life attributes was driven by the finding in the preference study that the majority of respondents appeared to favour the prioritisation of quality of life-improving treatments over life-extending treatments. The term "quality of life" was not

used in the survey itself; following the preference study and the approach adopted in the social weightings study being conducted by the Department of Health EEPRU (currently in progress), we described this attribute using a health scale ranging from “dead” (0%) and “full health” (100%).

Whilst other studies have presented quality of life using a large range of levels – see Baker et al. (2010), for example – we felt that it may be challenging for respondents to understand the concept of quality of life / health when described using percentage weights. We therefore included only two levels for the “quality of life without treatment” attribute: 50% and 100%. The concept of “50% health” was explained in the instructions as follows: “Suppose there is a health state which involves some health problems. If patients tell us that being in this health state for 2 years is equally desirable as being in full health for 1 year, then we would describe someone in this health state as being in 50% health.”

The three levels for the “quality of life gain from treatment” attribute were designed to represent treatments that: (i) offers no health improvement (0% gain); (ii) restores the patient from 50% health to full health (50% gain); and (iii) offers some improvement from 50% health but does not restore the patient to full health (25% gain).

Other potential attributes, such as the patient’s age or past health, were considered but eventually omitted from the final study design in order to restrict the complexity of the choice tasks. Whilst the literature is inconclusive with regard to the number of attributes that should be included in discrete choice experiments, some researchers have suggested when the tasks become too complex respondents may not make trade-offs but instead adopt other decision heuristics or lexicographic decision rules (Witt, Scott and Osborne, 2009). We therefore chose to focus on the attributes that are most salient to the policy context for NICE.

On the other hand, the results of the preference study suggested that people’s preferences regarding end of life treatments may be guided by how long they have known about their illness / prognosis (i.e. a patient who has only just found out about their illness may be prioritised differently from one who has known about their illness for some time, even if

both patients' prognoses are similar). This is something that we wished to explore further in this study. However, due to the complexities involved in incorporating a "time with knowledge of illness" attribute into the experimental design, we made a pragmatic decision to restrict attributes in the DCE tasks to those listed in Table 1, and to add two further "extension" pairwise choice tasks to the survey which focus specifically on the impact of this additional attribute. These extension tasks do not form part of the experimental design for the DCE but were designed so as to enable direct comparisons with the corresponding "standard" tasks.

Experimental Design

A full factorial design using the attributes and levels listed in Table 1 would have resulted in 5*2*6*3 = 180 possible profiles. Some combinations of levels on these attributes would result in implausible scenarios. The sum of quality of life without treatment and quality of life gain from treatment cannot exceed 100% as it is not possible to have a health state that is better than full health. We also imposed a constraint that the sum of life expectancy gain from treatment and quality of life gain from treatment must be greater than zero, or else the treatment would offer no improvement. After imposing these constraints, 70 of the 180 possible profiles were suppressed, leaving 110 profiles. This means that there were 5,995 possible pairwise choices sets to select from.

Using the STATA software package, 80 pairwise choice sets were constructed from these 110 profiles using a D-optimality algorithm (Carlsson and Martinsson, 2003) with the attribute coefficients set to zero. The design allowed for both main effects and selected interaction effects. All of the choice sets were checked for plausibility, and no manual alteration of the design was required.

The choice sets were organised into 13 different "choice types" according to the nature of the choice being depicted. For example, in 10 of the 80 choice sets, the patient with shorter life expectancy without treatment gains more quality of life from treatment than the patient with longer life expectancy. Similarly, in 11 of the 80 choice sets, both patients have the same amount of life expectancy and quality of life without treatment, but one patient gains

more life expectancy and more quality of life from treatment than the other. If we assume that (all else equal) larger health gains should always be preferred to smaller health gains (an assumption that is inherent to a QALY-maximisation approach to resource allocation), then choosing the patient who gains more life expectancy and quality of life from treatment can be regarded as the dominant option and should always be preferred. These choice sets therefore provide an opportunity to test whether respondents' preferences conform to this type of monotonicity (we might consider a large proportion of respondents failing to choose the dominant option to be a sign of poor data quality).

There is little guidance in the literature on the optimal number of DCE tasks to ask each respondent to complete in a single survey. The social preference DCE studies reviewed by Green (2007) used between one and 18 choice sets per respondent. We opted to organise the 80 choice sets into eight blocks of 10 choices. We sought to achieve a balance of choice types across the blocks. For example, all of the blocks contained at least one (but no more than two) choice sets in which both patients have the same amount of life expectancy and quality of life without treatment, but one patient gains more life expectancy and more quality of life from treatment than the other. Apart from this manual distribution of choice types, the choice sets were assigned to blocks at random.

When asked to choose between multiple options laid out next to each other, it is possible that a "left-hand-side" bias may exist if respondents (subconsciously or otherwise) treat the option on the left as the default choice (Dolan and Tsuchiya, 2011). Similarly, a "top-to-bottom" bias may exist when options are laid out one on top of the other (Spalek and Hammad, 2005). To control for this type of bias, eight "mirror" blocks were generated to match the eight blocks described above. These mirror blocks consisted of the same 10 choice sets but switched the labels assigned to the two alternatives – i.e. the alternative labelled as "patient A" in the original block choice set is labelled as "patient B" in the corresponding mirror block choice set (and vice versa). Including these mirror blocks meant that there were a total of 16 different versions of the survey.

Extension Tasks

As mentioned above, we included extension choice sets at the end of each block to examine whether respondents' choices are influenced by information about how long the patients have known about their illness. Each extension choice set replicated the scenario depicted in one of the DCE choice sets, but adding information that one of the patients had known about their illness for two years whereas the other patient had just found out about their illness. An example is shown in Table 2.

Table 2. Example of standard DCE choice set and corresponding extension choice set

	Standard DCE choice set		**Corresponding extension choice set**	
Attribute	**Patient A**	**Patient B**	**Patient A**	**Patient B**
Life expectancy without treatment	12 months	3 months	12 months	3 months
Quality of life without treatment	50%	50%	50%	50%
Life expectancy gain from treatment	1 month	6 months	1 month	6 months
Quality of life gain from treatment	25%	25%	25%	25%
How long patient has known about illness	No information provided		0 years (just found out)	2 years

In the standard DCE choice set, patient B is in poorer health than patient A without treatment (three months life expectancy at 50% quality of life vs. 12 months life expectancy at 50% quality of life). Choosing to treat patient B would be consistent with a preference for giving priority to those who are worse off without treatment. In the extension choice set, the respondent is told that patient B has known about their illness for two years whilst patient A has only just learnt of their illness. Some respondents who chose to treat patient B in the standard DCE choice set may have done so because of a concern about how little time they have to "get their affairs in order". If so, they may switch to choosing to treat patient A in the extension choice set, as patient A will have had less time to prepare than patient B

when taking into account the fact that patient B will have known about their prognosis for some time.

We hand-picked eight standard DCE choices sets to form the basis for the extension tasks. Our selection was guided by judgements about whether the choice sets depicted scenarios of particular interest (such as the one shown in the example in Table 2) and by considerations about whether the information could be presented graphically using the same format as used for the standard DCE choice sets.

Within each block we included two extension choice sets, presented to respondents after they had completed the 10 standard DCE choice sets. One of the extension choice sets replicated the scenario depicted in a standard DCE choice set that respondents in that block had already completed, to allow within-respondent comparisons; the other replicated the scenario depicted in a standard DCE choice set from a different block. The latter was always presented first, so respondents were never faced with an extension choice set immediately following the standard DCE choice set upon which the extension choice set had been based. For every choice set in which the time with knowledge was given to one of the patients, there was another choice set (in a different block) which was identical except that the time with knowledge was given to the other patient. As with the standard DCE choice sets, we sought to avoid top-to-bottom bias by creating mirror choice sets.

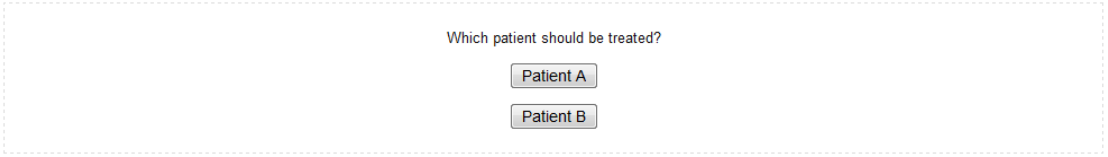
Questionnaire Design and Scenario Presentation

The choice sets formed the basis for questions in a self-completion survey administered over the Internet. The survey was developed in partnership with a software development company, EpiGenesys. The attributes and levels were presented as characteristics of two hypothetical patients (patient A and patient B) and the effects of the treatments available to them. Using a horizontal scale to represent life expectancy and a vertical scale to represent quality of life (described in the survey as “health”), we constructed diagrams of the sort shown in Figure 1. These diagrams appeared directly above the corresponding text descriptions of each patient’s attribute levels (presented using bullet points, as shown in Figure 1). The use of both text and diagrams to present the choice set information was

informed by feedback given by respondents in the preference study. The text/diagrams for patient A always appeared directly above the text/diagrams for patient B, with the "choice buttons" (see Figure 2) at the bottom of the screen.

Figure 1. Example of diagram and text used in the standard DCE tasks

Figure 2. "Choice buttons"



The survey began with a set of instructions which introduced the diagrams as a way of showing how different illnesses and treatments can affect people's health and life expectancy in different ways. The instructions are reproduced in full in the Appendix. Respondents were asked to indicate which patient they thought should be treated,

assuming that the health service has only enough funds to treat one of the two patients, and that there are no alternative treatments available. It was emphasised that there are no right or wrong answers to the questions.

Respondents were advised that they would be given information about the patients' health and life expectancy with and without treatment, but that no other information about the patients is available (except that they are both adults). To prevent respondents from making choices based on expectation or hope that a cure for the patients' illnesses may be found in the future, they were told that "the nature of the illnesses is such that further treatment will not be possible if either patient is not treated today – this is the only opportunity for treatment."

The 10 standard DCE tasks were presented to respondents in a random order so as to prevent order bias. After completing the standard DCE tasks, further instructions were provided to explain the additional "time with knowledge" attribute. The diagrams were modified to incorporate this attribute, as shown in Figures 3 and 4. Respondents were again asked to indicate which patient they thought should be treated.

Figure 3. Example of diagram used in the extension tasks (time with knowledge = 2 years)

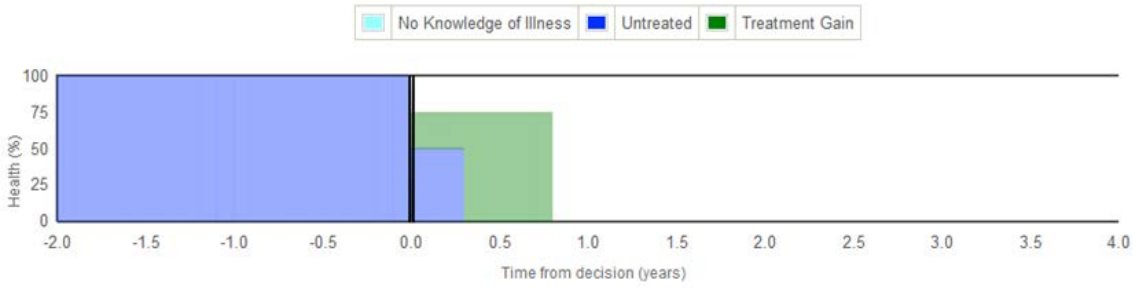
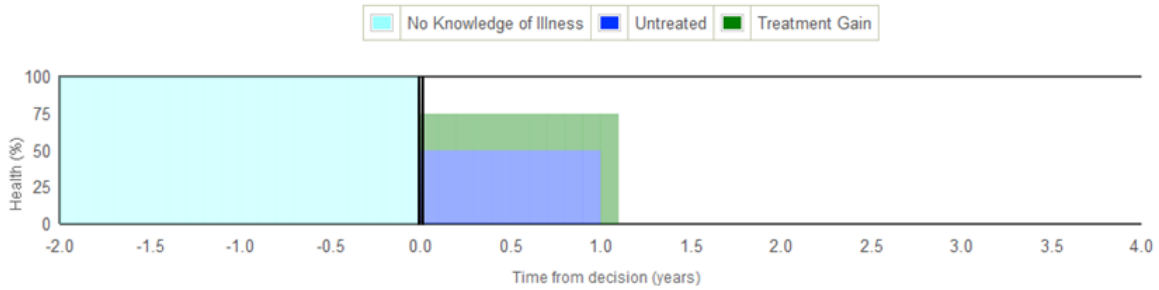


Figure 4. Example of diagram used in the extension tasks (time with knowledge = 0 years)



After completing the two extension tasks, respondents were asked some tick-box questions about their background and health. Finally, they were given the opportunity to leave feedback in an open-ended comment box if they so wished.

The decision to use questions involving forced choices without a "neither A nor B" option was informed by a number of considerations. First, it was felt that even if patients found it difficult to choose between the two patients, they would nevertheless prefer to treat one of them rather than to treat neither, since some health gain is preferable to the baseline of no health gain to either. Second, we suspected that such an option may be used as a default ("opt-out") choice, thus providing a way to avoid taking time to make difficult decisions. This was a particular concern due to the unsupervised, self-complete survey setting, which may encourage respondents to seek shortcuts in order to complete the survey as quickly as possible. Bridges et al (2011) advise that the inclusion of indifference options results in the censoring of data, which can limit researchers' ability to estimate the underlying preference structure. Finally, if respondents are genuinely indifferent between treating the two patients, this should simply result in a roughly even split between patient A and patient B in the choice data. The use of mirror choice sets controls for the possibility that respondents will revert to a default choice, such as the patient presented first, every time they are unable to choose between the patients.

The patients, illnesses and treatments were described in generic terms (e.g. patient A, patient A's illness, treatment for patient A) due to concern that the use of labels (e.g. stating that the illness could be cancer) would induce emotional and biased responses. This approach is consistent with the existing NICE policy which does not distinguish between

different illnesses or treatments. The generic presentation of health care priority setting scenarios is supported by the findings of Roberts et al (1999), who found that the level of respondent engagement was not sensitive to the provision of supporting clinical information.

Use of Web-Based Surveys / Online Panels

Web-based surveys offer a cost-effective means of collecting a large amount of data in a very short period of time. Large samples are difficult to achieve using other modes of administration: postal self-complete surveys have very low responses rates; surveys administered as part of face-to-face interviews (whether undertaken in homes of respondents or in a set of study sites) are expensive to manage; and the complexity of the questions precludes the use of telephone-based data collection for this type of survey. By comparison, web-based surveys can be custom-designed to present information and collect choice data in a clear, user-friendly manner. The ability to store data online and export securely into an electronic database should ease the usual concerns about the quality of data entry.

Interviewer-led survey administration is often preferred because the interviewer can explain the instructions more fully if required (Bridge et al, 2011) and respondents may be more likely to give their full attention to the survey whilst under supervision. However, the use of interviewers can lead to forms of interviewer bias – for example, if when explaining the instructions the interviewer gives subtle clues that influence the respondent towards certain preferences or choice strategies. With web-based surveys, the questions and instructions are presented in the same manner to all respondents (whilst presentation may differ according to the hardware/software being used, it is reasonable to assume that any variability will be random and unlikely therefore to result in systematic bias).

Whilst the vast majority of households in the UK now have access to the Internet (ONS, 2011), concerns remain about the extent to which a sample made up of members of an online panel can be said to be representative of the general population. Although quotas can be used to ensure representativeness in terms of certain observable characteristics (e.g.

age), it is likely that the sample will still be systematically different in terms of other unobservable characteristics. However, this issue is not specific to web-based data collection. The types of individuals who are willing to complete postal surveys or to allow interviewers into their homes for face-to-face interviews are similarly unlikely to be representative of the general population. Some market research agencies claim that providing an incentive for completing surveys can help to improve representativeness as an unpaid survey is more likely to be completed only by those passionately interested in the subject of that particular survey (YouGov, 2012).

Sample / Data Collection

The survey was administered on a sample of adult members of the general public in England and Wales, all of whom were members of a panel of a market research agency. The agency was responsible for inviting potential respondents to take part. A "minimum quota" approach, combined with a targeted invitation strategy, was used to ensure that the sample was representative of the general population in terms of age, gender and social grade (using data from the 2001 Census (ONS, 2011). The target sample size of 4,000 was determined on the basis that this was the largest sample that could be recruited within the required timelines.

Quotes were obtained from three different market research agencies. The selected agency, ResearchNow, was chosen on the basis of their quality control procedures and the large size of their panel. Individuals who had recently completed health-related surveys were not invited to take part. Respondents were compensated by way of "reward points" which can be redeemed for gift vouchers or charity donations. Completion statistics, including the age, gender and social grade of respondents who had completed the survey, were checked daily and used to guide the targeting of invitations. Once a quota for a particular subgroup had been reached, individuals attempting to access the survey who fell under that subgroup were "screened out" and informed that they were not eligible to take part. Once respondents had been "screened in" and given their informed consent to take part, they were randomly assigned to one of the 16 blocks.

The survey and sample recruitment procedures were given ethics approval by the Ethics Committee of the University of Sheffield's School of Health and Related Research.

Piloting

The main study was preceded by a pilot, which used a convenience sample of 12 members of non-academic staff and postgraduate research students at the University of Sheffield (excluding those in the Faculty of Medicine, Dentistry and Health or the Department of Economics). The pilot comprised face-to-face interviews conducted by one of the authors (KKS) in which respondents completed the survey (accessed via a laptop connected to the Internet) without assistance, and then answered some verbal probing questions designed to elicit feedback and concerns about the survey and approach.

The pilot was completed successfully, supporting the acceptability of the text and diagrams used in the survey and the feasibility of the proposed methods (e.g. choice of attribute levels, forced choice elicitation, web-based survey, randomisation processes). The instructions and choice tasks were described by most of the pilot respondents as being clear and easy to follow. Some of the wording of the instructions was improved following feedback from two of the respondents. All of the respondents stated that they were able to understand and complete the questions without assistance. The levels of understanding and engagement (as perceived by the interviewer) were high. Respondents spent between six and 14 minutes completing the questions (mean = 9 minutes 10 seconds).

Further testing was conducted by way of a "soft launch" data collection strategy. Once approximately 750 respondents had completed the survey, the survey was closed and the data were checked for any issues. Whilst the average time taken to complete the questions (choice tasks and follow-up questions) was consistent with the pilot (mean = 9 minutes 44 seconds), it was noted that 14 respondents (2%) completed the questions in less than three minutes. We questioned whether it was possible to complete the survey this quickly whilst paying adequate attention to the tasks at hand. We observed that when faced with choice sets in which one alternative dominated the other, nine of these 14 respondents (64%) failed to choose the dominant alternative. Since patient A and patient B were overall equally

likely to be represent the dominant alternative in these choice sets, we would expect a respondent who is not taking the survey seriously (e.g. making choices at random) to have a 50% chance of choosing the dominant alternative. By comparison, only 12% of respondents who spent at least three minutes completing the questions failed to choose the dominant alternative. We therefore judged that it is reasonable to exclude from the analysis data for respondents who completed the questions in less than three minutes on grounds of poor data quality.

The soft launch approach also provided an opportunity to examine the open-ended comments left by respondents, in case these highlighted any problems with the survey. Of the 100 or so comments that had been left, the majority were positive (e.g. "very well set out and easy to navigate"). Three respondents left comments about one of the background questions, stating that they were unsure about which category they belonged to when asked about the occupation of the chief income earner of their household. We amended the instructions to the question to address these comments. No other changes to the survey were deemed necessary in light of the soft launch data analysis.

Functional Form and Data Analysis

Choice data were modelled using a random utility maximisation framework (Louviere, Henscher and Swait, 2000) and STATA 11.2 software. The conditional logit model was used for modelling. This allowed us to include information about the attributes of the alternatives in the model.

The model estimated is of the form:

$$V = \beta_1 LE\ without\ treatment + \beta_2 QOL\ without\ treatment + \beta_3 LE\ gain + \beta_4 QOL\ gain$$

The deterministic component of the utility function (*V*) is a function of the attribute levels between alternatives, where the coefficients β_1-β_4 are estimated in the model. These coefficients can be summed to give the overall utility for each profile (combination of attribute levels). This gives us an indication of the relative social value of the 110 profiles in

the experimental design. The attributes are coded using dummy variables in order to allow for non-linear relationships, which means that each of the explanatory variables is treated as categorical.

It is expected that the coefficients β_3 and β_4 will have positive signs (larger QALY gains from treatment increase the probability of that treatment being chosen) whilst the coefficients β_1 and β_2 may have either positive or negative signs. Negative signs for β_1/β_2 would indicate an overall preference for prioritising the treatment of those who are worse off in terms of their quality of life/life expectancy without treatment. This would be consistent with the findings from the preference study and from a number of other empirical studies of priority setting preferences (for reviews, see Dolan et al (2005) and Shah (2009)). However, many such studies (including the preference study) have also reported that a sizeable minority of respondents do not support such a "priority to the worse off" approach.

We also defined three interactions that we felt *a priori* were likely to be influential. These were: (i) life expectancy without treatment against life expectancy gain (small gains in life expectancy may be increasingly important when life expectancy without treatment is short); (ii) life expectancy without treatment against quality of life gain (whether a quality of life improvement or a gain in life expectancy is preferred may depend on life expectancy without treatment); and (iii) life expectancy gain against quality of life gain (the important of a gain in life expectancy may depend on whether it is accompanied by a quality of life improvement).

As described earlier, we are able to transform the logit model results in order to show the probability of choosing a given profile from the complete set of profiles. Following the approach used by Green and Gerard (2009) we calculated the relative predicted probabilities for all of the 110 profiles, allowing us to compare the profiles with higher probabilities (those which are likely to be most preferred overall) with those with lower probabilities (those which are likely to be least preferred overall).

Finally, we defined *a priori* a selection of respondent subgroups whose choices may be expected to differ from those of the rest of the sample. These were: (i) respondents with

experience of close friends or family with terminal illness; (ii) respondents with responsibility for children under 18; (iii) respondents who opted to leave a comment in the open-ended box at the end of the survey (this was not mandatory); and (iv) respondents who completed the questions much quicker than average. Information on (i) and (ii) was obtained using the tick-box background questions. (iii) and (iv) may be indicators of engagement and attentiveness – we would expect that a respondent who took the survey seriously would spend longer on the questions and be more likely to leave a comment than one who did not. For each subgroup, we estimated the best fitting model and compared the results to those of the same model using the full sample.

RESULTS

In total, 43,000 individuals were invited by email to take part in the survey, of whom 5,308 clicked on the link to access the survey (response rate = 12.3%). Of the individuals who accessed the survey, 4,008 completed the survey in full (completion rate = 75.5%). The remainder either did not give consent to take part, or began the survey but dropped out without completing all of the questions. The response and completion rates for this survey are consistent with those of similar web-based surveys whose sample comprised members of ResearchNow's panel.

The survey allowed respondents to go back to previous questions and change their answers if they so wished. Only the final answers were used in the data analysis, leaving a data set comprising 48,096 pairwise observations. As described above, it was agreed that data for respondents who spent less than three minutes on the questions would be suppressed from the final data set. This cut-off excluded a further 39 respondents, leaving 3,969 respondents (47,628 pairwise observations). Of these 47,628 observations, 39,690 were for the standard DCE tasks to be analysed using the conditional logit model above; the remaining 7,938 were for the extension tasks.

By design, the sample was representative of the general population in England and Wales with respect to age and gender (ONS, 2011). Despite the use of quotas to seek representativeness in terms of social grade, the sample contains a larger proportion of

individuals in the highest grades and also those in the very lowest grade than in the general population (NRS, 2010), presumably due to changes in circumstances since these individuals joined the panel. The background characteristics of the sample are presented in Table 3.

Of the sample, 389 respondents (9.8%) failed to choose the dominant option when faced with choice sets in which one alternative dominated the other (i.e. where both patients have the same amount of life expectancy and quality of life without treatment, but one patient gains more life expectancy and more quality of life from treatment than the other). However, it is not necessarily the case that these preferences are "irrational" – Lancsar and Louviere (2006) warn against researchers imposing their own preferences by deleting responses that do not conform to their expectations. We therefore included data for these respondents in the analysis.

Table 3. Sample background characteristics

	#	%	gen pop[n] %
Total	3,969	100	100
Gender			
Male	1,942	49	49
Female	2,027	51	51
Age			
18-24	404	10	11
25-44	1,413	36	38
45-64	1,228	31	31
65+	924	23	21
Social grade [*]			
A (Higher managerial, administrative or professional)	221	6	4
B (Intermediate managerial, administrative or professional)	1,114	28	22
C1 (Supervisory or clerical and junior managerial, administrative or professional)	1,150	29	29
C2 (Skilled manual workers)	645	16	21
DE (Semi and unskilled manual workers)	357	9	15

	#	%	gen pop[n] %
E (Casual or lowest grade workers, pensioners and others who depend on the welfare state for their income)	482	12	8
Household composition			
With children	963	24	
Without children	3,006	76	
Education			
No education beyond minimum school leaving age	889	22	
Education beyond minimum school leaving age; no degree	1,244	31	
Education beyond minimum school leaving age; degree	1,836	46	
Self-reported general health level			
Very good	1,008	25	
Good	1,958	49	
Fair	770	19	
Poor	210	5	
Very poor	23	1	
Experience of close friends or family with terminal illness			
Yes	2,689	68	
No	1,197	30	
Question skipped by respondent	83	2	

* Refers to the occupation / qualifications / responsibilities of the chief wage earner of the respondent's household

Descriptive Statistics of The Choices Made

Overall, there was a tendency to choose to treat the alternative labelled patient B (the alternative appearing at the bottom of the respondent's screen) – the difference between the proportion of respondents choosing to treat patient A and the proportion choosing to treat patient B is statistically significant ($p = 0.00$).

We tested for differences between mirror blocks by comparing the proportion of respondents in each original block who chose to treat patient A (B) with the proportion of respondents in the corresponding mirror block who chose to treat patient B (A), across all choice sets (recall that the alternative labelled as patient A in the original choice set is

always labelled as patient B in the corresponding mirror block choice set, and vice versa). We found that with one exception, wherever the most popular choice was patient A (B) in the original block, the most popular choice was always patient B (A) in the mirror block. For each pair of blocks, we also tested whether the overall proportion choosing patient A (B) in the original block was statistically significantly different from the proportion choosing patient B (A) in the mirror block. In two of the eight pairs, there was a statistically significant difference ($p = 0.02$ and $p = 0.00$); in the remaining six there was no difference.

For each choice set, we calculated the "level of agreement" amongst respondents in terms of the proportion choosing the majority choice. Excluding choice sets comprising dominant/dominated alternatives, the level of agreement ranged from 50% to 95%.

The 20 choice sets with the highest levels of agreement had more than 85% of respondents choosing to treat the same patient. Of these choice sets, the majority (13) involved choosing between a patient who is *worse off* (in terms of QALYs) without treatment and *gains more* from treatment (also in terms of QALYs), and a patient who is *better off* without treatment and *gains less* from treatment. Other choice sets involved choosing between a patient who is *worse off* without treatment and *gains less* from treatment, and a patient who is *better off* without treatment and *gains more* from treatment. In all cases, the vast majority of respondents always chose to treat the patient who gains more from treatment, regardless of whether that patient is better or worse off without treatment.

By contrast, there were 25 choice sets in which the most popular choice was made by no more than 60% of respondents. Some of these choice sets involved choosing between a patient who is *worse off* without treatment and *gains more* from treatment, and a patient who is *better off* without treatment and *gains less* from treatment. *A priori*, we might expect most respondents to choose to treat the worse off patient who gains more from treatment when faced with this type of choice set (this would be consistent with both the QALY-maximisation and "priority to the worst off" approaches to health care priority setting), yet at least 40% of respondents expressed the opposite preference in these low-agreement choice sets.

Table 4 reports the average level of agreement for the choice sets belonging to each "choice type".

Table 4. Average level of agreement, by choice type

Description of choice type	No. choice sets	Level of agreement
Both patients have the same LE / QOL without treatment. One patient gains more LE and more QOL from treatment than the other.	11	92%
Patient with lower QOL without treatment gains more LE and more QOL from treatment.	5	85%
Patient with shorter LE and higher QOL without treatment gains more LE from treatment.	1	85%
Patient with longer LE without treatment gains more QOL from treatment.	14	78%
Patient with shorter LE without treatment gains more LE from treatment.	2	76%
Patient with longer LE without treatment gains more LE from treatment.	4	74%
Patient with longer LE and lower QOL without treatment gains more QOL from treatment.	2	72%
Both patients have the same LE / QOL without treatment. One patient gains more LE from treatment; the other gains more QOL from treatment.	10	69%
Patient with lower QOL without treatment gains more QOL from treatment. Patient with higher QOL without treatment gains more LE from treatment.	9	68%
One patient has longer LE without treatment; the other has higher QOL without treatment. Both patients gain same amount of LE / QOL from treatment.	4	68%
Patient with shorter LE without treatment gains more QOL from treatment.	10	68%
Patient with shorter LE and lower QOL without treatment gains more QOL from treatment.	5	68%
One patient has longer LE and higher QOL without treatment than the other. Both patients gain same amount of LE / QOL from treatment.	3	67%

Across all choice sets in which one of the patients is worse off and gains more from treatment, that patient was chosen by respondents on 71.8% of occasions. In a subset of these choice sets, the worse off patient remains worse off despite gaining more from

treatment. In these cases, the worse off patient was chosen by respondents on 68% of occasions. The choice sets with the lowest levels of agreement tended to involve smaller than average differences between the two patients' QALY gains from treatment (mean difference in the 25 low-agreement choice sets = 0.44 QALYs; mean difference in all 80 choice sets = 0.90 QALYs).

Discrete Choice Model Results

Table 5 reports the results of the conditional logit model for analysis of main effects only. All main effects coefficients are positive, and all but two are statistically significant. For the life expectancy without treatment and quality of life without treatment attributes, the signs of the coefficients suggest that the better off a patient is in terms of life expectancy and quality of life without treatment, the more likely respondents are to choose to treat that patient. The fact that the coefficients for all attribute levels are positive indicates that incremental increases in any of the attributes will lead to an increase in predicted utility.

Table 5. Conditional logit model results (main effects)

Attribute / level	Coefficient	p-value
LE without treatment		
3 months [baseline]	-	-
12 months	0.0155	0.63
24 months	0.1543	0.00
36 months	0.3350	0.00
60 months	0.3977	0.00
QOL without treatment		
50% [baseline]	-	-
100%	0.7093	0.00
LE gain		
0 months [baseline]	-	-
1 month	0.0349	0.41
2 months	0.4145	0.00
3 months	0.8135	0.00
6 months	1.4031	0.00
12 months	2.3295	0.00
QOL gain		
0% [baseline]	-	-
25%	1.0235	0.00
50%	1.7981	0.00

Akaike information criterion = 42797; Bayesian information criterion = 42908

As described in the methods section, we identified three interactions as explanatory variable candidates. We estimated three separate models that added each of these interactions to the main effects individually, and two further models that added combinations of the interactions. All interactions models were shown to fit better than the main effects model according to the Akaike and Bayesian information criteria and likelihood ratio tests (p = 0.00). The results of the best fitting model, which analysed main effects plus all of the three interactions, are reported in Table 6.

Table 6. Conditional logit model (main effects plus interactions)

Attribute / level	Coefficient	p-value
LE without treatment		
3 months [baseline]	-	-
12 months	0.1755	0.12
24 months	0.9307	0.00
36 months	0.7841	0.00
60 months	1.2625	0.00
QOL without treatment		
50% [baseline]	-	-
100%	0.6730	0.00
LE gain		
0 months [baseline]	-	-
1 month	0.1855	0.08
2 months	0.8517	0.00
3 months	1.0855	0.00
6 months	2.0433	0.00
12 months	2.9381	0.00
QOL gain		
0% [baseline]	-	-
25%	0.0632	0.47
50%	1.0212	0.00
Interaction: LE without treatment # LE gain		
12 months # 1 months	-0.1715	0.15
12 months # 2 months	-0.4220	0.00
12 months # 3 months	-0.1633	0.18
12 months # 6 months	-0.7294	0.00
12 months # 12 months	-0.6039	0.00
24 months # 1 months	-1.1308	0.00
24 months # 2 months	-1.0782	0.00
24 months # 3 months	-0.8614	0.00
24 months # 6 months	-1.2413	0.00
24 months # 12 months	-1.2601	0.00
36 months # 1 months	-0.7280	0.00
	-1.0428	0.00

Attribute / level	Coefficient	p-value
36 months # 2 months	-1.2252	0.00
36 months # 3 months	-1.6695	0.00
36 months # 6 months	-1.3963	0.00
36 months # 12 months	-1.3159	0.00
60 months # 1 months	-1.4933	0.00
60 months # 2 months	-1.2558	0.00
60 months # 3 months	-2.0434	0.00
60 months # 6 months	-1.7114	0.00
60 months # 12 months		
Interaction: LE without treatment # QOL gain		
12 months # 25%	0.4562	0.00
12 months # 50%	0.2139	0.00
24 months # 25%	0.2734	0.00
24 months # 50%	0.4123	0.00
36 months # 25%	0.8457	0.00
36 months # 50%	0.7374	0.00
60 months # 25%	0.5379	0.00
60 months # 50%	0.6676	0.00
Interaction: LE gain # QOL gain		
1 months # 25%	0.7649	0.00
1 months # 50%	0.5254	0.00
2 months # 25%	0.3197	0.00
2 months # 50%	0.3543	0.00
3 months # 25%	0.6321	0.00
3 months # 50%	0.3163	0.00
6 months # 25%	0.6661	0.00
6 months # 50%	0.2744	0.00
12 months # 25%	0.3466	0.00
12 months # 50% [baseline]	-	-

Akaike information criterion = 41579; Bayesian information criterion = 42034

The individual coefficients are difficult to interpret in isolation because overall utility is a function of both main effects and interactions combined. To enable interpretation of the model results, Table 7 presents the utility scores based on the best fitting model for all of the 110 profiles in the full factorial design, as well as the predicted probability of choosing each profile from all the full set profiles. The probabilities are standardised so as to sum to 1.000. The reference case profile (italicised) was selected as the one that we felt most closely matched the profile of a treatment that just meets the current NICE end of life criteria (life expectancy without treatment = 24 months; quality of life without treatment = 100%; life expectancy gain = 3 months; quality of life gain = 0%). This profile was ranked 60th out of the 110 profiles. Forty-three of the 59 profiles ranked higher than the reference case

profile involved at least some quality of life gain. By comparison, only 17 of the 50 profiles ranked lower than the reference case profile involved any quality of life gain.

Table 7. Estimated utility score and predicted probability of choice for all 110 profiles

Rank -best fitting model	Rank - main effects model	LE without treatment (mths)	QOL without treatment (%)	LE gain (mths)	QOL gain (%)	Utility	Prob.	Cumul. Prob.
1	1	60	50	12	50	4.17809	0.0155	0.0155
2	2	36	50	12	50	4.08461	0.0154	0.0309
3	3	24	50	12	50	4.04235	0.0153	0.0462
4	5	3	50	12	50	3.95938	0.0152	0.0614
5	4	12	50	12	50	3.74493	0.0148	0.0762
6	20	3	100	12	0	3.61116	0.0145	0.0908
7	7	36	50	12	25	3.58153	0.0145	0.1052
8	14	24	50	6	50	3.44063	0.0142	0.1194
9	6	60	50	12	25	3.43703	0.0142	0.1336
10	13	12	50	12	25	3.37581	0.0140	0.1476
11	15	3	50	12	25	3.34797	0.0139	0.1615
12	17	3	50	6	50	3.33890	0.0139	0.1754
13	10	24	50	12	25	3.29205	0.0138	0.1892
14	18	24	100	12	0	3.28184	0.0138	0.2030
15	8	60	50	6	50	3.22563	0.0136	0.2166
16	9	36	50	6	50	3.19091	0.0135	0.2302
17	19	12	100	12	0	3.18284	0.0135	0.2437
18	11	60	100	12	0	3.16229	0.0134	0.2571
19	21	60	50	3	50	3.09734	0.0133	0.2704
20	12	36	100	12	0	2.99902	0.0130	0.2833
21	16	12	50	6	50	2.99895	0.0130	0.2963
22	46	60	50	0	50	2.95135	0.0128	0.3091
23	40	3	50	12	0	2.93815	0.0128	0.3219
24	24	24	50	3	50	2.90469	0.0127	0.3346

Rank -best fitting model	Rank - main effects model	LE without treatment (mths)	QOL without treatment (%)	LE gain (mths)	QOL gain (%)	Utility	Prob.	Cumul. Prob.
25	37	3	50	6	25	2.77261	0.0122	0.3468
26	31	24	50	6	25	2.73544	0.0121	0.3589
27	25	36	50	6	25	2.73295	0.0121	0.3710
28	22	36	50	3	50	2.71935	0.0120	0.3831
29	51	3	100	6	0	2.71627	0.0120	0.3951
30	32	36	50	2	50	2.70598	0.0120	0.4071
31	36	12	50	6	25	2.67496	0.0119	0.4190
32	30	60	50	2	50	2.66403	0.0119	0.4309
33	28	12	50	3	50	2.64921	0.0118	0.4427
34	34	24	50	12	0	2.60883	0.0117	0.4543
35	49	36	50	0	50	2.54273	0.0114	0.4657
36	23	60	50	6	25	2.52970	0.0114	0.4771
37	48	36	50	1	50	2.52572	0.0113	0.4884
38	39	12	50	12	0	2.50983	0.0113	0.4997
39	38	24	50	2	50	2.49203	0.0112	0.5109
40	26	60	50	12	0	2.48928	0.0112	0.5221
41	29	3	50	3	50	2.42305	0.0110	0.5331
42	41	24	100	6	0	2.40570	0.0109	0.5440
43	54	24	50	0	50	2.36426	0.0107	0.5547
44	43	60	50	1	50	2.34635	0.0106	0.5653
45	27	36	50	12	0	2.32601	0.0106	0.5759
46	42	60	50	3	25	2.32540	0.0106	0.5865
47	57	12	50	3	25	2.24921	0.0103	0.5967
48	45	3	50	2	50	2.22724	0.0102	0.6069
49	44	12	50	2	50	2.19467	0.0100	0.6169
50	47	36	50	3	25	2.18539	0.0100	0.6269
51	50	12	100	6	0	2.16246	0.0099	0.6368
52	52	24	50	3	25	2.12350	0.0097	0.6466

Rank -best fitting model	Rank - main effects model	LE without treatment (mths)	QOL without treatment (%)	LE gain (mths)	QOL gain (%)	Utility	Prob.	Cumul. Prob.
53	79	3	50	6	0	2.04327	0.0094	0.6560
54	58	12	50	1	50	1.95006	0.0090	0.6650
55	53	24	50	1	50	1.94442	0.0090	0.6740
56	33	60	100	6	0	1.93544	0.0090	0.6829
57	80	36	50	1	25	1.91552	0.0089	0.6918
58	77	60	50	0	25	1.86368	0.0086	0.7005
59	35	36	100	6	0	1.83092	0.0085	0.7090
60	67	24	100	3	0	1.82784	0.0085	0.7175
61	65	36	50	2	25	1.82166	0.0085	0.7259
62	59	3	50	3	25	1.78076	0.0083	0.7342
63	70	12	100	3	0	1.77078	0.0083	0.7425
64	55	60	100	3	0	1.76521	0.0082	0.7507
65	71	3	100	3	0	1.75849	0.0082	0.7589
66	69	24	50	6	0	1.73269	0.0081	0.7670
67	61	3	50	1	50	1.73218	0.0081	0.7751
68	81	36	50	0	25	1.69303	0.0079	0.7830
69	60	60	50	2	25	1.54173	0.0073	0.7903
70	89	3	100	2	0	1.52474	0.0072	0.7975
71	74	60	50	1	25	1.49817	0.0071	0.8046
72	78	12	50	6	0	1.48945	0.0071	0.8117
73	91	12	50	1	25	1.47382	0.0070	0.8187
74	75	12	50	2	25	1.44432	0.0069	0.8256
75	62	12	50	0	50	1.41064	0.0067	0.8323
76	82	24	100	2	0	1.37723	0.0066	0.8390
77	68	24	50	2	25	1.36049	0.0065	0.8455
78	56	36	100	3	0	1.31743	0.0064	0.8519
79	72	60	100	2	0	1.29395	0.0063	0.8582
80	88	12	100	2	0	1.27830	0.0062	0.8644

Rank -best fitting model	Rank - main effects model	LE without treatment (mths)	QOL without treatment (%)	LE gain (mths)	QOL gain (%)	Utility	Prob.	Cumul. Prob.
81	85	24	50	0	25	1.26734	0.0062	0.8705
82	73	36	100	2	0	1.26611	0.0062	0.8767
83	63	60	50	6	0	1.26243	0.0062	0.8829
84	76	3	50	2	25	1.23459	0.0060	0.8889
85	66	36	50	6	0	1.15791	0.0058	0.8947
86	95	24	50	3	0	1.15483	0.0057	0.9004
87	97	12	50	3	0	1.09778	0.0055	0.9059
88	84	60	50	3	0	1.09221	0.0055	0.9114
89	83	24	50	1	25	1.08699	0.0055	0.9169
90	98	3	50	3	0	1.08548	0.0055	0.9224
91	64	3	50	0	50	1.02123	0.0052	0.9276
92	92	3	50	1	25	1.01365	0.0052	0.9329
93	90	36	100	1	0	0.91472	0.0049	0.9377
94	100	12	100	1	0	0.86256	0.0047	0.9424
95	102	3	100	1	0	0.85854	0.0047	0.9471
96	106	3	50	2	0	0.85173	0.0046	0.9517
97	87	60	100	1	0	0.80514	0.0045	0.9562
98	103	24	50	2	0	0.70422	0.0042	0.9604
99	93	12	50	0	25	0.69491	0.0041	0.9645
100	96	24	100	1	0	0.65849	0.0040	0.9686
101	86	36	50	3	0	0.64442	0.0040	0.9725
102	99	60	50	2	0	0.62094	0.0039	0.9765
103	105	12	50	2	0	0.60529	0.0039	0.9803
104	101	36	50	2	0	0.59310	0.0038	0.9842
105	107	36	50	1	0	0.24171	0.0029	0.9870
106	109	12	50	1	0	0.18955	0.0028	0.9898
107	110	3	50	1	0	0.18553	0.0028	0.9926
108	104	60	50	1	0	0.13213	0.0026	0.9952

Rank -best fitting model	Rank - main effects model	LE without treatment (mths)	QOL without treatment (%)	LE gain (mths)	QOL gain (%)	Utility	Prob.	Cumul. Prob.
109	94	3	50	0	25	0.06320	0.0025	0.9977
110	108	24	50	1	0	-0.01452	0.0023	1.0000

The highest ranked profiles (i.e. those with the greatest probability of being chosen) all involve substantial health gains. All of the profiles ranked between 1st and 24th involve a life expectancy gain of 12 months and/or a quality of life gain of 50%. This can be contrasted to the lowest ranked profiles, most of which involve a small life expectancy gain (one or two months) and no quality of life gain. A similar pattern with respect to life expectancy without treatment does not exist – profiles involving the highest and lowest levels for this attribute (60 months and 3 months, respectively) appear at both the top and bottom of Table 7. Quality of life without treatment is 50% in most of the highest ranked profiles, but this is always accompanied by a non-zero quality of life gain from treatment.

We also calculated utility scores and predicted probabilities using the results from the main effects only model. The rankings were broadly similar to those calculated using the best fitting model (Spearman's rho = 0.949), but profiles which involved lower levels of life expectancy without treatment (and, to a lesser extent, quality of life without treatment) had relatively higher utility scores and predicted probabilities when interactions were taken into account. For example, the profile ranked 79th according to the best fitting model, which involves the lowest possible levels for the life expectancy and quality of life without treatment attributes, was ranked 53rd using results from the main effects only model. This indicates that the apparent positive relationships between a patient's life expectancy and quality of life without treatment and the likelihood of choosing to treat that patient (as suggested by Table 5) may rely on a failure to take relevant interactions into account in the modelling.

Table 8 provides a summary of the highest and lowest ranked profiles by reporting mean levels of the four attributes as well as the mean number of QALYs without treatment and QALYs gained from treatment implied by those attribute levels. This indicates that there is little difference between the highest and lowest ranked profiles in terms of QALYs without treatment – the difference is driven by the difference in the sizes of the QALY gains from treatment. This finding is robust over a range of different definitions of "most preferred" and "least preferred".

Table 8. Summary (mean attribute levels) of the highest and lowest ranked profiles

	LE without treatment (months)	QOL without treatment (%)	LE gain (months)	QOL gain (%)	QALYs without treatment	QALYs gained
10 most preferred profiles	27	55	11	38	1.14	1.76
20 most preferred profiles	29	63	10	31	1.51	1.60
55 most preferred profiles	27	57	7	31	1.27	1.22
55 least preferred profiles	27	65	2	10	1.49	0.29
20 least preferred profiles	23	63	1	6	1.24	0.09
10 least preferred profiles	28	50	1	3	1.18	0.06

Figure 5 illustrates the levels of QALYs without treatment and QALYs gained from treatment associated with all of the 110 profiles, where the horizontal axis represents the standardised

predicted probabilities from the lowest (least preferred) to the highest (most preferred) profile. Whilst the patterns are noisy, the green linear trend line for QALYs gained from treatment has a clear upward slope (the larger the size of the QALY gains, the greater the probability of the profile being chosen). The blue linear trend line for QALYs without treatment is flat, indicating that the number of the QALYs without treatment does not have a major effect on the probability of the profile being chosen. By contrast, the corresponding QALYs without treatment trend line for the main effects only model data (not shown here) is slightly upward sloping. This further highlights the importance of taking into account the interactions between attributes.

Figure 5. Levels of QALYs without treatment / gained from treatment associated with all 110 profiles

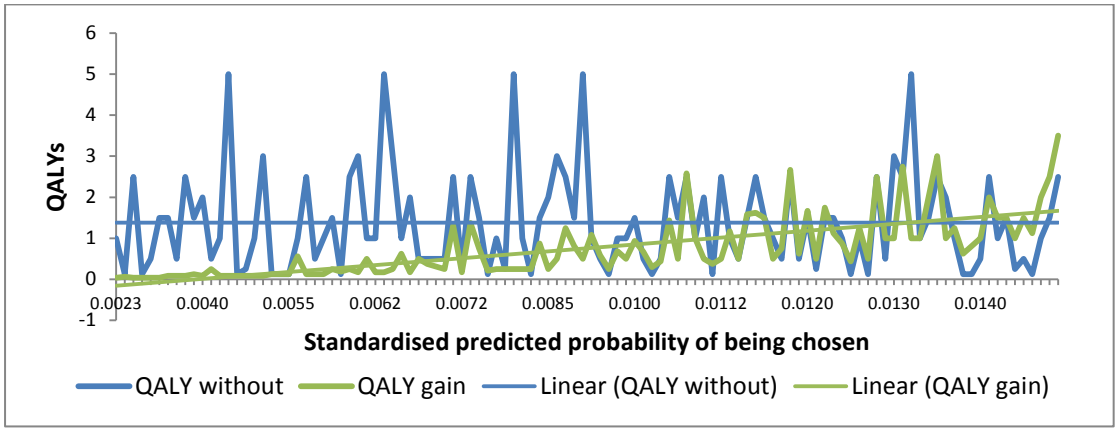
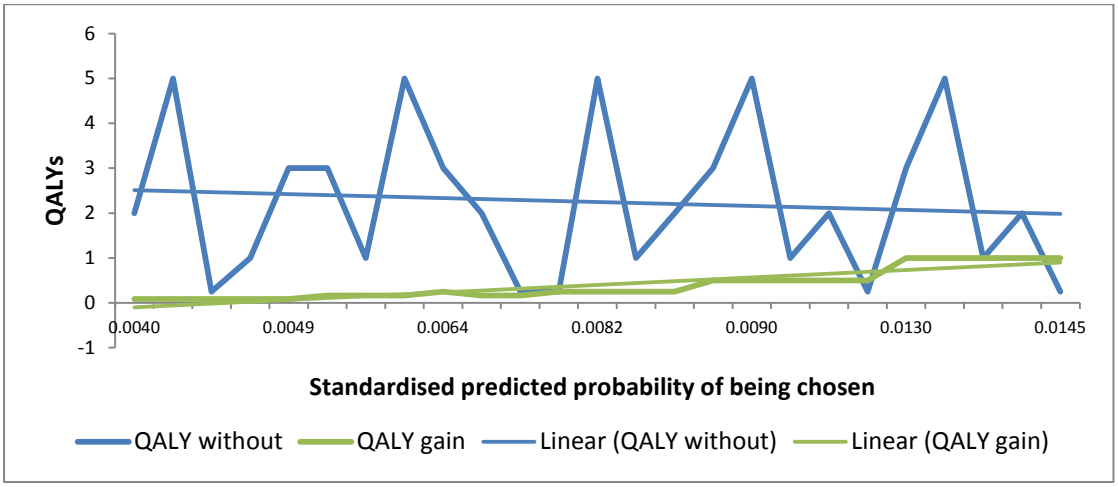


Figure 6 presents the same information for the profiles in which quality of life (with or without treatment) is 100%. Here the differences in QALYs between profiles are driven by differences in the levels of life expectancy with and without treatment. The upward slope for QALYs gained from treatment is now free of spikes, which indicates that when quality of life is controlled for, the probability of a profile being chosen is well explained by the size of the life expectancy gain. The linear trend line for QALYs without treatment is slightly downward sloping which indicates weak support for the claim that people prefer to treat patients with shorter life expectancy without treatment. However, the curve itself remains characterised by sharp spikes, and the trend line does not appear to fit to the data points very well.

Figure 6. Levels of QALYs without treatment / gained from treatment associated with the 25 profiles with QOL with / without treatment = 100%



As described above, the reference case profile (broadly corresponding to the NICE end of life criteria) was ranked 60th out of the 110 profiles. Of the 59 profiles ranked higher than the reference case profile, 14 involved a life expectancy gain of less than three months, so would not meet the existing criteria for being given special consideration as a life-extending, end of life treatment. However, in all 14 cases these profiles involved a quality of life improvement. None of the profiles ranked higher than the reference case profile involved an overall health gain of less than 0.25 QALYs (equivalent to three months in full quality of life).

Subgroup Analysis

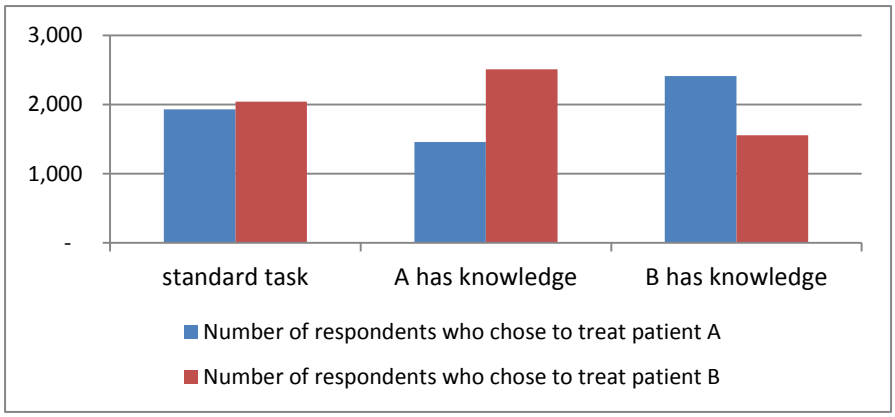
As described in the methods section, we defined respondent subgroups according to their responses to the background questions (whether or not they have experience of terminal illness in close friends or family; whether or not they have children) or to the ways in which they completed the survey (whether they left a comment or not; how quickly they completed the survey). We estimated the best fitting models for each subgroup and compared the results to those of the same model using the full sample. This analysis indicated no difference in the signs or approximate magnitude of the coefficients for any of the subgroups compared with the entire sample.

We also re-ran the best fitting model excluding the 389 respondents who failed to select the dominant alternative when faced with choice sets in which one alternative dominated the other. Excluding these respondents did not have an impact on the regression results.

Extension Tasks

Comparing the response data for the extension tasks with the data for the corresponding standard DCE tasks allows us to test whether respondents are more likely to choose to treat the patient who has just found out about their illness (e.g. due to concerns about how long they have to "prepare for death"). In all of the 16 extension tasks, being told that one of the patients has known about their illness for two years increases the proportion of respondents choosing the *other* patient compared to when no "time with knowledge" information is provided. In six of the 16 cases, this increase was sufficiently large that the majority choice in the extension task flipped from the majority choice in the corresponding standard task. Figure 7 presents the impact on choices of providing information on how long the patients have known about their illness, summed across all 16 extension tasks.

Figure 7. Impact on choices of providing "time with knowledge" information (all extension tasks)



During the study design phase, we identified one of the choice sets (and its mirror) as being of particular interest. This choice set formed the basis for the example in Table 2. Quality of life (both before and after treatment) is the same for both patients. One patient has shorter life expectancy without treatment than the other (3 months < 12 months), and despite gaining more life expectancy from treatment (6 months > 1 month), that patient continues to have shorter life expectancy after being treated (9 months < 13 months). If people wish

to give priority to those with shorter life expectancy and prefer large health gains to small health gains, then we would expect most people to choose to treat this "worse off, bigger gain" patient. This is indeed the case in the two standard DCE tasks that mirror each other (75% and 78%).

In two of the extension tasks based on these choice sets, respondents are told that the "worse off, bigger gain" patient found out about their illness two years ago, while the other ("better off, smaller gain") patient has only just found out about their illness. This means that despite their shorter remaining life expectancy, the "worse off, bigger gain" patient will have had longer to prepare for death (25 months > 12 months). With this new information, a smaller proportion (but still the majority) of respondents chose to treat this patient (57% and 57%).

DISCUSSION

This study used a web-based survey to elicit the preferences of a large sample of the general public in England and Wales, representative in terms of age and gender, over a range of health care priority setting scenarios, focusing on social preferences regarding the prioritisation of treatments for patients with short remaining life expectancy.

Dolan et al's (2005) review of the empirical ethics literature shows that a number of studies have reported evidence of people being willing to sacrifice overall health gain in order to pursue equity objectives such as the prioritisation of those who are severely ill. However, many of these studies used small, non-random samples, and involved elicitation methods that were not choice-based. Furthermore, the empirical literature more commonly defines severity in terms of quality of life than in terms of life expectancy or "proximity to death" (Shah, 2009).

The results show that choices about which patient to treat are influenced more by the sizes of the health gains achievable from treatment than by patients' life expectancy or quality of life in absence of treatment. The profiles most likely to be chosen are those with the highest

levels of both life expectancy gain (12 months) and quality of life gain (50%). Likewise, the profiles least likely to be chosen are those with very small health gains. On the other hand, the data suggest that the level of health without treatment in a given profile has little impact on the likelihood of that profile being chosen. There is certainly no indication that end of life is the driving factor; in fact, the average level of life expectancy without treatment in the 55 profiles most likely to be chosen is no different from that in the 55 profiles that are least likely to be chosen.

Analysis at the individual choice set level confirms this: in several of the choice sets showing the highest levels of agreement amongst respondents, the most popular choice was to treat the patient with longer left to live for whom treatment offered larger health gains, in favour of the patient with shorter remaining life expectancy. The overall view seems to be that giving priority to those who are worse off is desirable, but only if the gains from treatment are substantial.

In line with the findings of the preference study, the results show that people's preferences are heterogeneous. Some respondents appear to support a QALY-maximisation type objective throughout; a small minority always seek to treat those who are worse off without treatment; but the majority seem to advocate a mixture of the two approaches.

The finding in this study that respondents attach relatively little weight to how much life expectancy and quality of life without treatment patients have does not necessarily refute evidence elsewhere in the literature of popular support for the use of severity as a priority setting criterion (Shah, 2009). Our study focused on a small range of scenarios, all of which involve relatively poor prognoses (in terms of remaining life expectancy). Across all of the profiles included in the design, the patient who is "best off" without treatment would still die within five years. Thus in effect, *all* of the profiles in this study describe patients who are at, or near, their end of life to some extent. It is not possible from these data alone to infer whether the importance of the life expectancy without treatment attribute would be markedly different in a survey which asked respondents to choose between patients with very short life expectancies and patients with much longer life expectancies (e.g. those with 20 years of remaining life expectancy, who clearly cannot be described as "end of life").

The use of a web-based survey and online panel allowed us to collect a large amount of data in a very short period of time. However, this mode of administration offers limited opportunity for debriefing with respondents about their experience of completing the survey. We cannot know for certain the extent to which the choice data truly reflect respondents' beliefs and preferences. The study was designed in such a way that the results would not be biased if some respondents failed to pay adequate attention to the choice tasks (e.g. making choices at random rather than choosing the profile they actually prefer). However, if respondents had failed to read or understand the instructions, then this could be problematic. For example, they may mistakenly believe that the tasks requires them to choose which patient they would prefer to be in the position of, rather than which patient they would prefer the health service to treat. Alternatively, their choices may have been driven by a misguided hope or belief that a cure for one or both of the patients' illnesses may be discovered in the future.

Both of the above types of misunderstanding would be consistent with a preference for treating the patient with longer life expectancy. A sizeable proportion of respondents did indeed express this preference, something that is clear from the fact that the profile with the very highest probability of being chosen involves the maximum level of life expectancy without treatment (60 months). However, it is not necessarily the case that respondents who express a preference for giving priority to those with longer life expectancy do so because of misunderstanding. In the preference study, more than a quarter of respondents chose to give a six month life extension to a patient with 10 years of life expectancy without treatment in favour of an equal sized life extension to a patient with only one year of life expectancy without treatment. Data from follow-up questions suggest that this observed preference was genuine, and may have been driven by a belief that patients with longer remaining life expectancy are better placed to make the most out of a short life extension.

Whilst web-based surveys are limited in terms of the amount of data they can generate which describe the ways in which respondents completed the survey, it is possible to design follow-up questions which can be used to check whether respondents agree with the preferences implied by their answers to the DCE questions. For example, we might ask them

to rank a variety of statements describing different priority setting approaches according to the extent to which they agreed with each of them. A more explicit method would be to present them with a statement such as "The NHS should give priority to the treatment of patients who will die soon without treatment" and ask them to state whether they agreed or disagreed with that statement. This would allow us to check whether respondents agree with the policy implications of their responses to the DCE tasks. A high level of agreement would add legitimacy to the DCE results. These kinds of preference validation exercises could be a useful addition to future stated preference studies, in particular those administered in an unsupervised setting.

Results from the extension tasks show that including information about the amount of time that patients have known about their prognosis has a clear impact on preferences – specifically, holding everything else constant, respondents are less likely to choose to treat a patient if they have known about their illness for two years than if they had only just found out about their illness. This is consistent with findings from the preference study which suggested that the observed tendency to give priority to the end of life patient may be driven by concerns about the patient's ability to "prepare for death" rather than the amount of time they have left to live *per se*. The fact that this "time with knowledge" attribute was clearly the main subject of the extension task instructions and questions is likely to have resulted in a focusing effect whereby respondents placed more importance on this attribute than they otherwise might have done. Furthermore, the extension tasks in our study do not allow for the elicitation of the strength of preferences at the individual level. Nevertheless, the data clearly show this is a factor about which many people hold clear preferences, and further investigation of these preferences is recommended.

The current criteria for determining whether a treatment should be a candidate for special consideration are that it should be indicated for patients with less than 24 months of life expectancy and that it should offer a life extension of at least three months. Hence, a treatment offering a 12 month life expectancy gain (and no quality of life gain) to patients with 24 months life expectancy at 50% quality of life without treatment (scenario ranked 34th in Table 7) would meet these criteria. The gains from this treatment amount to 0.5 QALYs. An alternative treatment, offering a 25% quality of life gain (and no life expectancy

gain) to the same patients (scenario ranked 81st) would also deliver a treatment gain of 0.5 QALYs, but would not meet the criteria for being eligible for special consideration. The results of this study indicate that the profile representing the former treatment would be considerably more likely to be chosen (1.17%) than the profile representing the latter treatment (0.62%). This suggests that the focus on life extensions and absence of quality of life improvements in the criteria may be consistent with public preferences.

An examination of the impact of marginal changes in any of the attribute levels from the reference case profile suggests that amending the life expectancy without treatment criterion would not have a major effect on utility. The predicted probability of choosing a profile involving a life expectancy gain of three months is much the same regardless of whether the patient's life expectancy without treatment is three, 24 or 60 months. By comparison, a profile involving a life expectancy gain of six months is considerably more likely to be chosen than an otherwise identical profile involving a life expectancy gain of three months.

Overall, the results of this study do not suggest that the cut-offs implied by the existing NICE supplementary end of life policy require amending, and in fact call into question whether such a policy of giving higher priority to end of life treatments than to other types of treatments is supported by the public at all, particularly if the health gains offered by the treatments being "de-prioritised" are larger than those offered by the end of life treatments. Of course, there may be reasons other than social preferences for retaining the end of life policy, such as concerns about whether existing methods of technology appraisal are able to capture the value of the health gains deemed important by end of life patients (Garau et al, 2011). Nonetheless, this study indicates that when asked to make decisions about the treatment of hypothetical patients with relatively short life expectancies, most people's choices are driven by the size of the health gains offered by treatment.

APPENDIX

A.1 Text of On-Screen 'Information Sheet'

Note: this information was displayed to respondents after they had entered the survey but before they had been given the opportunity to give their informed consent to take part.

Health care priority setting preference project

You are being invited to take part in a research project. Before you decide whether you wish to take part, it is important for you to understand why the research is being done and what it will involve. Please read the following information carefully and discuss it with others if you wish. Take time to decide whether or not you wish to take part. Thank you for reading this.

What is the project's purpose?

The purpose of the project is to find out what the general public thinks about a range of hypothetical scenarios where the health service has to choose which types of treatment to allocate funding to. Better understanding of public preferences will help organisations such as the NHS to make decisions about which treatments to provide.

Why have I been chosen?

We are seeking to survey around 4,000 members of the general public.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be asked to complete an informed consent form and you can still withdraw at any time. If you decide to stop, then any information that you have provided will be destroyed. You do not have to give a reason for not taking part.

What will happen to me if I take part?

If you agree to take part, you will complete an online survey. The survey will involve looking at hypothetical scenarios in which a health care decision maker must allocate resources to one of two treatments for ill health. You will also be asked some questions about yourself.

What do I have to do?

You will be asked to answer a series of questions involving hypothetical scenarios. There are no right or wrong answers – we are simply seeking your views.

What are the possible disadvantages and risks of taking part?

Some participants may feel uncomfortable when asked to think about scenarios involving illness and death. However, previous research in this area has shown that participants are generally interested and engaged when taking part in these types of exercises. Remember – you are free to withdraw from participating at any time.

What are the possible benefits of taking part?

You will be contributing to research that will help health care decision makers to better understand the preferences of members of the general public.

What if something goes wrong?

Should you wish to raise a complaint about any aspect of the study, please send this to isabel@valuedopinions.co.uk.

Will my taking part in this project be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. All survey responses will be anonymised, so you will not be identified in any reports or publications.

What will happen to the results of the research project?

The results of the project will be written up in a report for the National Institute for Health and Clinical Excellence (NICE), and will be published in academic journals and presented at conferences. You will not be identified in any reports or publications. The anonymised data collected during the course of the project may be used for additional or subsequent research and analysis.

Who is organising and funding the research?

The project is being organised by Allan Wailoo, Aki Tsuchiya and Koonal Shah, of the University of Sheffield's School of Health and Related Research. It has been funded by NICE.

Who has ethically reviewed the project?

The project has been reviewed by the School of Health and Related Research Ethics Committee at the University of Sheffield.

Contact for further information

For further information about this survey, please contact Professor Aki Tsuchiya (a.tsuchiya@sheffield.ac.uk; 0114 222 0710).

If you wish to seek further information about the topics covered in this project, you may find it helpful to get in touch with the Dying Matters Coalition, a group set up by the National Council for Palliative Care. You can find information, resources and details of organisations providing support and counselling on their website, http://www.dyingmatters.org. To speak to someone for cancer support over the telephone, you may call the Macmillan Support Line: 0808 808 0000 (free).

A.2 Consent Form

Note: respondents were required to tick all four boxes in order to proceed to the main part of the survey.

Title of Project: Health care priority setting preference project

Name of Researcher: University of Sheffield

☐ I confirm that I have read and understand the project information provided and have had the opportunity to discuss my participation with others.

☐ I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.

☐ I understand that my responses will be anonymised before analysis. I give permission for my anonymised responses to be accessed by researchers in this project and in future research projects.

☐ I agree to take part in the above research project.

If you would like any further information about the research please contact Professor Aki Tsuchiya (a.tsuchiya@sheffield.ac.uk; 0114 222 0710)

To continue to the survey please click Continue below, or click the Decline consent button if you do not wish to take part in the survey.

A.3 Instructions (for standard tasks)

--

Thank you for agreeing to take part in this research project.

The main survey consists of 12 questions about hypothetical scenarios. Once you have completed these questions, you will be asked some further questions about yourself.

--

We are going to show you some hypothetical scenarios involving patients who are affected by illness. We will use the survey to ask you which patients you think the health service should treat.

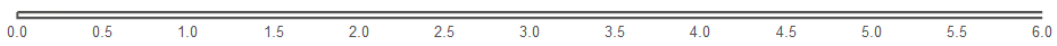
--

Illnesses and medical treatments affect people's health and how long they live.

Different illnesses affect people's health and how long they live in different ways; and different treatments offer different types of benefits.

We are going to use pictures to show these differences in illnesses and treatments. On the following pages, we will explain how the pictures work.

--



We can represent time with a line starting from 0 and going on to the right into the future.

Let's suppose that someone will live for 6 years from today. This can be shown by the line going from 0 years to 6 years.

--

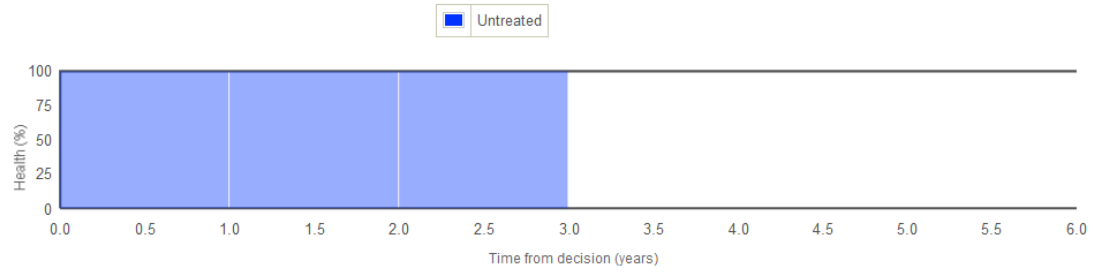
We can also show how good someone's health is using a health scale, where **'dead' is 0%** and **'full health' is 100%**.

Of course, full health for a young person may be different from full health for an elderly person. But to keep things simple, we show full health for everyone as 100%.

Someone who has health problems would have a health level of less than 100%.

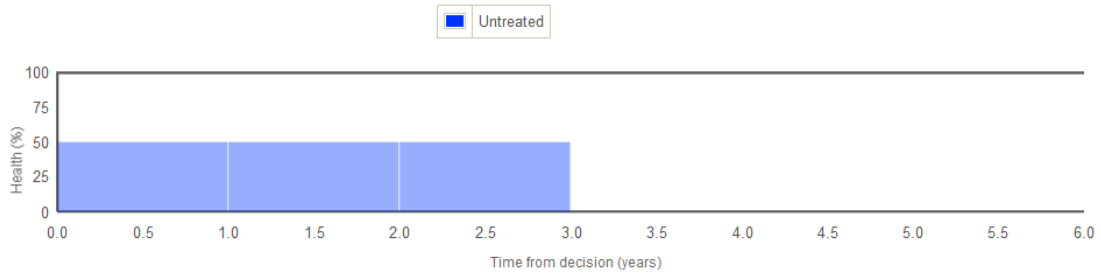
Suppose there is a health state which involves some health problems. If patients tell us that being in **this health state for 2 years** is equally desirable as being in **full health for 1 year**, then we would describe someone in this health state as being in **50% health**.

--



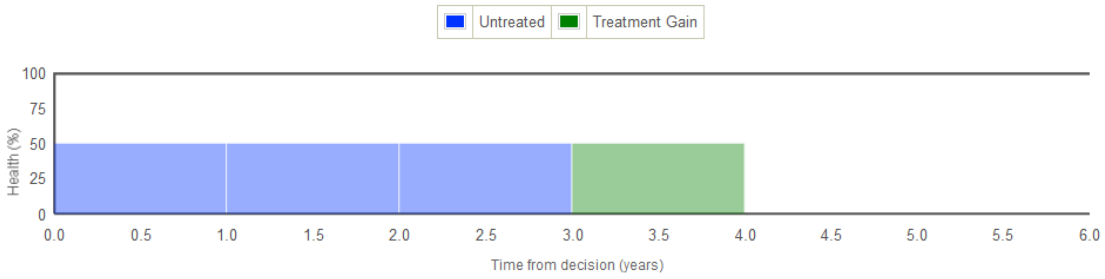
The blue area shows someone with an illness that gives the patient **3 years to live** from today, without treatment. This is shown by the end of the area at 3 years.

Note that the level of health is 100%, which represents full health. This means that although the illness leads to death in 3 years, it does not affect the patient's general health during those 3 years.



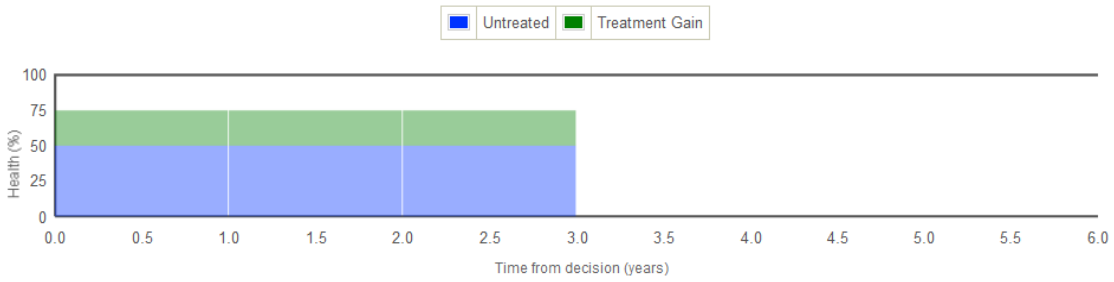
This blue area shows another illness. Without treatment, the patient shown here will live for 3 years in **50% health**, and then they will die.

--



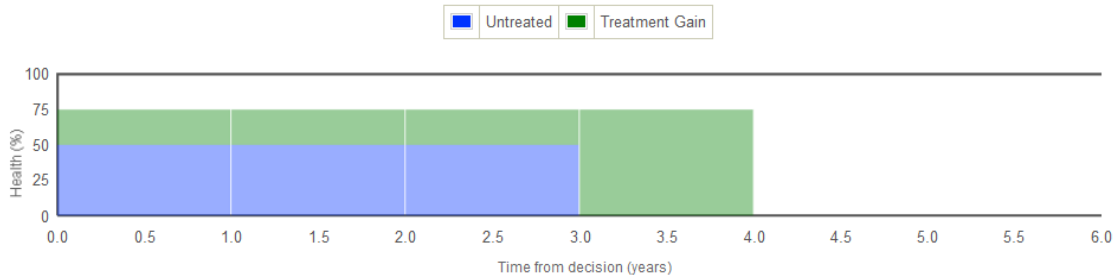
The green area shows a treatment for that illness. The treatment shown here gives the patient an **extra 1 year of life** at the same level of health (50%).

--



This treatment restores some of the patient's health (to 75%) but does not extend their life.

--



This treatment improves the patient's health to 75% AND gives them an extra 1 year of life.

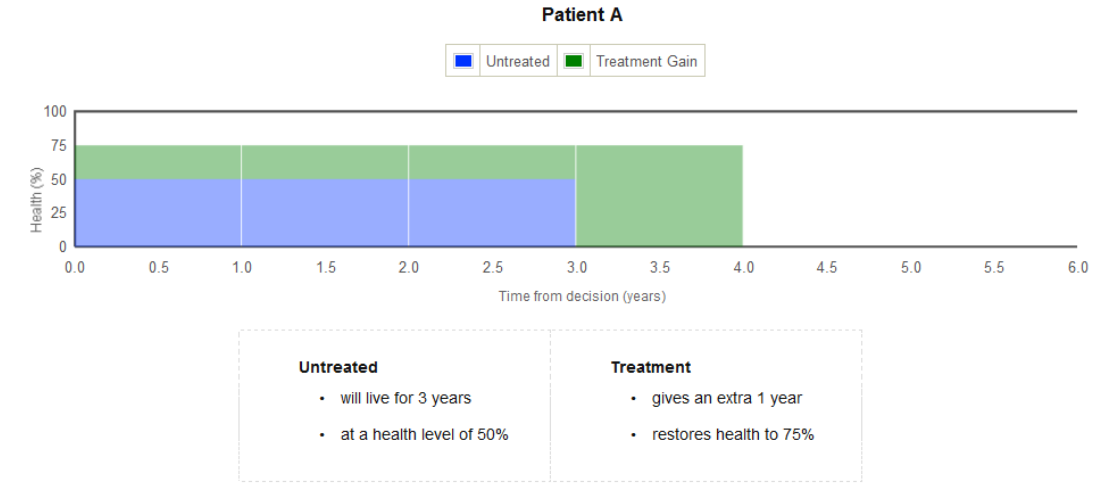
--

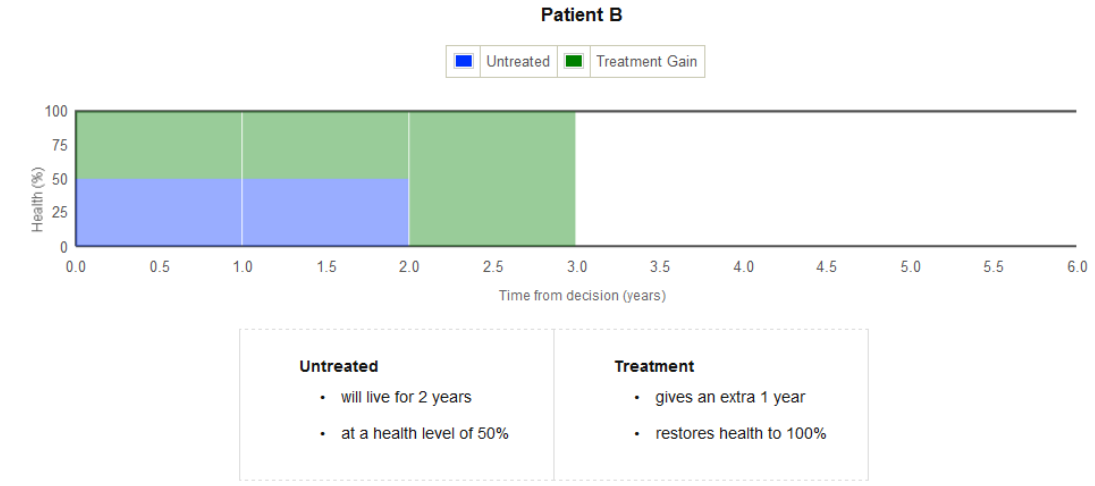
In the following questions you will asked to consider the situations of 2 hypothetical patients - patient A and patient B.

The patients will have different illnesses that affect their level of health and length of life in different ways.

The treatments available will also affect their health and length of life in different ways.

Scroll down to see an example of how the information about patient A and patient B will be shown in the questions.





--

No other information about the patients is available, except that they are **both adults**. You should therefore consider them to be equal in all other respects.

We want you to assume that the health service has only enough funds to treat one of the two patients, and that **there are no alternative treatments available**.

Furthermore, the nature of the illnesses is such that further treatment will not be possible if either patient is not treated today - this is the only opportunity for treatment.

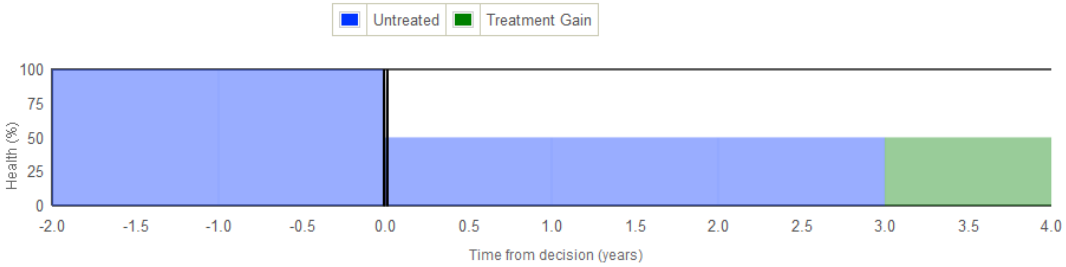
We want you to tell us which patient you think should be treated.
There are no right or wrong answers - we are simply seeking your view.

--

A.4 Instructions (for extension tasks)

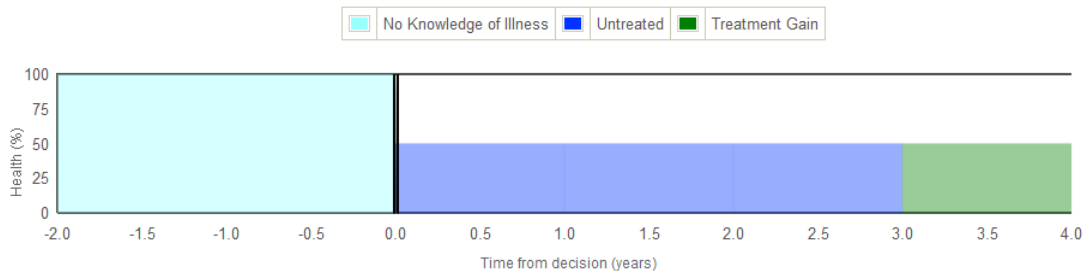
The next 2 questions will require you to consider slightly different scenarios.
Just as before, the patients will have different illnesses that affect their health and length of life in different ways.

But in these scenarios, **one of the patients has known about their illness for some time** while the other patient has only just learned of their illness.



This patient was told **2 years ago** that they have **5 years to live**. This means that from today, they have **3 years to live**, unless they receive treatment.

Note that the blue area to the left of 0 years is at 100% health. This means that up until today, the illness has not affected the patient's general health.



This patient has **just been told about their illness**. From today, they will live for 3 years before dying, unless they receive treatment.

The light blue area to the left of 0 years shows that the patient had **no knowledge of their illness** up until today.

Once again, we want you to tell us which patient you think should be treated.

ABBREVIATIONS AND DEFINITIONS

DCE	Discrete choice experiment
EEPRU	Policy Research Unit in Economic Evaluation of Health and Care Interventions
LE	Life expectancy
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
QALY	Quality adjusted life year
QOL	Quality of life

REFERENCES

Baker, R. et al. (2010) Weighting and valuing quality-adjusted life-years using stated preference methods: Preliminary results from the Social Value of a QALY project. *Health Technology Assessment*. 14(27), 1-162.

Bridges, J.F.P. et al. (2011) Conjoint analysis applications in health--a checklist: A report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value in Health.*14, 403-413.

Carlsson, F. and Martinsson, P. (2003) Design techniques for stated preference methods in health economics. *Health Economics*. 12(4), 281-294.

DH (Department of Health). (1997) *The new NHS: Modern, dependable*. London: Department of Health.

DH (Department of Health). (2001) *Involving patients and the public in healthcare: a discussion document.* London: Department of Health.

DH (Department of Health). (2010) *Impact assessment of proposal for a cancer drug fund.* London: Department of Health.

Dolan, P. and Tsuchiya, A. (2011) Determining the parameters in a social welfare function using stated preference data: An application to health. *Applied Economics*. 43(18), 2241-2250.

Dolan, P. et al. (2005) QALY maximisation and people's preferences: A methodological review of the literature. *Health Economics*. 14(2), 197-208.

Garau, M. et al. (2011) Using QALYs in cancer: A review of the methodological limitations. *Pharmacoeconomics.* 29(8), 673-685.

Green, C. (2007) *Justice, fairness and equity in health care: Exploring the social value of health care interventions.* PhD Thesis. Southampton: University of Southampton.

Green, C. and Gerard, K. (2009) Exploring the social value of health-care interventions: A stated preference discrete choice experiment. *Health Economics.* 18(8), 951-976.

Lancsar, E. and Louviere, J. (2006) Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics.* 15(8), 797-811.

Lancsar, E. and Louviere, J. (2008) Conducting discrete choice experiments to inform healthcare decision making: A user's guide. *Pharmacoeconomics.* 26, 661-677.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000) *Stated choice methods: Analysis and applications.* Cambridge: Cambridge University Press.

NICE (National Institute for Health and Clinical Excellence). (2009a) *Appraising life-extending, end of life treatments*. London: National Institute for Health and Clinical Excellence.

NICE (National Institute for Health and Clinical Excellence). (2008) *Guide to the methods of technology appraisal*. London: National Institute for Health and Clinical Excellence.

NICE (National Institute for Health and Clinical Excellence). (2009b) *End of life treatments: Summary response to consultation*. London: National Institute for Health and Clinical Excellence.

NICE (National Institute for Health and Clinical Excellence). (2009c) *Appraising life-extending, end of life treatments.* Revised July. London: National Institute for Health and Clinical Excellence.

NRS (National Readership Survey). (2010) Lifestyle data. Available at: http://www.nrs.co.uk/lifestyle.html [Accessed 5 December 2012].

ONS (Office for National Statistics). (2001) *Census: Standard area statistics (England and Wales)*. London: Office for National Statistics. Available at: http://www.neighbourhood.statistics.gov.uk/dissemination/ [Accessed March 2012].

ONS (Office for National Statistics). (2011) Internet access - households and individuals. *Statistical Bulletin*. 31 August. London: Office for National Statistics. Available at: http://www.ons.gov.uk/ons/dcp171778_227158.pdf [Accessed 5 December 2012].

Rawlins, M.D. and Culyer, A.J. (2004) National Institute for Clinical Excellence and its value judgments. *British Medical Journal.* 329(7459), 224-227.

Richardson, J. and McKie, J. (2005) Empiricism, ethics and orthodox economic theory: What is the appropriate basis for decision-making in the health sector? *Social Science & Medicine*. 60(2), 265-275.

Roberts, T. et al. (1999) Public involvement in health care priority setting: An economic perspective. *Health Expectations.* 2(4), 235-244.

Ryan, M. et al. (2001) Eliciting public preferences for healthcare: A systematic review of techniques. *Health Technology Assessment.* 5(5), np.

Ryan, M., Gerard, K. and Amaya-Amaya, M. (2008) *Using discrete choice experiments to value health and health care*. Dordrecht: Springer.

Shah, K.K. (2009a) Severity of illness and priority setting in healthcare: A review of the literature. *Health Policy*. 93(2-3), 77-84.

Shah, K.K., Tsuchiya, A. and Wailoo, A.J. (2012) *Valuing health at the end of life: An empirical study of public preferences.* Presentation at 3rd Joint CES/HESG Workshop. Aix-en-Provence. 11-13 January 2012.

Spalek, T.M. and Hammad, S. (2005) The left-to-right bias in inhibition of return is due to the direction of reading. *Psychological Science.* 16(1), 15-18.

Witt, J., Scott, A. and Osborne, R.H. (2009) Designing choice experiments with many attributes. An application to setting priorities for orthopaedic waiting lists. *Health Economics*. 18(6). 681-696.

YouGov. (2012) *Panel methodology*. London: YouGov plc. Available at: http://www.yougov.co.uk/about/about-methodology.asp [Accessed March 2012].