Social Science & Medicine 124 (2015) 48–56

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Valuing health at the end of life: A stated preference discrete choice experiment

Koonal K. Shah [a, b, *], Aki Tsuchiya [b, c], Allan J. Wailoo [b]

[a] Office of Health Economics, UK
[b] School of Health and Related Research, University of Sheffield, UK
[c] Department of Economics, University of Sheffield, UK

ARTICLE INFO

Article history:
Available online 13 November 2014

Keywords:
UK
End of life
NICE
Public preferences
Severity
Priority setting
Discrete choice experiment
Health economics

ABSTRACT

A source of debate in the field of health care priority setting is whether health gains should be weighted differently for different groups of patients. The debate has recently focused on the relative value of life extensions for patients with short life expectancy. However, few studies have examined empirically whether society is prepared to fund life-extending end-of-life treatments that would not meet the reimbursement criteria used for other treatments.

A web-based discrete choice experiment was conducted in 2012 using a sample of 3969 members of the general public in England and Wales. The study design was informed by the National Institute for Health and Care Excellence's supplementary policy for the appraisal of life-extending end-of-life treatments. The choice tasks involved asking respondents which of two hypothetical patients they would prefer to treat, assuming that the health service has enough funds to treat only one of them. Conditional logit regressions were used for modelling.

Choices about which patient to treat were influenced more by the sizes of treatment gains than by patients' life expectancy without treatment. Some respondents appear to support a health-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, we find little evidence that members of the general public prefer to give higher priority to life-extending end-of-life treatments than to other types of treatment. 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.

1. Introduction

Economic evaluation is used to estimate the efficiency of health technologies and subsequently to inform decisions about whether those technologies should be reimbursed. A common approach is to measure the health benefits of a given technology in terms of quality-adjusted life years (QALYs) (Weinstein and Stason, 1977). The cost-effectiveness of the technology can be expressed as cost per QALY gained. Decisions about whether to reimburse the technology can then be guided by comparing the cost-effectiveness of that technology to some threshold value that reflects displaced activities (Towse et al., 2002).

If it is assumed that the principal objective of health care is to maximise population health using available resources (Culyer, 1997) and that the QALY is an acceptable measure of health benefit, it follows that health care resources should be prioritised so as to maximise the total number of QALYs gained. This 'QALY-maximisation' rule (Dolan et al., 2005) entails distributive neutrality – it does not incorporate concerns for how the benefits are distributed across individuals.

However, maximising health may not be the only purpose: health care systems may also have other objectives, such as reducing health inequalities. As well as evaluating the evidence on cost-effectiveness, agencies carrying out health technology appraisals are often expected to make and apply social value judgements, about what is appropriate and acceptable for society.

* Corresponding author. Office of Health Economics, 7th floor Southside, 105 Victoria Street, London SW1E 6QT, UK.
E-mail address: kshah@ohe.org (K.K. Shah).

http://dx.doi.org/10.1016/j.socscimed.2014.11.022

Reviews by Schwappach (2002) and Dolan et al. (2005) show that people value QALYs differently depending on how they are distributed, the characteristics of the patients receiving them, and the characteristics of the health effects.

Health care decision makers in a number of countries have been considering whether and how to weight health gains to account for equity considerations, such as concern for those whose health prospects are poorest. In the Netherlands, for example, broad consensus has been reached to use the principle of 'proportional shortfall' as the basis for equity weighting (van der Wetering et al., 2013). This involves giving priority to patients who will lose the greatest proportion of their remaining health expectancy due to their condition. In England and Wales, the National Institute for Health and Care Excellence (NICE), the organisation responsible for producing advice on the use of health technologies in the National Health Service, has introduced a policy that effectively gives higher priority to life-extending, 'end-of-life' treatments than to other types of treatments. This constitutes a departure from the Institute's 'reference case' position (NICE, 2013) whereby all QALYs are deemed to be of equal social value, regardless of to whom they accrue and the context in which they are enjoyed. This paper examines society's preferences regarding the prioritisation of life-extending end-of-life treatments. The NICE policy is used as an example and as the framework for the study design. However, the issues explored have relevance in all countries seeking to understand the extent of societal support for giving priority to patients with short life expectancy.

1.1. NICE's end-of-life policy

In January 2009, NICE issued supplementary advice for appraising life-extending, end-of-life treatments (NICE, 2009a). This advice constitutes an explicit departure from the reference case position above. 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).

The current criteria, enshrined in NICE's updated methods guide (NICE, 2013), are set out below; if met, the Appraisal Committee is asked to consider the impact of giving greater weight to the treatment gains achieved in the later stages of disease.

1. The treatment is indicated for patients with short life expectancy (normally less than 24 months).
2. The treatment offers an extension to life compared to current NHS treatment (normally at least three additional months).
3. The treatment is licensed, or otherwise indicated, for small patient populations (normally less than 7000 patients).

In response to concerns expressed during the consultation 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), a few studies of people's preferences regarding end-of-life have been undertaken in the UK (Linley and Hughes, 2013; Brazier et al., 2013) and elsewhere (Olsen, 2013; Pinto-Prades et al., 2014). The findings are mixed, with evidence of support for an end-of-life premium reported by Brazier et al. (2013) and Pinto Prades et al. (2014) but not by Linley and Hughes (2013) or Olsen (2013). Notwithstanding these recent additions to the empirical literature, the evidence remains limited and there have been calls for further exploration of the issues (Green, 2011).

1.2. Objectives

The primary aim of this study is to investigate the extent to which the policy of giving higher priority to life-extending end-of-life treatments (as defined by NICE) than to other types of treatment is consistent with the stated preferences of members of the general public in England and Wales. Preliminary studies, reported elsewhere (Shah et al., 2011; 2014), tested the proposed methods and found weak evidence of public support for giving priority to end-of-life patients, all else being equal. A further aim is to add to the growing literature on public preferences regarding the prioritisation of health care, which can be used to support an 'empirical ethics' approach to allocating health care resources (Richardson and McKie, 2005).

2. Methods

2.1. Framework

There are many stated preference techniques that can be used to elicit public preferences regarding health care priority setting (Ryan et al., 2001). Health economists typically prefer choice-based techniques that reflect the view that the value of something is measured by how much one is willing to trade or sacrifice to obtain it. One such technique, the discrete choice experiment (DCE), produces quantitative trade-offs between different factors based on hypothetical choices (Louviere et al., 2000). DCEs are typically implemented in surveys comprising several 'choice sets', each containing competing alternative 'profiles' described using '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 relative contribution of each of the attribute levels to overall utility (Lancsar and Louviere, 2008).

DCE data are modelled within a random utility framework, which assumes 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 it as random. Assuming that the random terms are independently and identically distributed extreme value, the conditional logit model can be used to estimate the probability of alternative i being chosen from the complete set of alternatives ($j = 1, \ldots, J$):

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

2.2. Attributes and levels

The selection of attributes and levels (Table 1) was based on NICE's criteria (above) and informed by the findings of our preliminary studies (Shah et al., 2011; 2014). 'Life expectancy without

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

treatment' and 'life expectancy gain from treatment' form the basis for criteria 1 and 2. For life expectancy without treatment, a level representing the cut-off of 24 months was included, as well as two levels smaller and two levels larger than this cut-off (three months, 12 months; 36 months, 60 months). Larger levels were considered but omitted due to concerns about how the lives would be displayed visually using the computer-based diagrams. Similarly, the current 'life expectancy gain from treatment' cut-off of three months was included, as well as two smaller and two larger levels (one month, two months; six months, 12 months). In addition, 0 months was included in order to examine preferences for end-of-life treatments that offer no life extension.

The inclusion of quality-of-life attributes was driven by the finding in the preliminary studies that many respondents appeared to favour the prioritisation of quality-of-life-improving treatments over life-extending treatments (Shah et al., 2011; 2014). 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 wide range of levels (Baker et al., 2010a), our piloting work indicated that this may be challenging for respondents to interpret. We therefore included only two levels for the 'quality-of-life without treatment' attribute: 50% and 100%. The concept of '50% health' was explained as follows: "Suppose there is a health state which involves some health problems. If patients tell us that being in this health state for two years is equally desirable as being in full health for one 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) offer no health improvement (0% gain); (ii) restore the patient to full health (50% gain); and (iii) offer some improvement but do not restore the patient to full health (25% gain).

Other potential attributes, such as the patient's age, were considered but eventually omitted from the final 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 DCEs, some researchers have suggested that when tasks become too complex respondents may not make trade-offs but instead adopt other decision heuristics or lexicographic decision rules (Witt et al., 2009). We therefore chose to focus on the attributes that are most salient to the policy context for NICE.

2.3. Experimental design

A full factorial design using the attributes and levels in Table 1 results in 5*2*6*3 = 180 possible profiles, but some combinations would result in implausible scenarios. The sum of quality-of-life without treatment and quality-of-life gain from treatment cannot exceed 100%. 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. Imposing these constraints left 110 profiles, with 5995 possible pairwise choices sets to select from.

Using the STATA software (StataCorp, 2013), 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 the estimation of both main effects and selected interaction effects (see Section 2.7 for details). All of the choice sets were checked for plausibility, and no manual alteration of the design was required.

There is little guidance in the literature on the optimal number of DCE tasks to ask each respondent. The social preference DCE studies reviewed by Green (2007) used between one and 18 choice sets per respondent; whilst in a review of 79 conjoint analysis applications in health, Marshall et al. (2010) report that the majority of studies used between seven and 15 choice sets. We opted to organise the 80 choice sets into eight blocks of 10 choices. All 80 choice sets were classified into one of 13 'choice types' (see Table 3) which were spread across the blocks. For example, the experimental design included choice sets in which one profile could be said to 'dominate' the other (choice type 1 in Table 3). We ensured that all of the blocks contained at least one but no more than two such choice sets. Apart from this manual distribution of choice types, the choice sets were assigned to blocks at random.

To control for potential bias due to the positioning of choice options (Spalek and Hammad, 2005), 'mirror' blocks were generated to match the eight blocks described above. These consisted of the same 10 choice sets but switched the labels and positions assigned to the two alternatives – i.e. the alternative 'patient A' in the original block choice set appears in the corresponding mirror block choice set as 'patient B' (and vice versa). Including these mirror blocks meant that there were a total of 16 different versions of the survey.

2.4. Questionnaire design and scenario presentation

The choice sets were included in a self-completion survey administered over the Internet. Adapting the design of an existing survey used in research elsewhere (Brazier et al., 2013), we presented the attributes levels for two hypothetical patients using a combination of diagrams and text descriptions (Fig. 1).

The survey began with instructions which introduced the diagrams showing how different illnesses and treatments affect people's health and life expectancy. Respondents were asked 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.

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 hope that a cure 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." Although such 'hope effects' may exist and influence people's choices regarding priority setting, we did not consider them to be pertinent to this study. The treatments typically considered under NICE's end-of-life policy tend not to offer life extensions that are long enough for the realistic possibility of cures being discovered and made available for use during the intervening period.

No indifference or 'status quo' option was offered, in accordance with best practice guidelines (Bridges et al., 2011). The patients, illnesses and treatments were described in generic terms (e.g. "patient A's illness") since the use of labels (e.g. "cancer") may induce emotional and biased responses. This 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.

The 10 standard DCE tasks were presented to respondents in a random order so as to ensure that order bias is not systematic across the sample. After these, respondents were presented with two further tasks which sought to examine the impact of introducing additional information about how long the patients had known about their illnesses (not reported in this paper).

After completing these tasks, respondents were asked tick-box questions about their background and health. Finally, they were invited to leave comments if they so wished.

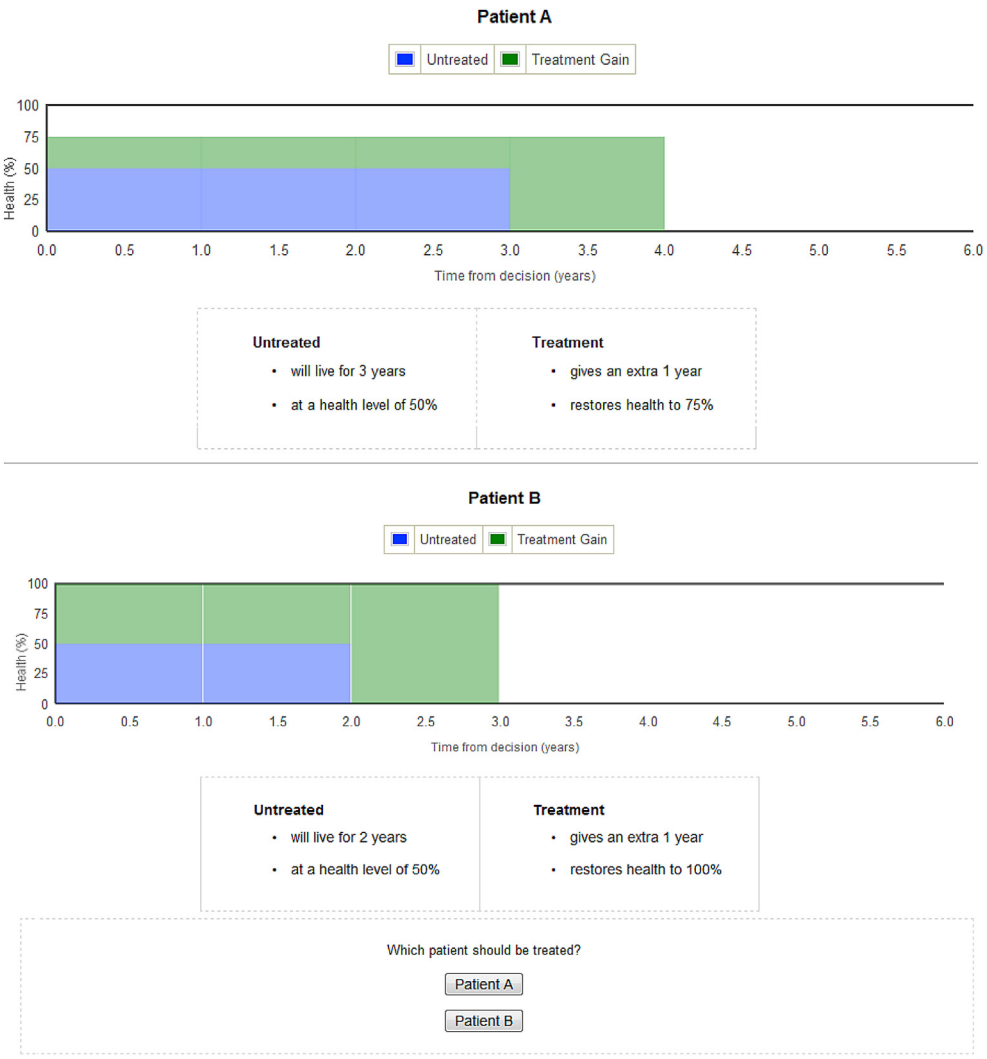

Fig. 1. Example of diagram and text used in the DCE tasks.

2.5. Sample/data collection

The sample comprised adult members of the general public in England and Wales, who were members of a panel of a market research agency, ResearchNow. 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 key observable characteristics. Individuals who had recently completed health-related surveys were not invited to take part. The average panel member completes six surveys per year. Respondents were compensated by way of 'reward points' which can be redeemed for gift vouchers or charity donations.

Web-based surveys offer a quick and cost-effective means of collecting a large amount of choice data, and can be custom-designed to present and elicit information in a clear, user-friendly manner. Interviewer-led survey administration is often preferred because the interviewer can explain the instructions more fully if required (Bridges et al., 2011) and respondents may be more attentive whilst under supervision. However, the use of interviewers can lead to forms of interviewer bias, which is not the case with web-based surveys.

Whilst the vast majority of households in the UK now have access to the Internet (Office for National Statistics, 2001), there remain concerns about the extent to which a sample of online panel members 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 allow interviewers into their homes for face-to-face interviews, for example, are similarly unlikely to be representative of the general population.

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. The ethics approval required us to destroy any information provided by respondents who did not complete the survey in full.

2.6. Piloting

The main study was preceded by a pilot, which used a convenience sample of 12 members of non-academic staff and post-graduate research students at the University of Sheffield. Face-to-face interviews were conducted by one of the authors (KKS) in which respondents completed the survey without assistance, and then answered verbal probing questions for feedback about the survey and approach.

The pilot was completed successfully, supporting the acceptability of the proposed methods. All of the respondents stated that they were able to understand and complete the questions without assistance.

2.7. Data analysis

Choice data were modelled using a random utility maximisation framework (Louviere et al., 2000) using STATA 11.2 software. As the data were binary choice data – '1' representing one option being chosen and '0' representing the other being chosen – conditional logit regressions were used.

The model estimated is of the form:

$$V = \beta_1 LE\ without\ treatment + \beta_2 QOL\ without\ treatment + \beta_3(QOL\ gain{*}LE\ without\ treatment) + \beta_4(LE\ gain{*}QOL\ without\ treatment) + \beta_5(LE\ gain{*}QOL\ gain) + \beta_6 EOL$$

The explanatory variables *LE* (life expectancy) *without treatment* and *QOL* (quality-of-life) *without treatment* represent the baseline health of the patients. The three interactions terms together make up QALY gains (quality-of-life improvement for a given level of life expectancy; life extension for a given level of quality-of-life; and life extension combined with quality of life improvement). These variables were treated as continuous. We also included an end-of-life dummy variable that took a value of 1 for profiles that would meet the NICE criteria for defining a life-extending end-of-life treatment (i.e. life expectancy without treatment of less than or equal to 24 months; life expectancy gain of greater than or equal to three months) and 0 otherwise. We assessed the performance of models both with and without the end-of-life dummy by examining the Akaike and Bayesian information criteria (Akaike, 1973; Schwartz, 1978) and conducting likelihood ratio tests.

The coefficients estimated in the model 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.

As described above, the probability of choosing a given profile from the complete set of profiles can be predicted from the model estimates. 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). This approach allows us to identify not only which attributes have statistically significant coefficients, but also which attributes are meaningful drivers of respondents' choices.

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 left a (optional) comment in the open-ended box at the end of the survey; and (iv) respondents who completed the questions much quicker than average. Family circumstances and personal experience of terminal illness were mentioned as influences on respondents' preferences in our earlier studies (Shah et al., 2011; 2014); whether a comment was left and how quickly the questions were completed may be indicators of respondent engagement. For each subgroup, we estimated the best fitting model and compared the results to those of the same model using the full sample.

Throughout our analyses we assumed a zero discount rate based on the lack of evidence from our preliminary studies (Shah et al., 2011; 2014) that time preference is a motivating factor for giving higher priority to end-of-life patients.

3. Results

Data collection was undertaken in early 2012. In total, 43,000 individuals were invited by email to take part in the survey, of whom 5308 clicked on the link to access the survey (response rate = 12.3%). Of the individuals who accessed the survey, 4008 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. Respondents who spent less than 3 min on the questions ($n = 39$) were excluded from the final data set, leaving 3969 respondents (39,690 pairwise observations).

Table 2 presents the background characteristics of the sample, which was representative of the general population in England and Wales with respect to age and gender (Office for National Statistics, 2011), and comprised a larger proportion of individuals in the highest and very lowest social grades (National Readership Survey, 2012-3).

Three hundred and eighty-nine 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

Table 2
Sample background characteristics.

	#	%	Gen pop[n]%
Total	3969	100	100
Gender			
Male	1942	49	49
Female	2027	51	51
Age			
18–24	404	10	11
25–44	1413	36	38
45–64	1228	31	31
65+	924	23	21
Social grade[a]			
A	221	6	4
B	1114	28	22
C1	1150	29	27
C2	645	16	22
DE	357	9	16
E	482	12	8
Household composition			
With children	963	24	
Without children	3006	76	
Education			
None beyond minimum school leaving age	889	22	
Beyond minimum school leaving age; no degree	1244	31	
Beyond minimum school leaving age; degree	1836	46	
Self-reported general health level			
Very good	1008	25	
Good	1958	49	
Fair	770	19	
Poor	210	5	
Very poor	23	1	
Experience of close friends or family with terminal illness			
Yes	2689	68	
No	1197	30	
Question skipped by respondent	83	2	

[a] Refers to the occupation/qualifications/responsibilities of the chief wage earner of the respondent's household; see National Readership Survey (2012–3).

Table 3
Average level of agreement, by choice type.

Choice type	No. choice sets	Description	Level of agreement (% respondents who chose patient X)
2	11	Both patients have the same LE/QOL without treatment. Patient X gains more LE and more QOL from treatment than patient Y.	92%
13	5	Patient X has shorter LE and higher QOL without treatment and gains more LE from treatment than patient Y.	85%
6	1	Patient X has lower QOL without treatment and gains more LE and more QOL from treatment than patient Y.	85%
1	14	Patient X has longer LE without treatment and gains more QOL from treatment than patient Y.	78%
12	2	Patient X has shorter LE without treatment and gains more LE from treatment than patient Y.	76%
9	4	Patient X has longer LE without treatment and gains more LE from treatment than patient Y.	74%
11	2	Patient X has longer LE and lower QOL without treatment and gains more QOL from treatment than patient Y.	72%
8	4	Patient X has shorter LE without treatment and higher QOL without treatment than patient Y. Both patients gain same amount of LE/QOL from treatment.	68%
7	5	Patient X has shorter LE and lower QOL without treatment and gains more QOL from treatment than patient Y.	68%
10	3	Patient X has longer LE and higher QOL without treatment than patient Y. Both patients gain same amount of LE/QOL from treatment.	66%
3	10	Patient X has shorter LE without treatment and gains more QOL from treatment than patient Y.	62%
4	10	Both patients have the same LE/QOL without treatment. Patient X gains more QOL from treatment; patient Y gains more LE from treatment.	59%
5	9	Patient X has lower QOL without treatment and gains more QOL and less LE from treatment than patient Y.	58%

expectations. We therefore included data for these respondents in the analysis.

3.1. Descriptive statistics of the choices made

For each choice set, we calculated the 'level of agreement' amongst respondents in terms of the proportion choosing the majority choice. Table 3 reports the average level of agreement for the choice sets belonging to each 'choice type'. The majority of respondents chose to treat the patient who gains more from treatment, regardless of whether that patient is better or worse off without treatment. Across the three choice sets in which the gains from treatment are the same for both patients and one patient is worse off without treatment in terms of both life expectancy and quality-of-life, the better-off patient was chosen 66% percent of the time.

Overall, there was a statistically significant tendency ($p < 0.01$) to choose to treat the alternative labelled patient B (the alternative appearing at the bottom of the respondent's screen).

3.2. Discrete choice model results

Table 4 reports the results of the conditional logit modelling. Note that the parameters have been coded such that that 1 year in full health is given a value of 1. We estimated two models – one with and one without the end-of-life dummy variable described in Section 2.7. The model *with* the end-of-life dummy performed better according to the Akaike and Bayesian information criteria and likelihood ratio tests ($p < 0.01$).

In both models, the coefficient for life expectancy without treatment is negative and statistically significant, which indicates that respondents are more likely to choose to treat the patient with shorter life expectancy without treatment, *ceteris paribus*. The coefficient for quality-of-life without treatment is not statistically significant in either model. The coefficients for the three interactions that make up QALY gains are all positive and statistically significant, and considerably larger in magnitude than the coefficient for life expectancy without treatment. The coefficient for the interaction between life expectancy gain and quality-of-life without treatment is substantially larger than the coefficient for the interaction between quality-of-life gain and life expectancy without treatment. This indicates that respondents' choices are driven by life extensions to a greater degree than by quality-of-life improvements. The coefficient for the end-of-life dummy is positive and statistically significant, which indicates that respondents are more likely to choose a treatment that meets the NICE criteria than one that does not.

To assist interpretation of the model results, Table 5 presents the utility scores based on the best fitting model for a selection of the profiles, as well as the predicted probability of choosing each profile from the full set of 110 profiles.

Table 4
Conditional logit modelling results.

Attribute	Model without end-of-life dummy[a]			Best fitting model (*with* end-of-life dummy)[b]		
	Coefficient	Std. error	*p*-value	Coefficient	Std. error	*p*-value
LE without treatment	−0.10715	0.00696	0.00	−0.06945	0.00736	0.00
QOL without treatment	−0.06357	0.04877	0.19	0.00051	0.04936	0.99
Interaction: QOL gain# LE without treatment	0.81567	0.01652	0.00	0.84535	0.01682	0.00
Interaction: LE gain# QOL without treatment	2.71342	0.05990	0.00	2.39408	0.06305	0.00
Interaction: LE gain# QOL gain	3.17557	0.10330	0.00	2.76204	0.10616	0.00
End-of-life dummy	N/A	N/A	N/A	0.37253	0.02510	0.00

[a] Akaike information criterion = 43,577; Bayesian information criterion = 43,623.
[b] Akaike information criterion = 43,358; Bayesian information criterion = 43,414.

Table 5
Estimated utility score and predicted probability of choice for the highest and lowest ranked profiles.

Rank	LE without treatment (mths)	QOL without treatment (%)	LE gain (mths)	QOL gain (%)	QALYs without treatment	QALYs gained from treatment	Utility	Prob.	Cumul. Prob.
1	60	50	12	50	2.500	3.500	4.3445	0.1351	0.1351
2	36	50	12	50	1.500	2.500	3.6380	0.0667	0.2018
3	12	50	12	50	0.500	1.500	3.3041	0.0477	0.2495
4	24	50	12	50	1.000	2.000	3.2848	0.0468	0.2964
5	60	50	6	50	2.500	3.000	3.0554	0.0372	0.3336
6	3	50	12	50	0.125	1.125	3.0392	0.0366	0.3702
7	3	100	12	0	0.250	1.000	2.7498	0.0274	0.3976
8	12	100	12	0	1.000	1.000	2.6977	0.0260	0.4237
9	60	50	12	25	2.500	2.000	2.5973	0.0235	0.4472
10	60	50	3	50	2.500	2.750	2.4109	0.0195	0.4668
11	12	50	12	25	0.500	1.000	2.4022	0.0194	0.4861
12	36	50	6	50	1.500	2.000	2.3490	0.0184	0.5045
13	36	50	12	25	1.500	1.500	2.3135	0.0177	0.5222
14	3	50	12	25	0.125	0.813	2.2958	0.0174	0.5396
15	24	100	12	0	2.000	1.000	2.2557	0.0167	0.5564
16	60	50	2	50	2.500	2.667	2.1961	0.0158	0.5721
17	36	100	12	0	3.000	1.000	2.1862	0.0156	0.5878
18	24	50	12	25	1.000	1.250	2.1716	0.0154	0.6031
19	60	100	12	0	5.000	1.000	2.0474	0.0136	0.6167
20	12	50	6	50	0.500	1.000	2.0150	0.0132	0.6299
–	–	–	–	–	–	–	–	–	–
10 most preferred profiles	2.750	0.600	0.875	0.375	1.438	2.038	3.1121	0.04668	
20 most preferred profiles	2.638	0.625	0.846	0.313	1.600	1.680	2.6677	0.03149	
55 most preferred profiles	2.268	0.600	0.558	0.277	1.305	1.201	1.7856	0.01570	
55 least preferred profiles	2.232	0.627	0.170	0.132	1.457	0.310	0.3081	0.00249	
20 least preferred profiles	2.488	0.625	0.117	0.050	1.644	0.093	0.0156	0.00179	
10 least preferred profiles	3.225	0.600	0.108	0.025	2.013	0.069	−0.0687	0.00164	
–	–	–	–	–	–	–	–	–	–
91	24	50	3	0	1.000	0.125	0.1606	0.0021	0.9662
92	12	50	0	25	0.500	0.250	0.1421	0.0020	0.9683
93	12	100	1	0	1.000	0.083	0.1306	0.0020	0.9703
94	12	50	2	0	0.500	0.083	0.1303	0.0020	0.9723
95	36	50	3	0	1.500	0.125	0.0912	0.0019	0.9742
96	3	50	0	50	0.125	0.125	0.0886	0.0019	0.9761
97	3	50	1	0	0.125	0.042	0.0826	0.0019	0.9780
98	24	100	1	0	2.000	0.083	0.0611	0.0019	0.9799
99	24	50	2	0	1.000	0.083	0.0609	0.0019	0.9817
100	60	100	2	0	5.000	0.167	0.0523	0.0018	0.9836
101	3	50	0	25	0.125	0.063	0.0357	0.0018	0.9854
102	12	50	1	0	0.500	0.042	0.0306	0.0018	0.9872
103	36	100	1	0	3.000	0.083	−0.0083	0.0017	0.9889
104	36	50	2	0	1.500	0.083	−0.0086	0.0017	0.9907
105	24	50	1	0	1.000	0.042	−0.0389	0.0017	0.9924
106	60	50	3	0	2.500	0.125	−0.0477	0.0017	0.9940
107	36	50	1	0	1.500	0.042	−0.1083	0.0016	0.9956
108	60	100	1	0	5.000	0.083	−0.1472	0.0015	0.9971
109	60	50	2	0	2.500	0.083	−0.1475	0.0015	0.9986
110	60	50	1	0	2.500	0.042	−0.2472	0.0014	1.0000

Note: all outcomes are undiscounted.

The highest ranked profiles all involve substantial treatment gains. All of the profiles ranked between 1st and 25th involve a life expectancy gain of 12 months and/or a quality-of-life gain of 50%. In contrast, the lowest ranked profiles mostly involve a small life expectancy gain 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 three months, respectively) appear at both the top and bottom of Table 5. 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. There is little difference between the highest and lowest ranked profiles in terms of QALYs without treatment – the key driver is the difference in the sizes of the QALY gains from treatment.

Fig. 2 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

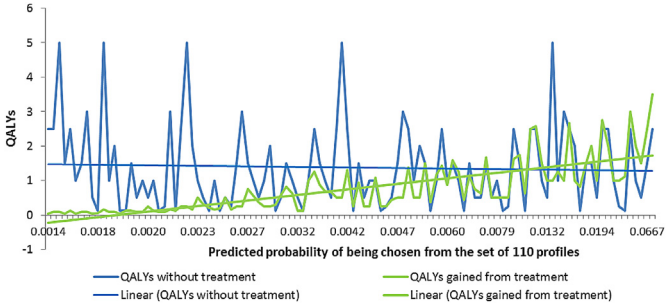

Fig. 2. Levels of QALYs without treatment/gained from treatment associated with all 110 profiles.

probabilities from the lowest (least preferred) to the highest (most preferred) profile. Whilst the patterns are noisy, the green linear trendline 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 trendline for QALYs without treatment is relatively flat, indicating that the number of QALYs without treatment does not have a major effect on the probability of the profile being chosen.

3.3. Subgroup analysis

As described in Section 2.7, we defined four respondent subgroups according to their responses to the background questions or to the ways in which they completed the survey. We estimated the best fitting model 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 (except for the coefficient for quality-of-life without treatment, which was not robust across models and never found to be statistically significant). 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 change our conclusions from the regression results.

4. Discussion

This study used a web-based survey to elicit the preferences of a large general public sample in England and Wales, representative in terms of age and gender, over a range of health care priority setting scenarios, focussing on social preferences regarding the prioritisation of treatments for patients with short life expectancy. The study used NICE's supplementary policy for the appraisal of life-extending end-of-life treatments as an example to explore issues that have potential relevance in other jurisdictions in which decision makers are considering whether health gains should be weighted differently for different groups of patients.

The results show that choices about which patient to treat are influenced more by the sizes of the gains achievable from treatment than by patients' life expectancy or quality-of-life in absence of treatment. There is certainly no indication that being at the 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 almost identical to that in the 55 profiles that are least likely to be chosen.

In line with the findings of our earlier work (Shah et al., 2011; 2014), the results show that people's preferences are heterogeneous. Although the conditional logit model is unable to account for the panel nature of the data, our analysis of choice frequencies at the individual respondent level showed that 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. These heterogeneous preferences do not appear to be well predicted by respondents' observable characteristics.

The main findings of this study do 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 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.

The outcomes examined in this study were not adjusted to account for any possible social time preference. Applying a positive discount rate would likely further strengthen the finding that respondents do not place special value on treating patients with short life expectancy, though we would expect the effect of discounting to be quite small given the relatively short timeframes included in the study design.

The web-based survey provided an efficient means of obtaining a large sample. The response rate observed is not unusual for a non-probability-based panel sample, and cannot easily be compared with response rates from studies using different modes of administration (for example, because many of the individuals invited to take part may not be active members of the panel) (Baker et al., 2010b). However, this mode of administration offers limited opportunity for debriefing with respondents about their experience of completing the survey (although our earlier studies and piloting were useful in this respect). The study was designed in such a way that the ranking of the profiles would not be expected to differ if some respondents failed to pay adequate attention to the choice tasks (e.g. making choices at random). Nevertheless, if respondents had failed to understand the instructions, then this could be problematic. For example, they may mistakenly believe that the tasks require 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. A useful addition to future stated preference studies, particularly those administered in an unsupervised setting, would be to design follow-up questions which can be used to check whether respondents agree with the policy implications of their responses to the questions. A high level of agreement would add legitimacy to the results.

The finding that respondents were more likely to choose to treat the alternative appearing at the bottom of the screen is consistent with findings reported elsewhere in the literature of bias due to the positioning of choice options (Spalek and Hammad, 2005).

NICE's current criteria for determining whether a treatment should be a candidate for special consideration are that it is indicated for patients with less than 24 months of life expectancy and that it extends life by at least three months. Hence, a treatment offering 0.5 QALYs through 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 would meet these criteria. An alternative treatment, also offering 0.5 QALYs through a 25% quality-of-life gain (and no life expectancy gain) to the same patients 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 more likely to be chosen (ranked 44th with a 0.51% probability of being chosen) than the profile representing the latter treatment (ranked 83rd; 0.23%). This suggests that the focus on life extensions and absence of quality-of-life improvements in the criteria may be consistent with public preferences, although some of the descriptive statistics analysis (Table 3) suggests otherwise.

An examination of the impact of marginal changes in any of the attribute levels from the profile representing a treatment that just meets the current NICE end-of-life criteria 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 36 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. We also tested an alternative model in which the end-of-life dummy was defined in terms of life expectancy without treatment but not life expectancy gain. The

coefficient for this end-of-life dummy was small and not statistically significant. This suggests that any observed support for the NICE policy amongst this sample requires that the policy includes a life extension criterion.

Overall, this study provides only limited evidence to suggest that members of the general public prefer to give higher priority to life-extending end-of-life treatments than to other types of treatment. 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 gains offered by treatment.

Acknowledgements

Financial support for this study was provided in part by a grant from the National Institute for Health and Care Excellence (NICE) through its Decision Support Unit. The views (and any errors or omissions) expressed are of the authors only. We would like to thank the anonymous reviewers for their helpful comments. We are also grateful for the contributions of Rachel Baker, John Brazier, Simon Dixon, Jenny Dunn, Clare Gardiner, Karen Gerard, Ulf Gerdtham, Anthony Hatswell, Julia Honea, Anju Keetharuth, Rachel Knowles, Oliver Rivero-Arias, Donna Rowen, Rob White; and to all of the respondents who took part in the study. We are particularly grateful for the contributions of Arne Risa Hole, who helped with the experimental design and analysis of the results.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest.

Baker, R., Bateman, I., Donaldson, C., et al., 2010a. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the social value of a QALY project. Health Technol. Assess. 14 (27), 1–162.

Baker, R., Blumberg, S., Brick, J.M., et al., 2010b. AAPOR report on online panels. Public Opin. Q. 74 (4), 711–781.

Brazier, J., Rowen, D., Mukura, C., Whyte, S., Keetharuth, A., Risa Hole, A., Tsuchiya, A., Shackley, P., 2013. Eliciting Societal Preferences for Burden of Illness, Therapeutic Improvement and End of Life for Value Based Pricing: a Report of the Main Survey. Report for the Department of Health. EEPRU, Sheffield/York.

Bridges, J.F.P., Hauber, A.B., Marshall, D., Lloyd, A., Prosser, L.A., Regier, D.A., Johnson, F.R., Mauskopf, J., 2011. Conjoint analysis applications in health – a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 14, 403–413.

Carlsson, F., Martinsson, P., 2003. Design techniques for stated preference methods in health economics. Health Econ. 12 (4), 281–294.

Culyer, A.J., 1997. Maximising the health of the whole community. In: New, B. (Ed.), Rationing: Talk and Action in Health Care. BMJ/King's Fund, London.

Dolan, P., Shaw, R., Tsuchiya, A., Williams, A., 2005. QALY maximisation andpeople's preferences: a methodological review of the literature. Health Econ. 14 (2), 197–208.

Green, C., 2007. Justice, Fairness and Equity in Health Care: Exploring the Social Value of Health Care Interventions. PhD Thesis. University of Southampton, Southampton.

Green, C., 2011. Looking for the 'values' to inform value-based pricing: a review of the empirical ethics (equity) evidence!. In: Paper Presented at the Health Economists' Study Group Meeting, Bangor.

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

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

Lancsar, E., Louviere, J., 2008. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 26, 661–677.

Linley, W.G., Hughes, D.A., 2013. Societal views on NICE, Cancer drugs fund and value-based pricing criteria for prioritising medicines: a cross-sectional survey of 4118 adults in Great Britain. Health Econ. 22 (8), 948–964.

Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. Stated Choice Methods: Analysis and Applications. Cambridge University Press, Cambridge.

Marshall, D., Bridges, J.F.P., Hauber, B., Cameron, R., Donnalley, L., Fyie, K., Johnson, F.R., 2010. Conjoint analysis applications in health — how are studies being designed and reported? Patient: Patient-Centered Outcomes Res. 3 (4), 249–256.

National Readership Survey, 2012-3. Social Grade. http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/ (accessed 18.09.14.).

NICE., 2009a. Appraising Life-extending, End of Life Treatments. NICE, London.

NICE., 2009b. End of Life Treatments: Summary Response to Consultation. NICE, London.

NICE., 2013. Guide to the Methods of Technology Appraisal 2013. NICE, London.

Office for National Statistics, 2001. Census: Standard Area Statistics (England and Wales). http://www.neighbourhood.statistics.gov.uk/dissemination/ (accessed 20.11.13.).

Office for National Statistics, 2011. Internet Access - Households and Individuals. http://www.ons.gov.uk/ons/dcp171778_227158.pdf (accessed 26.02.14.).

Olsen, J.A., 2013. Priority preferences: "end of life" does not matter, but total life does. Value Health 16 (6), 1063–1066.

Pinto-Prades, J.L., Sánchez-Martínez, F.I., Corbacho, B., Baker, R., 2014. Valuing QALYs at the end of Life. Soc. Sci. Med. 113, 5–14.

Rawlins, M.D., Culyer, A.J., 2004. National Institute for Clinical Excellence and its value judgments. Br. Med. J. 329 (7459), 224–227.

Richardson, J., McKie, J., 2005. Empiricism, ethics and orthodox economic theory: what is the appropriate basis for decision-making in the health sector? Soc. Sci. & Med. 60 (2), 265–275.

Roberts, T., Bryan, S., Heginbotham, C., McCallum, A., 1999. Public involvement in health care priority setting: an economic perspective. Health Expect. 2 (4), 235–244.

Ryan, M., Scott, D.A., Reeves, C., et al., 2001. Eliciting public preferences for healthcare: a systematic review of techniques. Health Technol. Assess. 5 (5).

Schwappach, D.L.B., 2002. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. Health Expect. 5, 210–222.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Statistics 6, 461–464.

Shah, K.K., 2009. Severity of illness and priority setting in healthcare: a review of the literature. Health Policy 73 (2–3), 77–84.

Shah, K.K., Tsuchiya, A., Wailoo, A.J., 2011. Valuing health at the end of life: an exploratory preference elicitation study. OHE Research Paper 11/06. Office of Health Economics, London.

Shah, K.K., Tsuchiya, A., Wailoo, A.J., 2014. Valuing health at the end of life: an empirical study of public preferences. Eur. J. Health Econ. 15 (4), 389–399.

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

StataCorp, 2013. Stata Statistical Software: Release 11. StataCorp LP, Texas.

Towse, A., Pritchard, C., Devlin, N., 2002. Cost-effectiveness Thresholds: Economic and Ethical Issues. King's Fund and Office of Health Economics, London.

Weinstein, M.C., Stason, W.B., 1977. Foundations of cost-effectiveness analysis for health and medical practices. N. Engl. J. Med. 296 (13), 716.

van der Wetering, E.J., Stolk, E.A., van Exel, N.J.A., Brouwer, W.B.F., 2013. Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. Eur. J. Health Econ. 14 (1), 107–115.

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