VALUE IN HEALTH 20 (2017) 224–229

Available online at www.sciencedirect.com

ScienceDirect

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

Comparing Increments in Utility of Health: An Individual-based Approach

Matthew Taylor, PhD[1],*, *Susan Chilton, PhD*[2], *Sarah Ronaldson, MSs*[3], *Hugh Metcalf, PhD*[2], *Jytte Seested Nielsen, PhD*[2]

[1]*York Health Economics Consortium, University of York, York, North Yorkshire, UK;* [2]*Newcastle University, Newcastle-upon-Tyne, Tyne and Wear, UK;* [3]*York Trials Unit, Department of Health Sciences, University of York, York, North Yorkshire, UK*

ABSTRACT

Background: Many economic evaluations of health care changes rely on quality-adjusted life year (QALY) estimates. Notably, though, the QALY approach values health *states* rather than *changes* in health states. Hence, a gain in utility of health is only indirectly valued through an *ex ante* preference elicitation of health states and the subsequent subtraction of health state values from one another, rather than being valued directly. There is therefore an underlying assumption that individuals, from an ex ante perspective *ceteris paribus*, would be indifferent between equal utility increments from health states with different baseline utilities. **Objective:** The aim of this paper is to develop a method that would allow us to measure individual-based preferences over utility increments from different baselines. We elicit our data using face-to-face interviews on a sample of UK individuals. **Results:** Overall, we find that gains of "equal" utility increments from different baselines are not found to be equally preferable by the individual. **Conclusions:** The results indicate that the subtraction approach could lead to sub-optimal resource allocations and suggest that a new approach which values health changes directly would better reflect individual preferences. This paper provides the foundations for a method to achieve this.
Keywords: QALY, quality of life, stated preference, utility.

Introduction

A cost-utility analysis involves a comparison of the quality-adjusted life-years (QALYs) gained for a given intervention with the incremental costs, where the QALY measure captures both quality and quantity gains [1]. Using the standard QALY procedure, the change in health utility resulting from a health care intervention is indirectly approximated through ex ante preference elicitation on health state and subsequent subtraction of health state values from one another [2]. This subtraction method will in practice be carried out using QALYs based on average health state values for several individuals (often a sample of the general population).

The method implicitly assumes that, on average and from an ex ante perspective ceteris paribus, individuals would be indifferent between an incremental gain in health utility from a health state associated with a higher utility and an equally sized incremental gain in utility from a health state associated with a lower utility. If gains of equal utility increments are not empirically found to be equally preferable by the individual on average, an assumption underlying the subtraction method is called into question and might suggest that a new approach, one that values health changes directly, would better reflect individual preferences.

The aim of this article is to develop a method that would allow us to measure individual-based preferences over utility increments from different baselines. Note that this issue differs from the one of equity when the potential recipients are different individuals [3–6]. We elicit our data using existing utility scores and face-to-face interviews on a sample of individuals in the United Kingdom. By asking individuals to make direct comparisons between equal utility increments (from health states with different baseline utilities), we obtain rankings at the individual level, allowing for a direct test of an assumption that underlies the subtraction methodology.

A test of the subtraction method is implicitly a test of the intraperson interval property, as defined by Torrance [7]: "a gain of equal utility increments anywhere on the scale should be equally preferable *for the individual whose utilities are being represented*. For example, if an individual's utilities are A = 0.2, B = 0.4, C = 0.6 and D = 0.8, the person should be indifferent to whether the change is from A to B or from C to D." According to Torrance

Conflicts of interest: The authors have indicated that they have no conflicts of interest with regard to the content of this article.

* *Address correspondence to:* Matthew Taylor, York Health Economics Consortium Ltd., Market Square, University of York, York YO10 5NH, UK.

E-mail: matthew.taylor@york.ac.uk.

1098-3015$36.00 – see front matter Copyright © 2017, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jval.2016.12.009

[7] and to the best of our knowledge, the intraperson interval property has, so far, not been tested.

Although not central to this article, we show that our data have the potential to measure the strength of such individual-based preferences and hence to indicate the size of the potential bias when using the subtraction method. We leave to the end of the article a discussion as to how the findings in this article may be translated into a health policy tool.

The article proceeds as follows: in section 2, we describe the empirical study and analytical approach; and section 3 presents the results; and the final section reflects on the implications for future research and policy formation.

The Framework for Comparing Increments in Utility

In the experiment, we follow the suggestion by Torrance [7] and test whether individuals looking through a "veil of ignorance," that is, not knowing which outcomes and choices might occur to them in the future, would agree in advance that a change from 0.2 to 0.4 is equally preferable compared with a change from 0.6 to 0.8. Hence, we compare two *changes* at a time (Gain X and Gain Y) from two proposed levels of an individual's baseline utility of health (A and C). The changes are represented using the visual analogue scale (VAS), the HUI Mark 3 (HUI:3) [8] descriptive system, or the EQ-5D-5L.

Using the VAS, we ask respondents for an individual assessment of the changes directly. However, for the HUI:3 and EQ-5D-5L, the health state change is described using the descriptive system, and associated individual scores are assumed to be the population scores generated by the HUI:3 algorithm [8] and the official crosswalk EQ-5D-5L values for the United Kingdom (http://www.euroqol.org). The ranking of the two changes constitute stage 1 of our analytical approach. By keeping Gain X fixed and varying Gain Y until the individual identifies his or her indifference point between the two utility increments from different proposed baselines, we also illustrate (as stage 2) how strength of individual preferences can be elicited for different gains and baselines, providing an indication of the size of the potential bias using the subtraction method.

EQ-5D is a widely used generic health status measurement method, and in the United Kingdom, preference values are derived using the time trade-off method. Throughout this article, we consider the elicited preference values to be a proxy for individual utility even though only preferences measured using standard gamble are based directly on von Neumann Morgenstern utility theory. The time trade-off method is framed in a riskless world [9], and because this article takes the existing application of the QALY framework as a starting point, our method correspondingly is designed in a world of certainty. However, it would be relatively straightforward to extend this method to a world of uncertainty.

Experimental Design

A "common" experiment was designed, varying only in terms of the method used to represent health states: the VAS, HUI:3, or EQ-5D-5L. Due to the potential for fatigue, if subjects were asked to answer all variants, the experiment was applied on two separate samples: sample I, students (from the Universities of York and Newcastle); and sample II, members of the general public (Newcastle-upon-Tyne). Subjects in sample I were randomized into either VAS_S or HUI_S, whereas subjects in sample II were randomized into either VAS_P or $EQ5D_P$ (Table 1).

Experimental Procedures

A total of 170 participants took part in the experiment. All experimental sessions were face-to-face interviews attended by only one participant and the interviewer. An example of the experimental instructions is included in the online Appendix (in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2016.12.009). Prior to the experiment itself, a piloting phase was carried out to test for comprehension. After the pilot, experimental instructions were further amended and more training exercises included. This process allowed us to verify that the experimental instructions were understood by as many participants as possible, a crucial feature of any new method, and, as much as is possible, to establish that subjects understood the concept of a health *change*.

Core procedures, questions, and so on, were identical across the samples. However, a more extensive introduction and explanation were required for sample II to ensure participants understood the task. Thus, although the experiment generates treatment-specific results, we refrain from any direct comparisons due to this difference in information communication. Nevertheless, we are still able to draw broad qualitative conclusions from the two sets of responses. After the warm-up and training exercises, the actual preference elicitation exercise was carried out on both sample I and II. This is described in the following section.

Eliciting Preferences Over Changes in Health States

Table 2 contains the full set of incremental changes offered to respondents. In each question, the respondent was shown two health changes, Gain X and Gain Y, described in terms of initial baseline health and a final health state. States corresponding to Gain X are labeled "A" and "B," and those corresponding to Gain Y are labeled "C and "D." The utility scores for these different states are contained in Table 2. It can be observed that the initial baseline state in Gain X is always better than the initial baseline state in the corresponding Gain Y.

In all variants (e.g., EQ5D; HUI; VAS), six questions were asked. In the question identifiers, the subscript refers to the specific utility gain (offered in both X and Y) that a respondent was asked to compare; that is, $Q1_{0.25}$ means that, in question 1, respondents had to choose between an increment of 0.25, in this case from baselines of 0.75 (A) and 0.5 (C). In $Q2_{0.25}$ and $Q3_{0.25}$, although the utility gains are the same, the initial baseline states differ. The same principle applies in the $Q4_{0.5}$ to $Q6_{0.5}$. In all questions, a respondent was asked to indicate whether X or Y offered the "better" change. These responses will be reported as stage 1 results. Examples of the visual presentation of Gain X are shown in the online Appendix (for $VAS_{S+P,}$ HUI and EQ5D; see the Appendix in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2016.12.009). It was explicitly emphasized that respondents should imagine themselves in the health states

Table 1 – Experimental design.

Sample	VAS	Method HUI:3*	EQ5D*
I	VAS_S n = 35	HUI_S n = 35	
II	VAS_P n = 51		$EQ5D_P$ n = 51

VAS, visual analogue scale; HUI:3, HUI Mark 3; EQ5D, EuroQol 5 Dimension 5 Levels.
Note: Subscript s denotes students and p denotes general public.

Table 2 – Comparisons used in the VAS_{P+S}, $HUI{:}3_S$, and EQ $5D_P$, Stage 1.

	VAS				HUI:3				EQ 5D			
	Gain X		Gain Y		Gain X		Gain Y		Gain X		Gain Y	
States	A	B	C	D	A	B	C	D	A	B	C	D
$Q1_{0.25}$	0.75	1.0	0.5	0.75	0.76	1.0	0.52	0.76	0.78	1.0	0.55	0.78
$Q2_{0.25}$	0.75	1.0	0.25	0.50	0.76	1.0	0.28	0.52	0.78	1.0	0.23	0.55
$Q3_{0.25}$	0.75	1.0	0	0.25	0.76	1.0	0	0.28	0.78	1.0	0	0.23
$Q4_{0.50}$	0.5	1.0	0.25	0.75	0.52	1.0	0.28	0.76	0.55	1.0	0.23	0.78
$Q5_{0.50}$	0.5	1.0	0	0.5	0.52	1.0	0	0.52	0.55	1.0	0	0.55
$Q6_{0.75}$	0.25	1.0	0	0.75	0.28	1.0	0	0.76	0.23	1.0	0	0.78

VAS, visual analogue scale; HUI:3, HUI Mark 3; EQ5D, EuroQol 5 Dimension 5 Levels.
Note: S, students; P, general public

described and choose the better one for themselves personally, as opposed to choosing on behalf of others.

After this, different alternatives to health state D were presented to the respondent. If Gain X is preferred to Gain Y, D is improved (and hence Gain Y becomes better), and if Gain Y is preferred to Gain X, D is made worse (and hence Gain Y becomes worse.). We denote D^* as the health state that makes the respondent indifferent between X and (the "new") Y, which we denote Y^*. D^* and Y^* will be reported as stage 2 results and discussed later.

To identify D^*, a total of 21 VAS cards were made available to the respondents (increments of 0.1). For HUI_S and $EQ5D_P$, 18 combinations were chosen to represent health-related quality-of-life utilities at equal intervals of around 0.06, ranging from 0 to 1. These utility levels were generated by the HUI:3 algorithm and taken from the official UK values for the EQ-5D-5L (see the Appendix Table A1–2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2016.12.009).

Care was taken to ensure that each of the 18 combinations dominated the adjacent lower card (i.e., either equal to or better than that for the corresponding health dimension). Due to this requirement, the utility of the initial baseline health states varies slightly across the different variants. Baseline health states are marked in italics (see the Appendix Table A1–2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2016.12.009). If the respondent did not explicitly indicate indifference between any two specific health gains, that is, did not identify a D^* health state, for the purposes of the subsequent analysis, we assumed a point of indifference exactly halfway between the two health gains where the preference changed from X to Y or Y to X. The ordering of questions was randomized. Participants were encouraged to ask questions throughout the interview with respect to procedures, and the interviewer was instructed to clarify any misunderstandings.

Analytical Approach

We report the experimental data in two stages. Stage 1 is based on whether each subject was indifferent between Gain X and Gain Y or whether one was preferred. The main research question we set out to test was RQ1: *Are individuals, on average, indifferent between "equal" increments in health from different baseline utility states?* Using unpaired, two-sided *t* tests, we tested whether the proportion of subjects who stated that Gain X and Gain Y were equally preferred was significantly different from the proportion who indicated a preference for either X or Y.

Table 3 – VAS_{S+P}; Stage 1 data and test results.

	Prefer Gain Y	Prefer Gain X	Indifferent	Two-sided *t* test Intraperson interval property (*P* value)	One-sided *t* test Prefer Y RQ1a (*P* value)	One-sided *t* test Prefer X RQ1b (*P* value)
VAS_S Sample I, n = 35						
$Q1_{0.25}$	30 (86 %)	1 (3 %)	4 (1%)	< 0.01	< 0.01	> 0.99
$Q2_{0.25}$	34 (97 %)	0 (0 %)	1 (0.3%)	< 0.01	< 0.01	> 0.99
$Q3_{0.25}$	34 (97 %)	1 (3 %)	0 (0%)	< 0.01	< 0.01	> 0.99
$Q4_{0.50}$	34 (97 %)	0 (0 %)	1 (0.3%)	< 0.01	< 0.01	> 0.99
$Q5_{0.50}$	35 (100%)	0 (0 %)	0 (0%)	< 0.01	< 0.01	> 0.99
$Q6_{0.75}$	35 (100%)	0 (0%)	0 (0%)	< 0.01	< 0.01	> 0.99
VAS_P Sample II, n = 51						
$Q1_{0.25}$	25 (49%)	15 (29%)	11 (22%)	< 0.01	$= 0.15$	$= 0.85$
$Q2_{0.25}$	25 (49%)	17(33%)	9 (18%)	< 0.01	$= 0.42$	$= 0.58$
$Q3_{0.25}$	29 (57%)	14 (27%)	8 (16%)	< 0.01	$= 0.18$	$= 0.82$
$Q4_{0.50}$	33 (65%)	14 (27%)	4 (8%)	< 0.01	< 0.01	> 0.99
$Q5_{0.50}$	32 (63%)	8 (16%)	11 (22%)	< 0.01	< 0.01	> 0.99
$Q6_{0.75}$	39 (76%)	8 (16%)	4 (8%)	< 0.01	< 0.01	> 0.99

VAS, visual analogue scale.
Note: S, students; P, general public

If subjects were not indifferent between the gains, we analyzed individual preferences in more details in RQ1a + b: *Do individuals on average prefer an "equal" increment in health from a poorer baseline utility state?* Using one-sided *t* tests, we tested whether the proportion of subjects who preferred Gain Y (i.e., an equivalent gain from a poorer starting point) was significantly different from the proportion of subjects who did not prefer Gain Y.

RQ1b: *Do individuals on average prefer an "equal" increment in health from a better baseline utility state?* For completeness, we used one-sided *t* tests to test whether the proportion of subjects who preferred Gain X (i.e., an equivalent gain from a better baseline) was significantly different from the proportion of subjects who did not prefer Gain X.

Taken as a whole, if we found that on average individuals rank "equal" increments differently, we could reject the intraperson interval property, which raises some doubt regarding the subtraction method. In addition, we report and examine stage 2 results (i.e., individuals' strength of preferences as elicited in the experiment). Hence, we report the average health state D* identified in the two samples and the associated Gain Y*.

Results

Tables 3 and 4 report the experimental stage 1 data for the six questions. We first state the number of individuals who indicated strict preference for either Gain Y or Gain X or were indifferent. In both samples and across all health state measurement methods, we rejected the hypothesis that Gain X and Gain Y were equally preferred, with P values < 0.01.

Turning to RQ1a, we tested whether participants prefer a change from a poorer baseline utility of health to an equally sized change from a better baseline utility of health state. The results from both the VAS_S and HUI_S are clear, and we cannot reject the hypothesis that these respondents prefer a change from a poorer health state. The results from sample II are more ambiguous. In VAS_P, when the utility Gain X is 0.5 or greater (Q4–Q6), we can draw the same conclusions. When the utility Gain X is 0.25 (Q1–Q3), however, we cannot reject the hypothesis that the participants were indifferent between the changes on the basis of a one-sided test. This pattern is not mirrored in $EQ5D_P$ where the only situation in which participants strictly preferred Gain X was from the worst utility state given in Q6.

In RQ1b, the one-sided alternative is that participants prefer a change from a good baseline utility to an equal change from a poor baseline utility state. The results from both VAS_S, VAS_P, and HUI_S are clear: This change is not preferred. In the $EQ5D_P$ results, however, we can identify that Gain X is strictly preferred or the subjects are indifferent between the health changes for 5 of the 6 questions; only in Q6 is Gain X not preferred.

Table 5 reports the stage 2 results. For a Gain X of 0.25 (Q1–Q3), we see that this is perceived by individuals to be equivalent to a Gain Y* ranging from 0.1 to 0.43 depending on sample, baseline, and elicitation method. A Gain X of 0.5 is equally preferred to a Gain Y* ranging from 0.24 to 0.54, whereas a Gain X of 0.75 is perceived to be equivalent to a Gain Y* in the range of 0.29 to 0.55. In general, we can see that Gain Y* is smaller than the corresponding Gain X in almost all situations, but given the small sample sizes and different elicitation methods used, we will only draw broad qualitative conclusions and not apply statistical inference at this time.

Overall, based on these findings, we can reject the hypothesis that participants in this data set comply with the intraperson interval property. Based on the results of RQ1a-b, we can infer that the main reason for this is explained by RQ1a; that is, individuals prefer an "equal" increment from a lower baseline utility state. In $EQ5D_P$, however, some subjects expressed a preference for a gain when the baseline health state was associated with higher utility.

Discussion

In this article, we set out to develop a method that would allow us to measure individual-based preferences over different utility increments. Our results show that in sample I, there was a clear tendency to prefer an equal utility increment from an initially more severe utility of health state compared to a better baseline state. Overall, the VAS results were similar in sample II, although the tendency to prefer incremental changes from a more severe

Table 4 – $HUI{:}3_S$ and EQ $5D_P$; Stage 1 data and test results.

	Prefer Gain Y	Prefer Gain X	Indifferent	Two-sided *t* test Intraperson interval property (*P* value)	One-sided *t* test Prefer Y RQ1a (*P* value)	One-sided *t* test Prefer X RQ1b (*P* value)
$HUI{:}3_S$ Sample I, n = 35						
$Q1_{0.25}$	22 (63 %)	8 (23 %)	5 (15 %)	< 0.01	= 0.01	= 0.99
$Q2_{0.25}$	22 (63 %)	11 (31 %)	2 (6%)	< 0.01	= 0.06	= 0.94
$Q3_{0.25}$	29 (83%)	6 (17 %)	0 (0%)	< 0.01	< 0.01	> 0.99
$Q4_{0.50}$	30 (86 %)	4 (11 %)	1 (3%)	< 0.01	< 0.01	> 0.99
$Q5_{0.50}$	29 (83 %)	6 (17 %)	0 (0%)	< 0.01	< 0.01	> 0.99
$Q6_{0.75}$	32 (91 %)	3 (9 %)	0 0 (%)	< 0.01	< 0.01	> 0.99
EQ $5D_P$ Sample II, n = 51						
$Q1_{0.25}$	23 (45%)	26 (51%)	2 (4%)	< 0.01	= 0.67	= 0.33
$Q2_{0.25}$	7 (14%)	41 (80%)	3 (6%)	< 0.01	= 0.99	= 0.01
$Q3_{0.25}$	16 (31%)	33 (65%)	2 (4%)	< 0.01	= 0.99	= 0.01
$Q4_{0.50}$	26 (51%)	23 (45%)	2 (4%)	< 0.01	= 0.67	= 0.33
$Q5_{0.50}$	16 (31%)	33 (65%)	2 (4%)	< 0.01	= 0.99	= 0.01
$Q6_{0.75}$	38 (75%)	11 (22%)	2 (4%)	< 0.01	< 0.01	> 0.99

HUI:3, HUI Mark 3; EQ5D, EuroQol 5 Dimension 5 Levels.
Note: S, students; P, general public

Table 5 – Stage 2 results: strength of preference for each gain from each baseline.

		VAS_S		VAS_P		HUI_S		$EQ\text{-}5D_P$	
	Gain X	D*	Gain Y*	D*	Gain Y*	D*	Gain Y*	D*	Gain Y*
$Q1_{0.25}$	0.25	0.66	0.16	0.72	0.22	0.73	0.21	0.74	0.19
$Q2_{0.25}$	0.25	0.37	0.12	0.49	0.24	0.51	0.23	0.63	0.35
$Q3_{0.25}$	0.25	0.10	0.10	0.24	0.24	0.15	0.15	0.43	0.43
$Q4_{0.50}$	0.50	0.53	0.28	0.61	0.36	0.64	0.36	0.72	0.42
$Q5_{0.50}$	0.50	0.24	0.24	0.37	0.37	0.29	0.29	0.54	0.54
$Q6_{0.75}$	0.75	0.42	0.42	0.53	0.53	0.43	0.43	0.55	0.55

VAS, visual analogue scale; HUI:3, HUI Mark 3; EQ5D, EuroQol 5 Dimension 5 Levels.
Note: S, students; P, general public
* D* denotes the health state that makes the respondent indifferent between X and (the "new") Y, which is denoted Y*

baseline was not significant when the utility score of the baseline health state in Gain X was high.

These results are similar to previous findings in other decision contexts where "social" values have been elicited. When potential recipients are different individuals and the questions relate more to equity concerns, it has been found that people prefer to give priority to severely ill patients [10,11]. In contrast, for the $EQ5D_P$ results, there was a tendency to prefer equal utility increments from the better baseline utility to incremental changes from an initially more inferior baseline. By varying one gain until the individual identifies indifference between two changes in utility of health from different proposed baselines, in stage 2 we also illustrate how not only individual rankings but also strength of individual preferences can be elicited.

On the whole, we find that gains of equal utility increments are not found to be equally preferable by the individual. By violating the intraperson interval property, our data raise some doubt regarding the subtraction method, which could lead to suboptimal resource allocations. This suggests that a new approach, which values changes in utility of health directly, would better reflect individual preferences. Our study has laid the foundations for a method that could address this question and, based on the stage 2 results, potentially determine a multiplier for changes in utility of health starting from a lower baseline level.

To illustrate, imagine a comparison of two treatments that both deliver an increment in utility of 0.25. One treatment would deliver an increment from 0.75 to 1 for the individual, whereas the other would provide an individual increment from 0.5 to 0.75. Our results (as reported in Table 5, VAS_S, $Q1_{0.25}$) would suggest multiplying the latter by the ratio between Gain X and Gain Y* (i.e., 0.25/0.16 = 1.6). In cost-utility analyses, the gain from the low baseline state would thereby be inflated to better reflect individual preferences. Hence, our approach provides a nonmonetary measure that could be used within the cost-utility framework. Moving toward eliciting willingness-to-pay for a change in health instead could be another way of estimating the benefit of a change in health directly, which could also take into account the results found in this article.

In the present article, our method was designed for a world of certainty, for reasons explained in the Introduction. Our method, however, can in principle be extended to an uncertain world, although this would undoubtedly be more cognitively challenging for respondents. In this case, respondents would be asked to choose between health gains in which the initial baseline health state and the new health states were described in terms of pairs of gambles, potentially resulting in a multiplier to be used in conjunction with EQ-5D scores elicited using the standard gamble.

A possible explanation for the difference in significance and results found between the two samples could be that the VAS method was better understood by participants and, therefore, led to more consistent results. However, as a direct valuation technique, the VAS method has been criticized for its lack of a theoretical foundation because it involves no choices and therefore no trade-offs [12,13].

We refrain from combining results from the different samples because one limitation of the current study is that population mean utility scores in $EQ5D_P$ and HUI_S were assumed to represent the utility score for the individual participant, whereas an individual assessment of the health changes was elicited using the VAS method. Preceding the elicitation task described here with an individual elicitation of utility scores would strictly speaking make the results of the different samples more comparable; this is a topic for further empirical investigation.

Our findings come with the caveat that they might be at least partially driven by biases in the heath state measurement method itself as well as by measurement error introduced by the method for comparing increments, in particular when eliciting strength of preferences. It is a question for future research whether standard elicitation methods can be adapted instead to minimize measurement bias and directly take into account the findings in this article and reflect them in the utility scores given to specific health states.

Based on our results generated from current health state elicitation methods, our findings might call into question the use of the subtraction method for assigning values to any change in health states delivered by an intervention, particularly for the range of health states within the study. Whether this holds for the entire range of utility improvements that could be delivered and/or is generalizable to the wider population is an open question and again one for further research. Likewise, a larger data set would be needed to test whether strength of preferences is significantly different across different baseline and measurement methods. Nevertheless, our study has laid the foundations for a method that could address this question empirically.

Conclusions

In the QALY approach, a gain in health is only indirectly valued through ex ante preference elicitation on health states and subsequent subtraction of average health state values from one another. Our results, however, raise some doubt regarding this subtraction method. Overall, we find that gains of equal utility increments from different baselines are not found to be equally preferable by the individual. Hence, our data violate an assumption that underlies the subtraction method.

In this case, a direct method would be preferable, although, up to now, one has not been available. This article provides the foundations for such an approach (i.e., one that allows us to measure individual-based preferences over different utility changes directly). Adopting this method in policy making would therefore have the potential to make cost-utility analyses of health care initiatives more representative of underlying preferences than current practice.

Source of financial support: Funding for this study was provided by the Newcastle University Business School, York Health Economics Consortium, and the Danish Strategic Research Council (under the IMPROSA project).

Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2016.12.008) or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

REFERENCES

[1] Drummond M, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. Oxford, UK: Oxford University Press, 2005.
[2] Nord E, Enge AU, Gundersen V. Is the value of treatment proportional to the size of the health gain? Health Econ 2010;19:596–607.
[3] Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. Health Econ 2005;14:197–208.
[4] Gyrd-Hansen D. Investigating the social value of health changes. J Health Econ 2004;23:1101–16.
[5] Lancsar E, Wildman J, Donaldson C, et al. Deriving distributional weights for QALYs through discrete choice experiments. J Health Econ 2011;30:466–78.
[6] Exel JV, Baker R, Mason H, et al. Public views on principles for health care priority setting: Findings of a European cross-country study using Q methodology. Soc Sci Med 2015;126:128–37.
[7] Torrance GW. Utility measurement in healthcare. The things I never got to. Pharmacoeconomics 2006;24:1069–78.
[8] Horsman K, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI®): concepts, measurement properties and applications. Health Qual Life Outcomes 2003;1:54.
[9] Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ 2002;11:447–56.
[10] Nord E. The person-trade-off approach to valuing health care programs. Med Decis Making 1995;15:201–8.
[11] Ubel PA. How stable are people's preferences for giving priority to severely ill patients? Soc Sci Med 1999;49:1049–55.
[12] Johannesson M, Jonsson B, Karlsson G. Outcome measurement in economic evaluation. Health Econ 1996;5:279–96.
[13] Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ 2006;15:653–64.