

Exploring a new method for deriving the monetary value of a QALY

Carl Tilling¹ · Marieke Krol^{2,3} · Arthur E. Attema³ · Aki Tsuchiya^{1,4} · John Brazier¹ · Job van Exel³ · Werner Brouwer³

Received: 11 December 2014 / Accepted: 5 August 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Several studies have sought to determine the monetary value of health gains expressed as quality adjusted life years (QALYs) gained, predominantly using willingness to pay approaches. However, willingness to pay has a number of recognized problems, most notably its insensitivity to scope. This paper presents an alternative approach to estimate the monetary value of a QALY, which is based on the time trade-off method. Moreover, it presents the results of an online study conducted in the Netherlands exploring the feasibility of this novel approach. The results seem promising, but also highlight a number of methodological problems with this approach, most notably nontrading and the elicitation of negative values. Additional research is necessary to try to overcome these problems and to determine the potential of this new approach.

Keywords Time trade-off method · QALY · Willingness to pay

JEL Classification I10

✉ Arthur E. Attema
attema@bmg.eur.nl

¹ School of Health and Related Research, University of Sheffield, Sheffield, UK

² Institute for Medical Technology Assessment, Erasmus University, Rotterdam, The Netherlands

³ Department of Health Policy and Management, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

⁴ Department of Economics, University of Sheffield, Sheffield, UK

Introduction

In light of increasing health care expenditure and the limited resources available, decision makers face the challenge of determining the appropriate allocation of these resources over health programs. To help determine an appropriate distribution, economic evaluations provide information on the costs and effects of health technologies. Within economic evaluations, health effects are typically expressed in quality adjusted life years (QALYs). The QALY is an outcome measure of health benefit that combines length of life with quality of life. Quality of life is typically expressed on a scale from zero to one, where zero represents a health state equivalent to being dead and one represents perfect health [1]. By expressing health outcomes on a common unit of measurement, outcomes can be compared across different health programs, which is helpful for making reimbursement decisions. Several countries use these economic evaluations to inform allocation decisions [2].

One intriguing question regarding the use of the outcomes of economic evaluations, typically taking the form of a ratio of incremental costs per QALY gained, is when to consider a technology to offer ‘value for money’ and hence to implement or fund it. That final judgment requires some threshold against which to evaluate the cost-per-QALY ratio. Different ideas regarding the nature and meaning of this threshold, and therefore the decision making context, exist [3–5]. It can represent either the amount a society is willing to pay for a QALY from private consumption or, in a fixed budget system, the opportunity cost of a QALY from displaced health care activities [6]. This paper, however, is concerned with the former interpretation, i.e. the societal value of a QALY. For either interpretation, introducing the technology can be deemed

cost-effective, i.e. welfare improving [3], only if the ratio of costs per QALY remains below the value of that QALY.

Finding the societal value of a QALY is a delicate matter and by no means easy. Recently, two large studies aimed at finding the monetary value of the QALY (MVQ) have been conducted: the UK Social Value of a QALY (SVQ) project [7] and an international study involving nine European countries (EuroVaQ, [8, 9]). Both studies experienced large difficulties related to the methodological approaches chosen. Like most other studies conducted to determine MVQ [10–18], these studies used a contingent valuation (CV) method to estimate the willingness to pay (WTP) for a health improvement (either life extension or quality of life improvement). However, CV has a number of recognised problems, most notably its insensitivity to scope [19], strategic behaviour [20], protest responses [21] and the restriction of personal income [22].

Insensitivity to scope (or scale) arises if respondents' WTP does not change in response to the size of the outcome being valued. Evidence of insensitivity to scope concerns economists because it contradicts the fundamental principles of neo-classical theory since 'more is better' consumers should be prepared to sacrifice more money to obtain more of some good (albeit at a diminishing rate). From a practical perspective, if WTP does not vary with the size of the gain, any possible MVQ could be obtained by varying the size of the gains. Although some studies found evidence against insensitivity to scope [23–25], quite a few others found evidence in support of scope insensitivity [19, 26–28].

Besides insensitivity to scope, a concern with WTP is the opportunity for strategic behaviour, depending on the payment vehicle (free-riding) [20, 29]. This may occur in two directions. Firstly, if respondents think they will actually have to pay the amount they reveal, they may underbid. Alternatively, if respondents do not believe they will actually have to pay their stated WTP amount, but they want to influence the provision of the good in question, they might overbid. There is limited available evidence regarding strategic behaviour in WTP studies in the health care field [20].

Another issue with WTP is the incidence of protest answers. For instance, people who indicate a WTP of zero may do so for several reasons, such as that they do not know their true WTP, they actually have a zero value for the good (real zeros), or they are protesting against the exercise or payment for the good or outliers [18, 21, 29]. In a contingent valuation survey of Dalmau-Matarrodona (aimed at determining the value of day case surgery as opposed to inpatient treatment) as much as 35 % of the respondents stated a zero WTP [21]. One-third of these were classified as protest zeros. An additional problem with WTP is the influence of ability to pay. This influence may

be considered particularly problematic in the context of health care, where the emphasis is on accessibility and equity [30]. In WTP, personal income acts as a budget constraint. The approach of WTP thus allows the wealthy to state higher values for the goods/treatments they prefer than the poor, which (depending on the use of the results) could bias health care decisions. This has led some to argue that WTP is a valid method only if we accept that the current distribution of income is appropriate [22], although Donaldson [31] has argued that one can correct and adjust WTP towards any desired distribution.

In the light of these issues with WTP, it seems useful to examine ways other than common WTP studies to obtain monetary valuations of health gains. This paper presents such an alternative approach¹ based on a time trade-off (TTO) exercise of income with health held constant at perfect health, which can be used to estimate the MVQ. We present the methods and theory underlying this experimental approach and some results from an online feasibility study in the Netherlands.

Methods

TTO is a widely used choice-based method of health state preference elicitation. Buckingham and Devlin [32] have outlined how the TTO method can be interpreted in the theoretical context of Hicksian utility theory and hence comply with welfare economic principles in a similar fashion to WTP derived through CV. We designed a TTO exercise in which respondents trade off length of life (in a certain health state) and income. People are thus asked to indicate their indifference between living longer (in health state X) with a lower income and living shorter (in health state X) but with a higher income. From these trade-offs, the implicit monetary value placed on a QALY can be derived. This is explained in more detail below.

Data and questionnaire

Data were gathered as part of a study seeking to determine whether respondents in TTO exercises consider the effects the states might have upon their income [33, 34]. Data were gathered through an online self-complete questionnaire administered in the Netherlands. Invitations were sent out to a subset of an existing panel of potential survey respondents in order to obtain a representative sample of 300 members of the Dutch general public. Respondents between the ages of 18 and 65 only were selected as questions about income were seen as being most relevant

¹ Note that a new approach may suffer from (some of) the same limitations.

for people in this age bracket. The data collection was performed by an online market research company (Survey Sampling International; <http://www.surveysampling.com>). Following a number of background questions including age, sex, marital status and self-assessed health by means of a visual analogue scale (VAS), respondents were presented with 14 different TTO exercises (see Tilling et al. [33] for more details). Two of these TTO exercises were relevant for this study, in which health is replaced by income so the trade-off becomes between longevity and income rather than longevity and health.

The wording of the first question was as follows:

TTO 1: Trading years to avoid an income loss in perfect health (equivalent variation of a loss)

“You can live for 10 years in perfect health with $(100 - Y) %$ of your current annual income for each year and then die or you can live for a shorter period of time in perfect health with your current annual income for each year and then die. How many years with your current income do you consider to be equally good as living 10 years with $(100 - Y) %$ of your current income?”

“I find living... years and... months with my current income equally good as 10 years with $(100 - Y) %$ of my current income”.

The indifference curves representing the trade-off are shown in Fig. 1. The x -axis represents length of life and the y -axis represents income. Each indifference curve represents a level of utility that can be achieved by different combinations of longevity and income, where

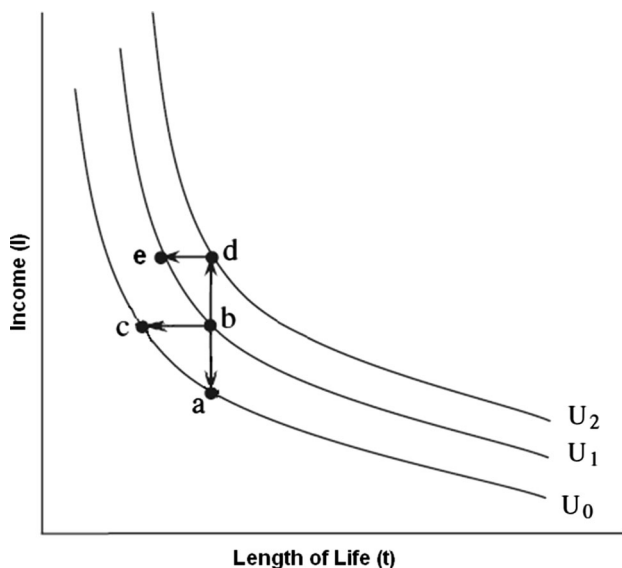


Fig. 1 Equivalent income loss and compensating income gain (adapted from Buckingham and Devlin [32], p 1151)

$U_2 > U_1 > U_0$. The first option asks the respondent to consider a move from point b on indifference curve U_1 (10 years in perfect health with current income) to point a on indifference curve U_0 (10 years in perfect health with less than current income). The second option involves a move from point b to point c (X years in perfect health with current income), which is again on U_0 . The respondent must thus specify a decrease in longevity that is equivalent to a decrease in income, both of which causing a decrease in utility from U_1 to U_0 .

The second question also asks respondents the decrease in longevity that would be required to compensate for an increase in income, but the reference point differs:

TTO2: Trading years to achieve an income gain in perfect health (compensating variation of a gain)

“You can live for 10 years in perfect health with your current annual income for each year and then die or you can live for a shorter period of time in perfect health with $(100 + Y) %$ of your current annual income for each year and then die. How many years with $(100 + Y) %$ of your current income do you consider to be equally good as living 10 years with your current income?”

“I find living... years and... months with $(100 + Y) %$ of my current income equally good as 10 years with my current income”.

Referring again to Fig. 1, the first option is to stay at point b on indifference curve U_1 (10 years with current annual income). Note, in TTO2 the first option is on a higher indifference curve (U_1) than in TTO1 (U_0), because income is set at current annual income. An increase in income (to a value greater than current income) places the individual onto a higher indifference curve U_2 , at point d . The respondent must then specify a decrease in longevity that returns them to their original indifference curve at point e on U_1 .

In other words, respondents have to consider an equivalent variation for a loss in TTO1. Equivalent variation is ‘the amount of money a consumer would pay to avert a price increase’ [35]. In TTO1, the consumer is faced with a fall in income of $X %$, which is essentially the same as an increase in prices. They are then asked how many years of life (rather than how much money) they would pay to avoid this ‘price increase’. Similarly, TTO2 can be viewed as asking a form of compensating variation. Compensating variation is ‘the amount of additional money a consumer requires to reach his initial level of utility after a change in prices [35]. For a drop in prices, the amount of additional money compensation will be negative. TTO2 corresponds essentially to a compensating variation that identifies the number of years payable that would let the individual

maintain the initial level of utility after a drop in prices, or increase in income. Essentially these questions can be interpreted as a WTP and a WTA question, respectively. However, while standard WTP (WTA) questions ask people to trade money for an improvement (deterioration) in length of life or health, these questions asked people to trade length of life for an improvement in income. Respondents were thus paying in years of life.

Three income change levels (Y) were used: in version 1 of the questionnaire 20 % was used, in version 2 40 % and in version 3, 60 %. Respondents were randomised to one of the three income change levels, which they then received in both TTO1 and TTO2. Since the survey was administered in an online self-complete fashion there was no iterative process. Respondents were simply asked to state how many years with higher income, was equivalent to 10 years with lower income. All respondents first received TTO1, followed by TTO2.

Analysis of responses

Our responses can be interpreted and analysed only after assuming the form of the utility function of respondents over health and income. In the current paper, given its explorative nature, we assume a simple additive function $W(\cdot)$ over health (H) and income (Y):

$$W(H, Y) = U(H) + Y \quad (1)$$

That is, individuals derive utility (U) from their health state H and have a linear utility function over income. This specification was used earlier by Eeckhoudt et al. [36]. The advantage of this function is that it becomes straightforward to elicit a monetary value of the utility of perfect health. Moreover, an additive way of thinking when answering this task is cognitively less demanding and appears more plausible than a multiplicative way of thinking.

To see how the results from these questions can be used to derive an MVQ, imagine that a respondent facing TTO1 states that 9 years with normal annual income of €100,000 is equivalent to 10 years with 80 % of this income, so €80,000. Using prospective lifetime income values and assuming a zero discount rate, this point of indifference gives us the following information:

$$10U(\text{PH}) + \text{€}800,000 = 9U(\text{PH}) + \text{€}900,000 \quad (2)$$

$$10U(\text{PH}) - 9U(\text{PH}) = \text{€}900,000 - \text{€}800,000 \quad (3)$$

$$U(\text{PH}) = \text{€}100,000 \quad (4)$$

where PH is perfect health. In reality, it is likely that the utility from a year in perfect health will be higher when combined with a higher amount of income, whereas we assume a constant marginal rate of substitution between

health and income. Relaxing this assumption would require us to estimate an indifference curve across a range of values, which is beyond the scope of this first empirical exploration of the method.

The compensating *gain* data from TTO2 is analysed in a similar fashion to the equivalent *loss* data in TTO1. Consider a respondent who is indifferent between 10 years with their current income and 9 years with 120 % of their current income. Their income is, once again, €100,000 per year:

$$10U(\text{PH}) + \text{€}1,000,000 = 9U(\text{PH}) + \text{€}1,080,000 \quad (5)$$

$$10U(\text{PH}) - 9U(\text{PH}) = \text{€}1,080,000 - \text{€}1,000,000 \quad (6)$$

$$U(\text{PH}) = \text{€}80,000 \quad (7)$$

Respondent income

In order to determine the level of “current annual income” for each respondent, respondents were asked to choose the income bracket within which their monthly income fell within the background characteristics questions. For our analysis, these income brackets were converted into numerical values using the mid-point of each bracket [37]. For respondents in the lowest income bracket, an income of two-thirds of the upper limit of the bracket was used. For respondents in the highest income bracket, an income of 1.5 of the lower income limit of the bracket was assumed [37].

Non-traders

Some respondents did not trade any time in any of the TTO exercises. For these respondents, calculating an MVQ becomes problematic because the left hand side of Eq. (2) becomes 0, meaning that the equation would give an indeterminate value. If such responses occur and are a protest against the exercise, this poses questions about the feasibility of the exercise. If such responses are a meaningful statement of preference for a seemingly infinite preference for life over income then this does not mean the exercises are infeasible, but rather that the calculation method above is not capable of calculating a finite MVQ for such individuals based on these meaningful responses. A respondent with lexicographic preferences of this nature would not give up any length of life to increase their income. In the context of the equivalent variation for a loss question, the decrease in income facing the respondent (from current income to less than current income) does not decrease their utility; therefore, they stay on their initial indifference curve, implying their equivalent loss in longevity is zero, because otherwise their utility would drop below this level.

It should be noted that non-trading in the equivalent variation for a loss or compensating variation for a gain question does not necessarily mean that the indifference curve is perfectly vertical; it just means that the curve is sufficiently steep that the utility gained from the increase in income is less than the amount of utility that would be lost through giving up the smallest amount of longevity possible (the smallest unit of trade was 1 month). Furthermore, non-trading for a given income change level does not mean that the entire indifference curve is vertical (or sufficiently steep), it only determines the slope of the indifference curve between the two income points on the y axis that the respondent is being questioned on.

Regardless of whether non-trades are protest responses or a true reflection of lexicographic preferences, if an individual calculation method (i.e. calculate an MVQ for each individual and then compute the mean) is to be used, then non-traders must be excluded, because their answers would imply an infinite MVQ [38]. Therefore, we excluded all ‘extreme non-traders’ (i.e. respondents who did not trade across all 14 TTO questions of the questionnaire). An alternative is to use an aggregate approach, where we divide the sum of the income differences by the sum of the life time reductions (‘ratio of means’) [38]. This can be compared to the disaggregate approach (‘mean of ratios’), where one divides the income difference by the reduction of life time for each respondent. These approaches are likely to generate different results, especially because we have a lot of non-traders, who could be included in the aggregate approach but not in the disaggregate approach. The results from both approaches are presented.

Negative values

One further problem of our approach is the potential generation of negative MVQ values. For TTO1, if the percentage of life years the respondent is prepared to give up is larger than the percentage income loss they are faced with, their MVQ will be negative. For example, if a respondent is faced with 20 % income loss and is willing to trade more than 2 years of life to avert this, her MVQ value will be negative (while if she trades exactly 2 years, her MVQ value will be zero). In other words, for the 20 % loss respondents, trading more than 2 years leads to a negative MVQ; for the 40 % (60 %) loss respondents, this holds for trades of more than 4(6) years. For TTO2 the relationship is not linear. For a 20 % (40 %, 60 %) gain in income, trades of more than 1.67 (2.86, 3.75) years result in negative values. In the disaggregate approach, we truncated negative MVQ values at 0. In the aggregate approach we left the number of years traded unchanged.

Results

Data were available from 321 members of the Dutch general public. After exclusion of 80 ‘extreme non-traders’ the relevant sample size fell to 241. The sample consisted of slightly more males than females, and 41.5 % of the sample was not employed. Just under one-half of the sample had children, and less than one-half of the sample was married. The mean VAS score for own health was 0.75. The results of χ^2 tests showed that background characteristics did not differ significantly across the three versions of the questionnaire. Only employment differed slightly across the versions, with a smaller proportion of respondents in version 2 being in employment than in the other two versions.

Even after excluding the ‘extreme non-traders’, a substantial number of the respondents did not trade time in the compensating gain and/or equivalent loss questions. The proportion of non-traders in the equivalent loss questions decreased as the level of loss increased: 72 % were non-traders for 20 % loss, 54 % for 40 % loss and 45 % for 60 % loss. In the compensating gain questions the proportion of non-traders was fairly constant across the three income gain levels: 63 %, 65 % and 64 % were non-traders for 20 %, 40 % and 60 % gain, respectively. Trading off life duration for income increases hence invokes a large degree of non-trading.

Table 1 shows the mean number of years respondents were willing to trade, in both the compensating *gain* and equivalent *loss* questions. Looking at the values including the non-traders, for two of the income change levels, respondents were willing to trade more years to avoid an income *loss* than they were to achieve an income *gain*. However, these differences were significant only for the 60 % income change level (at the 1 % level). The median values were 0 in all but one case, which was a product of the large numbers of non-traders. Mann–Whitney rank-sum tests were performed to compare the values for the different income levels, both for equivalent *loss* and compensating *gain* values. Number of years traded was significantly different between 20 % and 40 % equivalent *loss* (5 % level) and between 20 % and 60 % equivalent *loss* (1 % level). For the equivalent loss questions the standard deviations generally increased as the level of loss increased, while no clear relationship was observed for the gain questions.

Table 2 shows the MVQ estimates calculated according to the disaggregate approach. As described, this approach excludes all non-traders, resulting in a much smaller sample for analysis. The mean MVQ values ranged from €17,439 to €65,957. A larger proportion of respondents gave negative MVQ values (which were truncated to zero

Table 1 Number of years traded

	Version 1: 20 % (<i>n</i> = 78)		Version 2: 40 % (<i>n</i> = 80)		Version 3: 60 % (<i>n</i> = 83)	
	Loss	Gain	Loss	Gain	Loss	Gain
Number of years traded to either avoid an income loss or achieve an income gain						
Mean	0.99	1.47	1.81**	1.33	2.45	1.51***
SD	2.23	2.96	2.74	2.63	3.28	2.89
Median	0	0	0	0	1	0

** Significant at 5 % level, *** significant at 1 % level

Table 2 Monetary value of the QALY (MVQ) values calculated at the individual level (excluding non-traders)

	Version 1: 20 %		Version 2: 40 %		Version 3: 60 %	
	Loss	Gain	Loss	Gain	Loss	Gain
Number of respondents	22	29	37	28	46	30
Mean number of years traded	3.5	3.95	3.91	3.81	4.43	4.17
Mean annual income (€)	15,042	16,375	14,834	15,675	21,041	18,630
Number of negative responses (truncated to zero)	11	17	16	14	13	14
Value of a QALY (€)						
Mean	17,439	42,212	43,564	65,957	56,827	48,846
SD	44,561	166,650	13,8097	193,760	126,109	108,570
Median	0	0	0	1020	8673	10,994

for the analysis) for the compensating *gain* questions than for the equivalent *loss* questions. In general, the mean MVQ values increased as the level of income change increased, 60 % income gain being the only exception. The monetary values for a QALY were higher for the gain questions than the loss questions, except in the case of the 60 % income change level. The mean values were consistently higher than the median values, implying that the data were skewed. In half of the cases the median was 0, caused by the large number of respondents who traded enough years to generate a negative MVQ value, which was then truncated to zero.

Table 3 shows the MVQ values calculated using aggregate values. The estimates ranged from €2805 to €49,437. Similar to the individual approach, the mean MVQ values increased as the level of income change increased. Except in the case of the 20 % income change level, the MVQ was higher for the gain questions than for the loss questions.

Table 3 MVQ values calculated at the aggregate level

	Version 1: 20 %		Version 2: 40 %		Version 3: 60 %	
	Loss	Gain	Loss	Gain	Loss	Gain
Number of respondents	78	78	80	80	83	83
Mean number of years traded	0.99	1.47	1.81	1.33	2.45	1.51
Mean annual income (€)	17,471	17,471	15,771	15,771	20,829	20,829
Value of a QALY (€)	17,824	2805	19,082	25,353	30,181	49,437

As shown in Table 4, we tested whether weighted mean monetary values for a QALY for both the disaggregate and the aggregate approach differed between respondents in different income brackets. We found no clear relationship between respondents' income and mean QALY values. For the disaggregate approach, values were broadly similar across income levels, suggesting that the MVQ values generated by our method were not a function of respondent income.

Discussion and conclusions

The aim of this study was not to present a definitive MVQ for the Netherlands, but rather to test the feasibility of an alternative method of eliciting an MVQ. The results from the small-scale online study suggest that the compensating *gain* and equivalent *loss* TTO exercises have potential, but a number of problems must be overcome before its use can be advocated more widely for purposes other than research.

Table 4 Weighted mean QALY values for different income brackets

	Respondent income level (€)	Weighted mean QALY value
Disaggregate approach	Less than 12,000	45,837
	12,000–17,999	39,097
	18,000–23,999	66,060
	>24,000	43,240
Aggregate approach	Less than 12,000	10,401
	12,000–17,999	41,770
	18,000–23,999	30,986
	>24,000	30,137

Generally, respondents in our new method gave up more years when faced with a larger income change level rather than a smaller income change, suggesting some sensitivity to scope. However, these differences were not always significant and never significant without the ‘non-traders’, due to the small numbers in the sample. Surprisingly, we did not find a clear relation between respondent income and MVQ. Maybe this is related to the relatively small sample size of our explorative study. Studies with larger sample sizes may be able to provide more insight into the relationship between income and MVQ values generated with this new approach. Moreover, larger sample size would allow further investigation of sensitivity to scope in the TTO method in this context.

Since respondents are forced to consider giving up years of life from a finite 10-year survival, one could claim that the method introduced here forces respondents to trade-off income and health in a very direct way. Furthermore, the method makes strategic behaviour difficult as it is not obvious to the respondent how the results from the exercise will be used, although the results from this feasibility study do not allow us to specifically test this.

Amongst the sample analysed (excluding 80 ‘extreme non-traders’), 60 % of responses in the equivalent loss and compensating gain questions were non-trades. This is considerably higher than the 35 % found in the study by Dalmau-Matarrodona [21] in the context of a WTP exercise. We have no means of determining what proportion of these trades revealed true lexicographic preferences and what proportion were protest responses. The high proportion of non-trades may also be related to the use of an online survey. Van Nooten et al. [39] found that numerous respondents opted not to trade in conventional TTO exercises in their online questionnaire. It may well be that trading off life time for income is considered in some way ‘unethical’ by respondents or a trade-off they are even less willing to make than trading off length and quality of life. This requires further investigation. The use of discrete choice experiments to elicit WTP could be a fruitful direction for future research in this respect.

A serious problem with the TTO-based approach, and one not encountered when using WTP, is the elicitation of negative MVQ values. It is not easy for respondents to see that they are making choices that imply negative valuation of health, which they may not support if they were shown the implication. This is where the proportion of health traded off exceeds that of the income change. However, in reality, it is plausible that individuals may wish to live for a shorter period of time with higher income than for a longer period of time with lower income, even though their total lifetime income may be lower. For instance, they may feel that the lower income is not enough to be able to sustain themselves and their significant others, so that they would rather live for a shorter time and with a lower total, but higher monthly, income. This also relates to the shape of the utility function assumed here. The additive, linear utility function may not adequately describe people’s actual preferences. In addition, the zero discounting assumption we used here may not hold. If respondents instead discount future income very steeply, a short lifespan with high yearly income will give more discounted utility than a long lifespan with a lower yearly income. It is also likely that respondents may not have been able to calculate exactly at which point their lifetime income in one prospect became lower than that in the other prospect. In that sense, applying this method in an interview elicitation procedure, potentially using visual aids and providing feedback to respondents whose answers imply negative WTP, could support the decision-making process of respondents. This may reduce the number of respondents trading ‘too many years’, yielding negative valuations, but not being aware of this implication.

In this study respondents were told to imagine being in perfect health in both scenarios. In future work it may be preferable to tell respondents they would be in their own current state of health. Their current health could then be valued through either conventional TTO or VAS and the income changes obtained could be divided by the value of the respondents’ current health to give MVQ values. This may reduce the number of hypothetical aspects and hence make the task more manageable for respondents who are

currently not in full health. However, this approach would entail further dependence upon the assumption of no interactions between health and income. This assumption, one of the impossibility theorem criteria set out by Dolan and Edlin [40], is not avoided in this study. The MVQ value elicited is determined essentially by the choice of income change level. A large-scale study would make it possible to obtain values for enough income change levels to estimate an indifference curve between health and income. MVQ values across a range of income change levels could then be estimated. If it is found that the utility of health depends on income and vice versa, this would suggest that an additive utility function is not descriptively valid. In that case, a multiplicative utility function over health and income would be a logical alternative [41].

Another limitation of this study is that we used large income losses, which may be perceived to be unrealistic. Hence, future research may attempt to use more realistic scenarios in order to reduce the hypothetical nature of the data. However, care should be taken that the use of smaller losses does not result in differences becoming too small to be meaningful for the respondents.

Finally, because there is evidence of a lack of the constant proportional trade-off, the willingness to trade years (and thus the trade-off between income and length of life) may depend on the baseline length of life [42, 43]. Moreover, answers to TTO questions may depend on real remaining life expectancy, which in turn depends on income. For this reason, it has been suggested to use real remaining life expectancy in TTO exercises as opposed to an arbitrary number of years of life (10 years in this study), at least for subjects where real life expectancy diverges from preset life expectancy [39, 44]. Future research may investigate this possibility further.

At this moment, the aggregate approach seems to be preferred over the disaggregate approach. Even though it may include some responses of individuals who strategically did not trade, the alternative (the disaggregate approach) left a small number of ‘trading’ respondents after excluding non-traders and truncating negative values to zero. The aggregate approach represents a movement away from standard welfare economics (societal welfare as the sum of individual welfare), but might be considered acceptable in an extra-welfarist framework, although further discussion remains warranted. Further research using face-to-face interviews is needed to try to determine whether the non-trades are strategic or true indicators of preference, and hence whether the calculation method needs to be able to accommodate them.

In summary, the search for the monetary value of QALYs is ongoing, yet remains problematic. Here, we presented an alternative method for the elicitation of MVQ based on the TTO and a first empirical test found it to be

feasible for respondents to answer. Still, the empirical exploration highlighted numerous important issues with the method, most notably the elicitation of ‘non-trades’ and negative values. Future research could address these issues, also looking at the shape of the utility function over income and health. An interview-based study that requires respondents to engage in an iterative process, and that can be supplemented by a visual aid, is required to determine whether this approach is valid and should be taken forward, also as an alternative for WTP valuations.

Acknowledgments We would like to thank Jan van Busschbach, Paul McNamee, Phil Shackley and Richard Edlin who provided helpful feedback and comments on earlier versions of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Weinstein, M.C., Stason, W.B.: Foundations of cost-effectiveness analysis for health and medical practices. *N. Engl. J. Med.* **296**, 716–721 (1977)
- NICE: Guide to the methods of technology appraisal. <http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp> (2008). Accessed 20 February 2014
- Gravelle, H., Brouwer, W., Niessen, L., Postma, M., Rutten, F.: Discounting in economic evaluations: stepping forward towards optimal decision rules. *Health Econ.* **16**(3), 307–317 (2007)
- Claxton, K., Paulden, M., Gravelle, H., Brouwer, W., Culyer, A.J.: Discounting and decision making in the economic evaluation of health-care technologies. *Health Econ.* **20**, 2–15 (2011)
- Culyer, A., McCabe, C., Briggs, A., Claxton, K., Buxton, M., Akehurst, R., Sculpher, M., Brazier, J.: Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *J. Health Serv. Res. Policy* **12**, 56–58 (2007)
- Baker, R., Chilton, S., Donaldson, C., Jones-Lee, M., Lancsar, E., Mason, H., Metcalf, H., Pennington, M., Wildman, J.: Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE. *Health Econ. Policy Law* **6**, 435–447 (2011)
- Donaldson, C., Baker, R., Mason, H., Jones-Lee, M.W., Lancsar, E., Wildman, J., Bateman, I., Loomes, G.C., Robinson, A., Sugden, R., Pinto Prades, J.L., Ryan, M., Shackley, P., Smith, R.: The social value of a QALY: raising the bar or barring the raise?. *BMC Health Serv. Res.* Vol. 11, arte no. 8 (2011)
- Pennington, M., Baker, R., Brouwer, W., Mason, H., Hansen, D.G., Robinson, A., Donaldson, C., the EuroVaQ Team: Comparing wtp values of different types of Qaly gain elicited from the general public. *Health Econ.* **24**, 280–289 (2015)
- Robinson, A., Gyrd-Hansen, D., Bacon, P., Baker, R., Pennington, M., Donaldson, C.: Estimating a WTP-based value of a QALY: the ‘chained’ approach. *Soc. Sci. Med.* **92**, 92–104 (2013)
- Bobinac, A., van Exel, N.J.A., Rutten, F.F.H., Brouwer, W.B.F.: Willingness to pay for a quality-adjusted life-year: the individual perspective. *Value Health* **13**, 1046–1055 (2010)

11. Bobinac, A., van Exel, J., Rutten, F.F.H., Brouwer, W.B.F.: Get more, pay more? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. *J. Health Econ.* **31**, 158–168 (2012)
12. Johnson, F.R., Desvousges, W.H., Ruby, M.C., Stieb, D., De Civita, P.: Eliciting stated health preferences: an application to willingness to pay for longevity. *Med. Decis. Mak.* **18**, S57–S67 (1998)
13. Abelson, P.: The value of life and health for public policy. *Econ. Rec.* **79**, S2–S13 (2003)
14. Johannesson, M., Meltzer, D.: Editorial: some reflections on cost-effectiveness analysis. *Health Econ.* **7**, 1–7 (1998)
15. Gyrd-Hansen, D.: Willingness to pay for a QALY. *Health Econ.* **12**, 1049–1060 (2003)
16. Hirth, R.A., Chernew, M.E., Miller, E., Fendrick, A.M., Weissert, W.G.: Willingness to pay for a quality-adjusted life year. in search of a standard. *Med. Decis. Mak.* **20**, 332–342 (2000)
17. Johannesson, M., Johannsson, P.: Is the valuation of a QALY gained independent of age? Some empirical evidence. *J. Health Econ.* **16**, 589–599 (1997)
18. Helvoort-Postulart, D., Dirksen, C., Kessels, A., Engelshoven, J., Hunink, M.M.: A comparison between willingness to pay and willingness to give up time. *Eur. J. Health Econ.* **10**, 81–91 (2009)
19. Olsen, J.A., Donaldson, C., Pereira, J.: The insensitivity of willingness-to-pay to the size of the good: new evidence for health care. *J. Econ. Psychol.* **25**, 445–460 (2004)
20. Hackl, F., Pruckner, G.J.: Warm glow, free-riding and vehicle neutrality in a health-related contingent valuation study. *Health Econ.* **14**, 293–306 (2005)
21. Dalmau-Matarrodona, E.: Alternative approaches to obtain optimal bid values in contingent valuation studies and to model protest zeros. Estimating the determinants of individuals willingness to pay for home care services in day case surgery. *Health Econ.* **10**, 101–118 (2001)
22. O'Brien, B., Drummond, M.: Statistical versus quantitative significance in the socioeconomic evaluation of medicines. *Pharmacoeconomics* **5**, 389–398 (1994)
23. Kartman, B., Andersson, F., Johannesson, M.: Willingness to pay for reductions in angina pectoris attacks. *Med. Decis. Mak.* **16**, 248–253 (1996)
24. Kartman, B., Stålhammar, N., Johannesson, M.: Valuation of health changes with the contingent valuation method: a test of scope and question order effects. *Health Econ.* **5**, 531–541 (1996)
25. O'Connor, R.M., Johannesson, M., Hass, S.L., Kobelt-Nguyen, G.: Urge incontinence. *Pharmacoeconomics* **14**, 531–539 (1998)
26. Chestnut, L.G., Keller, L.R., Lambert, W.E., Rowe, R.D.: Measuring heart patients willingness to pay for changes in angina symptoms. *Med. Decis. Mak.* **16**, 65–76 (1996)
27. Jones-Lee, M.W., Loomes, G.C., Philips, P.R.: Valuing the prevention of non-fatal road injuries: contingent valuation vs. standard gambles. *Oxf. Econ. Pap.* **47**, 676–695 (1995)
28. Smith, V.K., Desvousges, W.H.: An empirical analysis of the economic value of risk changes. *J. Polit. Econ.* **95**, 89–114 (1987)
29. Fonta, W., Ichoku, H.E., Kabubo-Mariara, J.: The effect of protest zeros on estimates of willingness to pay in healthcare contingent valuation analysis. *Appl. Health Econ. Health Policy* **8**, 225–237 (2010)
30. Brouwer, W.B.F., Culyer, A.J., van Exel, N.J.A., Rutten, F.F.H.: Welfarism vs. extra-welfarism. *J. Health Econ.* **27**, 325–338 (2008)
31. Donaldson, C.: Valuing the benefits of publicly-provided health care: does 'ability to pay' preclude the use of 'willingness to pay'? *Soc. Sci. Med.* **49**, 551–563 (1999)
32. Buckingham, K., Devlin, N.: A theoretical framework for TTO valuations of health. *Health Econ.* **15**, 1149–1154 (2006)
33. Tilling, C., Krol, M., Tsuchiya, A., Brazier, J., van Exel, J., Brouwer, W.: Does the EQ-5D reflect lost earnings? *Pharmacoeconomics* **30**, 47–61 (2012)
34. Tilling, C., Devlin, N., Tsuchiya, A., Buckingham, K.: Protocols for time tradeoff valuations of health states worse than dead: a literature review. *Med. Decis. Mak.* **30**, 610–619 (2010)
35. Hicks, J.R.: Value and capital: an enquiry into some fundamental principles of economic theory. Clarendon, Oxford (1939)
36. Eeckhoudt, L., Godfroid, P., Marchand, M.: Risque de santé, médecine préventive et médecine curative. *Rev. d'Econ. Politique* **108**, 321–337 (1998)
37. Layard, R., Mayraz, G., Nickell, S.: The marginal utility of income. *J. Public Econ.* **92**, 1846–1857 (2008)
38. Gyrd-Hansen, D., Kjær, T.: Disentangling WTP per QALY data: different analytical approaches, different answers. *Health Econ.* **21**, 222–237 (2012)
39. van Nooten, F.E., Koolman, X., Brouwer, W.B.F.: The influence of subjective life expectancy on health state valuations using a 10 year TTO. *Health Econ.* **18**, 549–558 (2009)
40. Dolan, P., Edlin, R.: Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? *J. Health Econ.* **21**, 827–843 (2002)
41. Levy, M., Nir, A.R.: The utility of health and wealth. *J. Health Econ.* **31**, 379–392 (2012)
42. Attema, A.E., Brouwer, W.B.F.: On the (not so) constant proportional tradeoff in TTO. *Qual. Life Res.* **19**, 489–497 (2010)
43. Bleichrodt, H., Pinto, J.L., Abellán-Perpinán, J.M.: A consistency test of the time trade-off. *J. Health Econ.* **22**, 1037–1052 (2003)
44. Heintz, E., Krol, M., Levin, L.A.: The impact of patients subjective life expectancy on time trade-off valuations. *Med. Decis. Mak.* **33**, 261–270 (2013)