The corrective approach for health utility measurement using time trade-off - a direct test of validity

EsCHER Working Paper No. 2023012
November 2023

Jacob A. Richelshagen, M.Sc.
Katja C. Senn, Ph.D.
Klaus Nagels, Ph.D.
Stefan A. Lipman, Ph.D.

Erasmus University Rotterdam
Making Minds Matter

Title
The corrective approach for health utility measurement using time trade-off - a direct test of validity

Authors
Jacob A. Richelshagen, M.Sc. University of Bayreuth, Chair of Healthcare Management and Health Services Research, Bayreuth, Germany
Katja C. Senn, Ph.D. University of Bayreuth, Chair of Healthcare Management and Health Services Research, Bayreuth, Germany
Klaus Nagels, Ph.D. University of Bayreuth, Chair of Healthcare Management and Health Services Research, Bayreuth, Germany
Stefan A. Lipman, Ph.D. Erasmus School of Health Policy & Management, Rotterdam, Netherlands

Corresponding author: Jacob A. Richelshagen, email: j.richelshagen@web.de

Keywords
QALYs; corrective approach; loss aversion; prospect theory; time trade-off

JEL classification
I12; D91

Cite as
Richelshagen, J.A.; Senn, K.C.; Nagels, K.; Lipman, S.A. (2023). The corrective approach for health utility measurement using time trade-off - a direct test of validity. EsCHER Working Paper Series No. 2023012, Erasmus University Rotterdam. Available from: https://www.eur.nl/en/research-groups-initiatives/escher/research/working-papers

The corrective approach for health utility measurement using time trade-off - a direct test of validity

Richelshagen[a,*], J.A.; Senn[a], K.C.; Nagels[a], K.; Lipman[b], S.A.

[a] University of Bayreuth, Chair of Healthcare Management and Health Services Research, Bayreuth, Germany
[b] Erasmus School of Health Policy & Management, Rotterdam, Netherlands
* Corresponding author: j.richelshagen@web.de

Funding: No funding was received for conducting this study

Competing interest: The authors have no relevant financial or non-financial interests to disclose.

Title: The corrective approach for health utility measurement using time trade-off - a direct test of validity

Abstract:

Objective: Quality adjusted life years (QALYs) are used to measure health gains of an intervention. A classical method to elicit utility-weights for QALYs is time trade-off (TTO). TTO is often applied assuming that linear utility of life duration holds. It has been shown that this assumption is often not valid, leading to biased elicitation outcomes. The corrective approach introduced by Lipman et al. (2019) corrects QALY-weights based on prospect theory (PT). Published studies have yielded results in line with theoretical predictions, but arguably low external validity. The present study investigates if corrected TTO-weights reflect preferences of individuals more accurately than uncorrected TTO-weights.

Method: The following method was applied: First, QALY-weights were elicited using TTO. Second, participants loss aversion was quantified based on PT, and QALY-weights were corrected accordingly. Third, corrected QALY-weights were validated by presenting corrected and uncorrected QALY-weights and asking participants to choose which more accurately reflects their preferences and enabling adjustments using a visual analogue scale (VAS).

Results: QALY-weights for three EQ-5D-5L health states $ß_{1-3}$ were elicited from 66 German university students, using online interviews. Corrected QALY-weights were lower than uncorrected QALY-weights. The direction of the corrections was in line with the direction of participant´s adjustments using VAS. Most participants indicated that uncorrected QALY-weights reflect their preferences more accurately than corrected QALY-weights (Binomial test, $p<0.05$ for $ß_1$, $p>0.05$ for $ß_{2-3}$).

Discussion: These results suggest that a need for correction exists but that the corrective approach does not yield QALY-weights with better validity. Future work should replicate this study with a representative sample and individual-level measures of participants´ utility curvature.

1. Introduction

In the context of increasing resource pressure in healthcare, economic evaluations provide transparency enabling informed allocation decisions. One of the most frequently used methods for health economic evaluation is cost-utility analysis (CUA) [1], in which the costs of an intervention are contrasted to the utility associated with health gains realized through the intervention. Prerequisite for CUA is precise measurement of utility. Typically, quality adjusted life years (QALYs) are used to measure utility of health gains of an intervention [2, 3, 4], combining quality and quantity dimensions of life in certain health states. More specifically, QALYs are calculated by multiplying the number of life years a health gain is experienced in, with a utility weight (i.e., QALY-weight), representing the health status [5].
Although many methods exist to elicit QALY-weights, a method that has been used extensively is Time trade-off (TTO). In TTO, individuals are presented with a choice between living a perfect health for a given number of years, and living in a given health state for a typically higher given number of years. The number of years lived in perfect health is varied, until indifference between the two health profiles is reached, which enables quantifying the utility of the given health state. In practice, TTO is often used for e.g., EQ-5D valuation studies. Another frequently used method is standard gamble (SG), in which participants are presented with a choice between living in a given health state for a number of years and a gamble with the outcomes of living in perfect health for a number of years, or dying immediately. The probability of occurrence of the two alternatives is varied, until indifference between the two alternatives is reached, which enables quantifying the utility of the given health state [1]. Lastly visual analogue scales (VAS) can be used, in which participants are asked to indicate the relative value of a given health state on a visual scale between death and full health [1]. The present study focuses on TTO, as research indicates it is better suited than SG to measure preferences in the context of health [6-8] and is preferred by surveyed individuals vis-à-vis SG [9, 10].
TTO is based on the linear QALY-model, which assumes that expected utility (EU) holds. This means its application implies a certain set of assumptions about the decision-making behaviors of individuals, e.g., as risk neutrality with regards to life years [11]. Research has shown that these assumptions are descriptively not valid, e.g., that most individuals are risk averse with regards to life years [12], which leads to potentially biased elicitation outcomes [11, 13]. The *corrective approach*, as introduced by Lipman et al. aims to correct elicited QALY-weights by applying calculations based on prospect theory (PT) [14]. PT was developed by Kahneman and Tversky (1979) based on empirical observations of EU violations, and is characterized by four principles [16, 16]:

- (I) Reference-dependence, i.e., utilities are dependent on a reference point
- (II) Loss aversion, i.e., losses loom larger than equally sized gains
- (III) Non-linear utility, i.e., utility functions are concave/convex for gains/losses respectively
- (IV) Probability weighting, i.e., high/low probabilities are under-/overweighted respectively

As (II) loss aversion and (III) non-linear utility are expected to impact TTO elicitations, the present study will focus on corrections based on these two principles [17].
On a high level, the approach comprises two steps: First, depending on the elicitation method, biases applicable to the decision making of surveyed individuals are quantified. Second, QALY-weights elicited from these individuals are corrected for these biases [18].
There are empirical findings supporting the corrective approach. For example, Lipman et al. observed convergence between corrected SG and TTO values when applying the corrective approach [14]. In a subsequent study, they found that if respondents were asked to reflect on and adjust their composite TTO values, their adjustments were in accordance with the direction of corrections indicated by the corrective approach [18]. However, the application of the corrective approach in practice is currently limited. The key reason for this is the difficulty to establish if corrected QALY-weights do adequately reflect the preferences of surveyed individuals [17].
Hence, the present study aims to investigate the hypothesis that corrected TTO-weights reflect the preferences of individuals more accurately than uncorrected TTO-weights. Therefore, the following three-step approach was applied: a) Elicitation of QALY-weights from individuals for a set of EQ-5D-5L health states using TTO; b) quantification of loss aversion, and correction of QALY-weights elicited using the approach outlined by Lipman et al. [14]; c) validation of QALY-weights by explaining participants the concept of QALYs, presenting both corrected and uncorrected QALY-weights and asking participants directly to choose which QALY-weight more accurately reflects their preferences for each health state. Additionally, participants were asked to directly assign and confirm QALY-weights to each health state using a visual analogue scale (VAS), applying an approach like Lipman et al. [14, 18].
In the next chapter, the theoretical background of the corrective approach will be described, showing how the general QALY-model can be adapted to model loss aversion and non-linear utility curvature in TTO. In the following chapter, we will show the methods applied in the present study to elicit, correct, and validate QALY-weights, before presenting and discussing our key findings.

2. Theoretical Background

At first, the notation applied, and assumptions are outlined. We model preference for a decision-maker choosing between different health profiles. Health profiles are denoted as $(ß, t)$. ß represents a health state and t the age, in which this health state ends. The age in which a health states begins is denoted as t_a. Indices (x, y) will be used to denote alternative results of a decision. Hereafter, these alternatives will be abbreviated as $(ß_x, T_x)$ with $T_x = t_x - t_a$. Preference, weak preference, and indifference are denoted as $\succ, \succcurlyeq$ and $\sim$ respectively, immediate death and full health are denoted as D and FH respectively. Full health is defined as a health state, in which an individual does not experience problems in any of the EQ-5D-5L dimensions. We furthermore make a set of simplifying assumptions. First, all health states ß are chronic and preferred to D. Second, we assume the zero condition [19] holds, i.e., individuals are indifferent between all health profiles with $T = 0$,., i.e., profiles equivalent to D. Furthermore, it will be assumed that if $x \succ y$ and $y \succ z$, then $x \succ z$ (i.e., transitivity) and $(ß_x, T_x) \succ (ß_x, T_y)$ if $T_x \succ T_y$ (i.e., monotonicity) for all ß. Notation and assumptions made correspond to Lipman et al. [14]. Under these assumptions, preferences for health profiles (preferred to D) can be represented by the general QALY-model as developed by Miyamoto et al. in which the utility of health profiles is evaluated as in equation 1 [19].

$$\boldsymbol{V}(ß_x, \boldsymbol{T}_x) = \boldsymbol{U}(ß_x)\boldsymbol{L}(\boldsymbol{T}_x) \tag{1}$$

In this model, $U(ß_x)$ describes an individual's utility function over a health state, and $L(T_x)$ describes an individual's utility function over life years.
As outlined, TTO is often used to elicit utility values in the context of health to be used in the QALY-model. In TTO, individuals are asked how many life years they are willing to give up to prevent a certain health state. Specifically, individuals are presented with a choice between living for T_x years in $ß_x$, and living T_y years in FH, with $T_y < T_x$. T_y is then varied until indifference in the form $(ß_x, T_x) \sim (FH, T_y)$ is reached. Typically, TTO is applied assuming linear utility, i.e., $L(T_x) = T_x$ which allows the derivation of uncorrected QALY-weights as $U(ß_x) = \frac{T_y}{T_x}$, assuming that $U(FH) = 1$. In this application, no additional information is required from respondents beyond their indifference between the two health profiles evaluated. Applying however the non-linear QALY-model to abovementioned indifference and assuming that $U(FH) = 1$, this yields equation 2.

$$\boldsymbol{U}(ß_x) = \frac{\boldsymbol{L}(\boldsymbol{T}_y)}{\boldsymbol{L}(\boldsymbol{T}_x)} \tag{2}$$

As outlined previously, PT provides an alternative to EU in modeling individual decision-making behavior, taking biases into consideration. Based on the insights of PT, Bleichrodt conjectured that first, loss aversion causes TTO-values to be biased upwards and second, non-linear utility causes TTO-values to be biased downwards (for a detailed explanation see Bleichrodt, 2002) [11]. Hence, Lipman et al. expanded the equation 3 by including additional parameters (PT-parameters) to model - and thus correct - for these biases [14].

$$\boldsymbol{U}(ß_x) = \frac{\boldsymbol{L}^{-}(\boldsymbol{T}_y^*) + \mathbf{1}}{(\mathbf{1} - \boldsymbol{\lambda})\boldsymbol{L}^{-}(\boldsymbol{T}_y^*) + \mathbf{1}} \tag{3}$$

L is assumed to be sign-dependent, i.e., taking a different shape for losses and for gains in life duration. In equation 3, L^{-} denotes an individual's utility function for life years in the domain of losses. L^{-} applies in TTO exercises as described earlier, since Lipman et al. assume that individuals use $(ß_x, T_x)$ as reference point $(ß_r, T_r)$, with $T_y^* = T_x - T_r$, characterizing T_y^* as a loss. In other words, TTO is assumed to involve trading off losses in life duration for improvements in quality of life. Loss aversion is denoted by λ, as defined by Köbberling and Wakker [14, 20]. For a detailed derivation of equation 3 see Lipman et al. [14].

3. Methods

Our empirical cross-sectional study was conducted between 01.06.2021-30.06.2021 using a computer-based questionnaire programmed with R Shiny. 66 students were recruited as participants, to facilitate comparability to the results of Lipman et al. [14]. Participants were recruited using social media (WhatsApp, Facebook) as well as the eLearning platform of the University of Bayreuth. As inclusion criteria we set a current enrollment at a German university. Nine Amazon vouchers (Total worth of 200€) were raffled among participants as incentives. The experiment was conducted in one-on-one interviews using the video communication software Zoom. This approach was taken to comply with COVID-19 measures. Recent research indicates that the data quality of TTO elicitations in virtual settings are comparable to TTO elicitations in real-world settings [21, 22]. The role of the interviewer in the experiments was limited to answering clarifying questions while participants were completing the questionnaire autonomously to ensure objectivity. In the sessions, purpose, duration, and procedure were explained to the participants, and they were reminded that stopping the interview at any given time is possible. To finally start the questionnaire, participants had to give their informed consent via the R Shiny questionnaire. The questionnaire was divided into three parts, analogous to the structure previously outlined: a) elicitation of QALY-weights, b) elicitation of PT-parameters, and c) validation of elicited QALY-weights. Lastly, age and gender of participants were collected. This approach is like that of Lipman et al., who applied a comparable approach to test if QALY-weights elicited with SG or TTO have higher validity [8].

4.1 Elicitation of QALY-weights with TTO

QALY-weights were elicited for three hypothetical EQ-5D-5L health states which are described in table 1 below along five dimensions (mobility, self-care, usual activities, pain / discomfort, anxiety / depression) with 5 levels respectively (Not / no problems, slight problems, moderate problems, severe problems, unable to / extreme / extremely).

	21211	**31231**	**33342**
I have … problems walking about	slight	moderate	moderate
I have … problems washing or dressing myself	no	no	slight
I have … problems doing my usual activities	slight	slight	moderate
I have … pain or discomfort	no	moderate	severe
I am … anxious or depressed	not	not	slightly

Table 1 EQ-5D-5L health states selected

The health states 21211, 31231, and 33342 (Hereafter denoted as β_1, β_2, and β_3 for brevity) were selected for two reasons. First, they are relatively moderate which helps to avoid participants preferring immediate death, which would violate our applied assumptions. Second, to facilitate comparability to the results of Lipman et al. [14], who have used the same set of health states. It should be noted that the health states in the order as depicted in table 1 have either equal or worse levels across all dimensions. Hence, it should be expected participants prefer for example β_1 over β_3. Participants were instructed to imagine having lived until the age of 50 in full health, when they would contract a disease which would reduce the quality of life in the next 20 years, after which they would die an immediate painless death. These instructions aim to ensure comparability of QALY-weights elicited from different participants as preferences are always indicated for the same timespan (20 years). The participants were asked 3 questions per health state in which they were able to choose between living in full health for a reduced amount of time T_y or 20 years in the respective health state ß$_x$. T_y^* was varied after each question depending on the participants answers in decreasing steps, to “zoom in” until participants were indifferent between the two options presented, using a bisection approach. After the third iteration participants could use a slider to adjust T_y so that they were indifferent between the two options. If indifference was already reached earlier, participants had the option to indicate this during the initial questions. In addition, bar charts were used as a visual aid (see appendix I).

4.2 Elicitation of PT-parameters

To correct TTO-weights based on the approach introduced by Lipman et al. [14], two parameters need to be elicited to correct for, *utility curvature* and *loss aversion* [15]. To elicit these parameters, a modified version of the approach taken in the aforementioned paper was applied in the present study. To reduce interview duration, a parametric assumption was made with regards to utility curvature of participants in the form of $(T^*) = -\left(-\left(\frac{T}{T_r}\right)^{\alpha}\right)$ with $\alpha = 0.88$, i.e., using the original utility curvature estimates by Tversky and Kahneman, as a fully non-parametric elicitation would have required a number of additional questions [14]. To include this assumption into the model introduced in chapter 2, equation 3 is modified to equation 4. As it has been shown that effect of correcting for loss aversion is larger than the effect of correcting for utility curvature [14], limited impact of this assumption on the results of the study is expected.

$$U(ß_x) = \frac{\left(-\left(-\left(\frac{T_y}{T_r}-1\right)^{\alpha}\right)\right)+1}{(1-\lambda)\left(-\left(-\left(\frac{T_y}{T_r}-1\right)^{\alpha}\right)\right)+1} \tag{4}$$

In equation 4, loss aversion is applied with λ. In accordance with the approach taken by Lipman et al. based on a non-parametric method to measure PT developed by Abdellaoui et al., λ is defined as the kink of the utility function at the reference point [14, 20, 23]. λ is calculated by dividing the first point on the utility function above a given reference point x_1^+ by the first point below the same reference point x_1^-. Hence, to elicit λ, participants answered three questions in which they were asked to choose between a gamble in the form of $(\beta_x, T_x^*)_p(\beta_x, T_y^*)$, in which the first / second alternative´s probability of occurrence is p / $1-p$ respectively, and a certain outcome in the form of (β_x, T_z^*) with $T_y^* < 0 < T_z^* < T_x^*$. Participants were asked to imagine living in full health until the age of 70, when they would contract a deadly disease which requires treatment with one of two drugs. The effects of the respective drugs are characterized by a certainty equivalent (CE) outlined above. Participants were given a visual aid in form of a probability wheel (see appendix II) [24]. Based on the participants' replies T_x^* was varied to "zoom-in" on indifference, again using a bisection approach.
After the third iteration per question participants could use a slider to adjust T_x^* so that they were indifferent between the two options, eliciting indifferences in the form of $(\beta_x, T_x^*)_p(\beta_x, T_y^*) \sim (\beta_x, T_z^*)$. For a detailed description on the calculation of λ see Lipman et al. [13] Probability weighting, i.e., the tendency to underweigh high probabilities and overweigh low probabilities, was not corrected as we only expected a minor effect at $p = 0.5$ [15]. Due to the applied assumption with regards to utility curvature and having obtained λ, QALY-weights elicited could be corrected, which were validated by participants in the next part of the experiment.

4.3 Validation of QALY-weights

To examine the validity of corrected QALY-weights, a *direct test* was conducted. Participants were asked directly if corrected or uncorrected QALY-weights reflect their preferences with regards to the respective health state better. Before this exercise, participants were explained the QALY-concept in text format, using an explanation validated by Bleichrodt et al. [25]. Analogously to Lipman et al., a two-step approach was applied [8]. First, participants were presented the corrected and uncorrected QALY-weights for a health state. Participants were asked to choose the value which more accurately represented their conceptions of the value of life in this health state. In a second step, participants could use a slider to choose a value between 0 and 1 (whereas 0 represented D and 1 FH), which represented their conceptions of the value of life in the respective health state, to confirm the utilities elicited using a VAS. Again, bar charts were used as visual aid (see appendix III).

4. Results

SPSS version 28 was used for data analysis. In total, 66 participants were included. 68% were female (n=45) and 32% male (n=21); mean age was 25.3 (18-37) and median loss aversion was 3.8 (86% loss averse; 9% gain seeking; 5% neutral). The QALY-weights elicited are depicted in table 2 including QALY-weights from the German EQ-5D-5L value set [26].

	β_1		β_2		β_3	
	U	SD	U	SD	U	SD
EQ-5D-5L-Index	0.938	-	0.831	-	0.419	-
Elicited	0.809	0.164	0.733	0.173	0.459	0.256

Table 2 QALY-weights elicited in comparison to German EQ-5D-5L value set

The QALY-weights elicited with TTO were not distributed normally (Shapiro-Wilk-test, $p < 0.05$ for all β), hence non-parametric tests were applied. The results of the correction of QALY-weights are depicted table 3, including a breakdown by effect to better represent the relative effect size.

	β_1		β_2		β_3	
	U	SD	U	SD	U	SD
Loss aversion and utility curvature	0.587	0.247	0.481	0.242	0.249	0.212
Loss aversion $(\alpha = 1)$	0.634	0.244	0.528	0.279	0.279	0.229
Utility curvature $(\lambda = 1)$	0.777	0.171	0.695	0.177	0.427	0.250
VAS	0.797	0.136	0.631	0.135	0.380	0.188

Table 3 Corrected QALY-weights including breakdown by effect

The first test of the hypothesis outlined was to check whether a statistically significant larger share of participants selected corrected QALY-weights as reflecting their preferences more accurately than uncorrected QALY-weights. As the following figure 1 shows, a majority of participants indicated that the uncorrected QALY-weights reflect their preferences more accurately than the corrected QALY-weights (Binomial test, $p < 0.05$ for β_1, $p > 0.05$ for β_{2-3}).

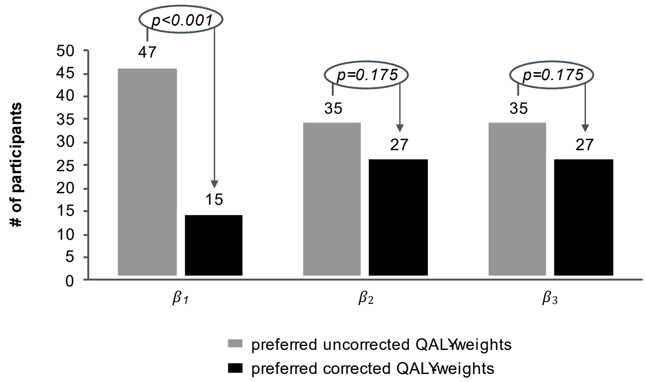

Fig. 1 Results of validation exercise

The difference however is only statistically significant for the mildest health state β_1, while with increasing severity of the health states the difference is not statistically significant. This result suggests that the corrective approach has lower validity for mild health states, for severe states it predicts preferences just as well as uncorrected QALY-weights.
The second test of the hypothesis outlined was whether the difference between the utilities directly assigned by the participants after an explanation of the QALY-concept was smaller for the corrected QALY-weights than it was for the corrected QALY-weights. However, as figure 2 shows, the differences were actually smaller for corrected QALY-weights. The difference is statistically significant for β_1 and β_3 (Wilcoxon-Sign-Rank-Test, $p < 0.05$) and not for β_2 (Wilcoxon-Sign-Rank-Test, $p > 0.05$).

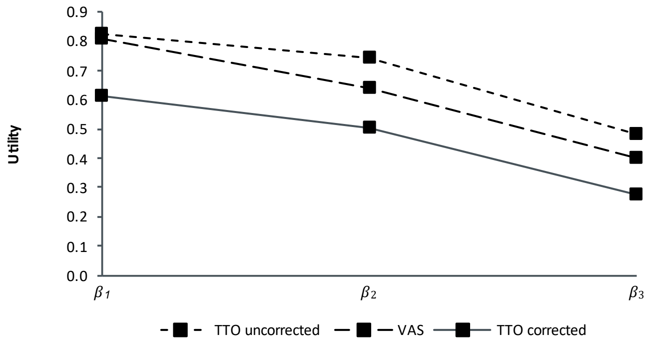

Fig. 2 Average utility per health state

When only correcting for utility curvature, the difference to the validated VAS-weights was statistically significantly smaller for corrected QALY-weights for the health states β_{2-3} (Wilcoxon-Sign-Rank-Test, $p > 0.05$). This mirrors the interpretation of the results of the first test, according to which the difference in validity between uncorrected and corrected QALY-weights seems to diminish with increasing severity of health states.
It should be noted that two participants did not want to give up any life years during the experiment (e.g., for religious reasons) and eleven participants violated assumptions taken (e.g., monotonicity across health states). The exclusion of these participants from the analysis did not alter the results materially (see appendix IV).

5. Discussion

Our starting hypothesis was that the application of the corrective approach, based on the model developed by Lipman et al. leads to more accurate elicitation of QALY-weights, equivalent to a higher validity according to the definitions taken as a basis [14]. We did not find evidence supporting this hypothesis.
Generally, corrected QALY-weights were *lower* than uncorrected QALY-weights, which corresponds to the results of Lipman et al. [14, 18]. Furthermore, the direction of the corrections made are in accordance with the direction of the adjustments made by participants using the VAS. Bleichrodt expected TTO-weights to be biased downwards, however as outcomes in the present study were characterized as losses, a correction downwards is in line with PT as it is expected that the utility function of participants is convex [11]. Furthermore, it is evident that the correction for loss aversion has a relatively stronger effect on the corrected QALY-weights than the effect of correction for utility curvature. This is in line with the results of Lipman et al. [14], however in this study necessarily driven by the assumed α.
However, participants indicated for a set of different health states, that uncorrected TTO-weights reflected their preferences more accurately than the corrected values. This indicates too low corrected values and that the corrective approach does not necessarily elicit more valid TTO-weights, which has been indicated in previous validation studies as well [17]. In the wider context of health state valuation, this means that further research is required to enable adequate debiasing of the elicitation of QALY-weights with TTO, but also other classical elicitation methods. Much of the corrective approach hinges on the assumptions about the reference-point in TTO, i.e., the assumption that respondents give up life duration for improved health status. A crucial direction for subsequent research is identifying if this reference-point (and its’ corresponding strong downward direction correction on QALY weights) applies.

Some evidence exists already suggesting that respondents use different reference-points in TTO, which may be a potential explanation for the limited validity of the approach used in the present study [27, 28]. Nevertheless, various limitations must be noted. First, the external validity of the results is limited due to sample characteristics as participants of this study were young, predominantly female and university educated. In addition, for a number of participants problems of comprehension were indicated, for example by inconsistent answers given or multiple comprehension questions asked by single participants. Considering these limitations, it is suggested to rerun this study with corresponding amendments, including a larger and representative sample to increase external validity applying a fully non-parametric corrective approach, randomizing health states. Should the results of this study be confirmed, modifications to the theoretical model need to be made as QALY-weights appear to be too low and compressed, as previously indicated in another publication [17]. Second, there is potential bias due to anchoring effects as the sequence of health states was not randomized in the health state valuation of the experiment. Anchoring describes a heuristic in which decision makers adjust a familiar starting value, which can be completely unrelated to the decision at hand [29]. Hence, it can be considered that participants used the QALY-weight assigned to one health state as an anchor when assigning the QALY-weight to the following health state, potentially leading to a downwards bias. This effect could potentially also have affected the validation step of the experiment, as the sequence of health states was not randomized. Third, there is potential bias due to the simplifying parametric assumption made about the utility curvature of participants. As this assumption implies a certain shape of participants´ utility function, it could cause corrected QALY-weights to inaccurate for those participants for which this assumption does not hold true. Fourth, to enable broad application of the corrective approach, future investigations of the corrective approach should apply the variation introduced by Lipman et al. which allows the correction of QALY-weights elicited with lead-time-TTO (LT-TTO), as also health states being considered worse than death can be evaluated [18]. It should be noted that for all health states, QALY-weights elicited were lower than corresponding values in the German value set. This is however to be expected due to the characteristics of the sample and different methods in the present study. First, it is generally understood that younger participants tend to assign lower QALY-weights to health states. Second, in EQ-5D valuations typically 10-year durations are used, beginning at the point in time of the survey. However, in the present study a 20-year duration starting at age 50 was applied.

6. Conclusion

To conclude, the results of the present study provide no support for the validity of the corrective approach, cautioning against its' application in economic evaluation practice. Nevertheless, the corrective approach should be investigated further as from a theoretical perspective it still appears like a promising steppingstone on the path to more eliciting more valid QALY-weights.

Erasmus University Rotterdam
Erasmus Centre for Health Economics Rotterdam
Burgemeester Oudlaan 50
3062 PA Rotterdam, The Netherlands
T +31 10 408 8555
E escher@eur.nl
W www.eur.nl/escher