On the (not so) constant proportional trade-off in TTO

Arthur E. Attema · Werner B. F. Brouwer

Abstract

Purpose The linear and power QALY models require that people in Time Trade-off (TTO) exercises sacrifice the same proportion of lifetime to obtain a health improvement, irrespective of the absolute amount. However, evidence on these constant proportional trade-offs (CPTOs) is mixed, indicating that these versions of the QALY model do not represent preferences. Still, it may be the case that a more general version of the QALY model represents preferences. This version has the property that people want to sacrifice the same proportion of *utilities* of lifetime for a health improvement, irrespective of the amount of this lifetime.

Methods We use a new method to correct TTO scores for utility of life duration and test whether decision makers trade off utility of duration and quality at the same rate irrespective of duration.

Results We find a robust violation of CPTO for both uncorrected and corrected TTO scores. Remarkably, we find higher values for longer durations, contrary to most previous studies. This represents the only study correcting for utility of life duration to find such a violation.

Conclusions It seems that the trade-off of life years is indeed not so constantly proportional and, therefore, that health state valuations depend on durations.

Keywords Constant proportional trade-offs ·
QALY model · TTO method · Utility of life duration

A. E. Attema (✉) · W. B. F. Brouwer
iBMG/iMTA, Erasmus University Rotterdam,
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
e-mail: attema@bmg.eur.nl

W. B. F. Brouwer
e-mail: brouwer@bmg.eur.nl

Abbreviations

QALY	Quality-adjusted life year
TTO	Time trade-off
CPTO	Constant proportional trade-offs
MET	Maximum endurable time
CE	Certainty equivalence
FH	Full health
D	Death
BP	Back pain

Introduction

The quality-adjusted life year (QALY) model has become an important model in valuing health benefits. To make the model practical, measurement methods are needed in order to elicit the quality of life weights used in this model. One such method is the time trade-off (TTO) method, which is often used to derive (standard) quality of life weights for health states to be used in economic evaluations [1, 2]. The popularity of the TTO, however, cannot be explained by the absence of methodological problems. On the contrary, the TTO has been shown to be prone to several influences such as loss aversion, scale compatibility and utility of duration (i.e., discounting) [3].

One important and necessary assumption of the QALY model is that of constant proportional trade-offs (CPTO). In the context of TTO, CPTO basically requires that the estimated TTO value should be the same for different durations. For example, if in valuing some imperfect health state β using a 10-year TTO, people would indicate they are willing to trade off 2 years (that is 20% of total time), then CPTO requires them to give up 2 months when using a 10-month TTO or 2 days when using a 10-day TTO. The proportion traded should always be equal (i.e. 20%). CPTO

is also practically important when one considers the use of the valuation of health states in economic evaluations and medical decision making: they are attached to such health states regardless of the duration of the health problem, normally. If, therefore, the assumption of CPTO does not hold, health state valuations could be time dependent—that is, health states could be valued differently when their durations differ.

The evidence on the validity of the CPTO assumption is mixed. Some empirical studies found support [4–6], while others rejected it [7, 8], or found mixed results [9]. Given the importance of the assumption and the mixed evidence for it, more research in this area seems warranted. In this paper, we therefore discuss the current evidence regarding CPTO on the basis of a literature review and highlight the role of the utility of life duration in this debate. So far, most studies that found violations of CPTO assumed linear utility of life duration (i.e. no discounting of future life years), but it seems implausible that their subjects would satisfy that assumption. Therefore, if one was to correct for utility of life duration curvature, these subjects might satisfy CPTO in terms of utilities for life duration after all. That is, TTO values corrected for utility of life duration curvature may still be the same for different durations, despite the fact that uncorrected TTO values vary with duration. It is important to investigate this possibility, because it might indicate that the QALY model does hold in a more general form and that only the assumption regarding the shape of the utility of life duration function has to be relaxed.

We present the results of an experiment to test the CPTO assumption, in which we used both uncorrected TTO values and TTO values that were corrected for utility of life duration curvature. For this correction, we used the risk-free utility of life duration elicitation method proposed by Attema et al. [10]. Its advantages are that it does not need to make specific parametric assumptions about the utility for life duration function and that it is not influenced by biases due to probability weighting and the inclusion of the problematic outcome death.

This paper is organized as follows: we describe the theoretical background of CPTO in the next section and review the existing literature that tested the CPTO assumption in the literature review. Our experimental test is described in the experimental section, followed by a presentation of the results. The last section discusses the results and provides a possible explanation for our findings in the form of a generalized relationship between duration and trade-offs.

Theory

As indicated, the TTO method is based on the QALY model. The QALY model is a common way to describe preferences over health profiles. Let $h = (h_j,…,h_T)$ denote a health profile, where h denotes the health state in period $t = j,…,T$, where T is the decision maker's final period of life. Further, $v(h_t)$ is a value function that represents the individual's preferences over health quality and $\delta(t)$ denotes the corresponding weight attached to the value in this period. It can then be shown that, under some reasonable assumptions, $h \geqslant h'$ if and only if $\sum_{t=j-1}^{T} \delta(t) * v(h_t) \geq \sum_{t=j-1}^{T} \delta(t) * v(h'_t)$ (we use $t = j - 1$ instead of $t = j$ because the first period ($j = 1$) starts at $t = 0$) [11]. We call $U(t, h_t) = \sum_{t=j-1}^{T} \delta(t) * v(h_t)$ the *generalized QALY model* and assume that health profiles are evaluated by this function.

In order to avoid confusion, we term $v(h_t)$ a (TTO) *value* throughout the paper, whereas the function $\sum_{t=j-1}^{T} \delta(t)$ is termed the *utility of life duration* for the period between $t = j - 1$ and $t = T$. Notice that this terminology differs from some of the literature where preferences for health states elicited in a riskless setting are termed 'values' and preferences for health states elicited in a risky setting are termed 'utilities'.

CPTO holds if the proportion of remaining life years that one is willing to give up for an improvement in health status from any health state β_1 to any health state β_2 does not depend on the absolute number of remaining life years [12] (Pliskin et al. [12] considered a probabilistic framework, but the definition is similar for deterministic frameworks [13, 14]). This means that for any health state β_1 and any health state β_2, the individuals are prepared to give up the proportion $(n_{\beta 1} - n_{\beta 2})/n_{\beta 1} = 1 - n_{\beta 2}/n_{\beta 1}$. Thus, there exists a number $q \geq 0$, such that $q = n_{\beta 2}/n_{\beta 1}$ and individuals are willing to give up the same proportion $(1 - q)$ of lifetime irrespective of its duration $(n_{\beta 1})$.

Generalized CPTO holds if the proportion of remaining *utility* of life years that one is willing to give up for an improvement in health status from any health state β_1 to any health state β_2 does not depend on the *utility* of the absolute number of remaining life years. This would mean that individuals are prepared to give up the proportion $\left(\sum_{t=0}^{n_{\beta 1}} \delta(t) - \sum_{t=0}^{n_{\beta 2}} \delta(t)\right)/\sum_{t=0}^{n_{\beta 1}} \delta(t) = 1 - \sum_{t=0}^{n_{\beta 2}} \delta(t)/\sum_{t=0}^{n_{\beta 1}} \delta(t)$. Then, there may exist a number $q \geq 0$, such that $q = \sum_{t=0}^{n_{\beta 2}} \delta(t)/\sum_{t=0}^{n_{\beta 1}} \delta(t)$ and individuals are willing to give up the same proportion $(1 - q)$ of utility of life duration irrespective of its duration $(n_{\beta 1})$. Generalized CPTO is consistent with the earlier defined generalized QALY model.

Investigators of TTO often assume linear utility of life duration, so that $\sum_{t=j-1}^{T} \delta(t) = T - j + 1$ or $\delta(t) = 1$ for each t. The model then simplifies to the *linear* QALY model, where equal weight is assumed to be attached to all health state values regardless of their timing. Then, the value of a health state $\beta_1 = \beta$ can be elicited by asking the

subject to give some period n_{FH} in full health (FH), followed by death (D), which makes him indifferent to a stated period in health state β (n_β), also followed by death. That is, the indifference relation ($h_1 = FH$, ..., $h_{nFH} = FH$, h_{nFH+1}=D, ..., $h_{n\beta} = D$) ~ ($h_1 = \beta$, ..., $h_{n\beta} = \beta$) is obtained. Under the linear QALY model, this indifference can be represented by the following equation:

$$n_{FH} * v(FH) + (n_\beta - n_{FH}) * v(D) = n_\beta * v(\beta). \quad (1)$$

If the value function over health is normalized so that $v(FH) = 1$ and $v(D) = 0$, we get the following simple expression for $v(\beta)$:

$$v(\beta) = \frac{n_{FH}}{n_\beta}. \quad (2)$$

This means that the value of health state β is obtained by dividing the number of years in full health by the number of years in health state β that is equivalent for the subject. Under this model, $q = n_{FH}/n_\beta$ and CPTO holds.

Pliskin et al. [12] gave an axiomatic derivation of a particular version of the QALY model for constant health profiles and proved that for the utility function of life duration to be a power function ($\sum_{t=j-1}^{T} \delta(t) = (T - j + 1)^c$ for $c > 0$, $\sum_{t=j-1}^{T} \delta(t) = -(T - j + 1)^c$ for $c < 0$ and $\sum_{t=j-1}^{T} \delta(t) = \log(T - j + 1)$ for $c = 0$), it is necessary that CPTO be valid (note that the linear function is a special case of this function for $c = 1$). For example, for $c = 0.5$, $n_\beta = 10$ and $v(\beta) = 0.6$, we have $U(t, \beta) = 10^{0.5} * 0.6 = 1.90$. The response to a TTO question (n_β) would then be $U(t, \beta) = (n_{FH})^{0.5} * 1 = 1.90 \Leftrightarrow n_{FH} = 1.90^2 = 3.61$, so $q = 3.61/10 = 0.36$. Similarly, for $n_\beta = 20$, we get $U(t, \beta) = 20^{0.5} * 0.6 = 2.68$ and the TTO response would be $U(t, \beta) = (n_{FH})^{0.5} * 1 = 2.68 \Leftrightarrow n_{FH} = 2.68^2 = 7.18$, so $q = 7.18/20 = 0.36$ again.

However, individuals may have a utility of life duration function that does not belong to the power family, but instead to some other parametric family. In that case, CPTO does not need to be confirmed, but it may very well be that the answers of such individuals do satisfy generalized CPTO. In other words, the generalized QALY model may still hold. When the utility function for life duration is exponential instead of a power, for example, CPTO may hold in terms of utilities of life years but not in terms of the absolute number of life years. In other words, the conventional CPTO assumption is not necessary for the generalized QALY model to hold.

Suppose, for example, that an individual has an exponential utility for life duration function, given by $\sum_{t=j-1}^{T} \delta(t) = 5(1 - e^{-0.1(T-j+1)})/(1 - e^{-2})$. However, as commonly the case, researchers may assume the function to be linear: $\sum_{t=j-1}^{T} \delta(t) = T - j + 1$. Figure 1 graphically shows these functions. The figure makes clear that the linear function implies no discounting, whereas future life

years are discounted in the exponential function, with the gap between linear and exponential utility increasing the larger the number of life years. Now suppose that this individual has completed two TTO tasks, one for $n_\beta = 10$ and one for $n_\beta = 20$. Her indifference values are $n_{FH} = 5$ and $n_{FH} = 7.73$, respectively. According to the linear function, we obtain $q = \sum_{t=0}^{5} \delta(t) \big/ \sum_{t=0}^{10} \delta(t) = 5/10 = 0.5$ for the former task, and $q = 7.73/20 = 0.39$ for the latter, which obviously violates CPTO. However, according to the exponential function, we have $q = 2.28/3.66 = 0.62$ and $q = 3.11/5 = 0.62$, respectively, so generalized CPTO is not violated.

Testing the generalized CPTO assumption requires the correction of TTO values for utility of life duration curvature. Recently, Attema et al. [10] have developed a new procedure to correct for this curvature. Attema and Brouwer [15] showed how this method can be used to correct TTO values.

CPTO can be violated due to other reasons than discounting as well (e.g. [3]), like loss aversion (i.e. the phenomenon that people are more sensitive to losses than to gains when viewed from a particular reference point [16]), and maximum endurable time (MET, i.e. the fact that some bad health states can only be endured during some period of time after which its value becomes negative). Depending on the magnitude and direction of these effects, CPTO may be violated in both directions or the effects may
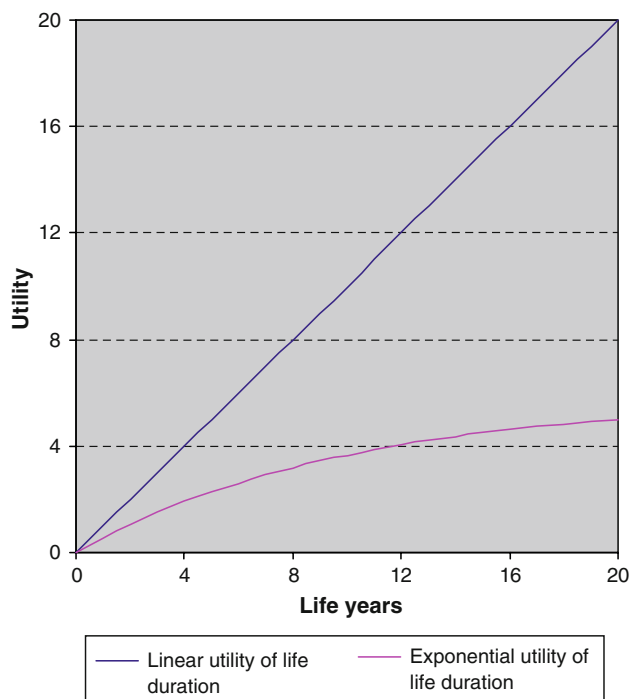

Fig. 1 Example of different utility for life duration functions

cancel out so that CPTO is not violated on the aggregate—or may mistakenly be perceived as not being violated if the violation itself is balanced by other effects. A better understanding of the validity of the assumption of CPTO, therefore, depends on a better understanding of the magnitudes of these effects and correcting for them as far as possible. In the conventional TTO procedure, loss aversion may cause subjects to be overly reluctant to give up life years, leading to relatively high health state values. MET will lead to higher values for bad health states for short durations, because for longer durations, extra time in that health state will be valued negatively. While loss aversion is more likely to be present in all TTO valuations, the presence of MET depends on the health state valued. Moreover, the influence of utility of life duration curvature is also present in normal TTO valuations, but it will likely vary with the time horizon chosen and can be corrected for. In the next section, we highlight the existing evidence regarding CPTO. In doing so, we indicate whether the performed studies corrected for utility of life duration curvature or used health states that could lead to MET effects.

Literature review

As noted previously, the evidence about the empirical validity of the CPTO assumption is mixed. This section highlights the existing evidence on the validity of the CPTO assumption. Our review adds to the one by Tsuchiya and Dolan [17], in that we also take into account studies published after 2002 and consider correction for utility of life duration. Table 1 presents the empirical studies regarding CPTO and summarizes their main results. The second column of this table shows the stimulus durations (n_β). The third and fourth columns indicate, respectively, whether or not a health state that may be susceptible to MET was used and whether or not the authors corrected for

Table 1 Overview of CPTO studies

Study	Life years used	MET health state	Utility correction	TTO long …TTO short	Sample
Sackett and Torrance [8]	3 months 8 years Life expectancy	Mixed	No	<	General population Patients
Pliskin et al. [12]	5 years 15 years	No	Yes	=	Pilot sample
Miyamoto and Eraker [18]	1, 2, 15, 16, 20, 24	No	No	<(p values not reported)	Patients
Hall et al. [19]	10% of LE 50% of LE Life expectancy	Mixed	No	=	Women 40–70 50% patients
Cook et al. [20]	1 year 12 years	No	No	=	Patients
Stiggelbout et al. [7]	3, 10, 15 years 3, 5, 10 years 5, 20, LE	No	No	<	Patients
Stalmeier et al. [25]	5, 10, 25, 50	Yes	Yes	=Except for $t = 5$	Students
Stalmeier et al. [21]	5, 10, 25, 50	Yes	No	≥	Students
Bleichrodt and Johannesson [4]	10 and 30 years	No	No	=	Students
Unic et al. [22]	5, 10 and higher	No	No	>	Healthy women
Kirsch and McGuire [23]	2 and 10 years	Mixed	No	≤	General population
Martin et al. [26]	5, 10, 15 years	No	Yes	<	Older cardiovascular disease patients (mean age 61)
Stalmeier et al. [24]	10 and 20 years	Yes	No	<(p values not reported)	Students Patients
Bleichrodt et al. [9]	13, 19, 24, 31, 38 years	No	No	≥	Students
Dolan and Stalmeier [5]	10 and 20 years	Yes	No	<	Students
Van der Pol and Roux [6]	20 and 50 years	No	Yes	=	Students

utility of life duration. The fifth column shows the results in terms of significance of the difference between shorter and longer durations at the 5% level. Finally, the sixth column indicates which kind of sample was analyzed. The table emphasizes the amount of variation in results. Some studies confirmed CPTO, but most studies rejected it. These violations, however, are not easily interpretable since CPTO was violated in both directions, i.e., when compared to longer durations, sometimes the proportion traded was relatively small for shorter durations and sometimes the proportion traded was larger. The finding that the TTO values of some health states were higher for short durations (i.e. relatively fewer life years were being traded in that case) was somewhat more common.

Sackett and Torrance [8] used a short, an intermediate and a long duration for a general population sample and a patient sample, and reported higher TTO values for short durations. Miyamoto and Eraker [18] reported evidence of no trade at all for durations smaller than 1 year for 25% of the subjects, whereas these people did trade off time for longer durations. Therefore, TTO values were higher for short durations than for long durations. Hall et al. [19] compared three different durations, 10, 50, and 100% of life expectancy. No violations of CPTO were found. Cook et al. [20] interviewed patients. Their TTO values were not significantly different between a duration of 1 year and one of 12 years. Stiggelbout et al. [7] used short and intermediate durations and found a violation of CPTO with TTO values for short durations being higher than those for long durations. Bleichrodt and Johannesson [4] used an intermediate and a long duration. They found no violation of CPTO at the aggregate level.

Stalmeier et al. [21] used bad health states to test for MET. They could not reject CPTO when comparing durations of 10 years and longer, but did find significantly lower TTO values for the 5-year horizon. Unic et al. [22] estimated TTO values for several durations in a sample of healthy women and found lower values for shorter durations. Kirsch and McGuire [23] compared a short and an intermediate duration and found mixed evidence. They reported higher TTO values for the short duration for bad health states, but no significant differences for moderate health states. They attributed this to subjects who valued additional time in a bad health state as worse than dead after some duration (i.e. MET). Stalmeier et al. [24] and Dolan and Stalmeier [5] found smaller TTO values for longer durations when comparing two intermediate durations in a severe health state. This may again have been caused by MET. Bleichrodt et al. [9] used five different durations that were not multiples of 5, so that they were not very susceptible to a proportional heuristic. These durations were of an intermediate and long-term nature. They found higher TTO values for long durations than short

durations for one procedure, indicating that people are willing to trade off relatively fewer life years for higher amounts. For another procedure, however, they could not reject CPTO.

We found four studies that corrected for utility of life duration curvature. Pliskin et al. [12] used a certainty equivalence (CE) question to correct for utility of life duration and found no violation at the aggregate level. However, their sample was very small. Stalmeier et al. [25] also found no violation of CPTO. They also corrected for utility curvature by means of the CE method and estimated several parametric models. Martin et al. [26] used three short and short-intermediate durations and corrected for utility curvature by means of CE questions. In a sample of cardiovascular patients, they found smaller TTO values for longer durations. Van der Pol and Roux [6] compared TTO values for a long-intermediate duration (20 years) and a very long duration (50 years). Further, they corrected for discounting (which is equivalent to utility of life duration in our definition) by means of one discounting question. They found no violation of CPTO, neither for uncorrected nor for individually corrected values.

To summarize, sixteen empirical studies of CPTO were found. Six of these did not reject CPTO, six found lower TTO values for longer durations, one found the opposite and three found mixed results. There is no clear influence of correcting for discounting nor is there a clear influence of MET. It appears difficult, therefore, to derive any definite answers from the literature regarding CPTO. Most evidence points toward higher values for short durations, yet all but one of these studies did not correct for utility of life duration curvature, which can strongly influence results, given the time horizons chosen. It appears that more evidence is required in order to understand the relationship between health state duration and valuation better.

Experiment

We performed an experimental study testing the conventional and generalized CPTO assumptions in the context of a larger study (see [10] and [15] for more details). Seventy students were recruited from different faculties of Erasmus University Rotterdam and were paid a fixed amount of €12.50 to join the experiment in a behavioral laboratory. The questions were administered by computer, using a program which was specially designed for this study. The experimental sessions were run with one or two subjects at a time, and lasted at most 45 min. One of the authors was present during the entire session to answer questions. Our experiment entailed two phases. In both phases, we used regular back pain as the health state of interest (β). Back pain is a common, easily understandable and non-severe

health state. The latter aspect minimizes the influence of MET. We described the health state in terms of the Euro-Qol 5D terminology as 11221. Further, we stressed to the subjects that the health state 'full health' (*FH*) meant they were able to function perfectly on all the five EuroQoL dimensions (11111), irrespective of their age. The descriptions of back pain and full health were printed on cards and handed to the subjects (see Appendix A). In the first phase of the experiment, we elicited the utility function for life duration, used to correct the answers to the TTO questions that were posed in the second phase.

The first phase was based on the notion that if utility of life duration is non-linear, the estimate obtained by (3) can be biased. Rather, one would then need to use the generalized QALY model. Without restrictions on $\delta(t)$ (and using the notation introduced earlier), the indifference implied by the TTO under this model would give the following equation (setting $v(FH) = 1$ again):

$$\sum_{t=0}^{n_\beta} \delta(t)v(\beta) = \sum_{t=0}^{n_{FH}} \delta(t). \tag{3}$$

Eliciting $v(\beta)$ thus requires the elicitation of the weights $\delta(t)$. We accomplished this by employing the recently proposed risk-free elicitation method of Attema et al. [10]. In short, the method presents a subject with two health scenarios. In the first scenario, the subject is in a good health state (g) at first. After some time, t, she moves to a worse health state (h) for the remainder of the total time period P. In the second scenario, the subject first is in the worse health state (h) and at time t moves to the better health state (g) for the remainder of P. The value of t is elicited that makes the subject indifferent between these scenarios. This value indicates the point where the period before t yields as much utility as the period after t. When t is smaller than the midpoint of the period P, this indicates concavity of the utility function over life duration. Then, as a result, raw TTO values will be biased downwards and correcting for this concavity results in higher utility scores. More detailed information about the exact shape of the utility function can be obtained by repeating this procedure (using the first estimate of t as input in the next exercise, etc). (See [10] for details on the risk-free method and [15] for details on correcting TTO values using this method.)

In the second phase of the experiment, we used two approaches to value the specified health state. First, in a conventional procedure, we fixed the duration of the health state with back pain (n_β) at 14 (*BP14*) and 27 years (*BP27*), respectively, and asked for the number of years in full health (n_{FH}) that they considered equivalent (Appendix B shows the formulation of the associated questions). We used a 1-shot response to obtain indifferences. If the generalized QALY model holds, we obtain the following equalities:

$$\sum_{t=0}^{n_{14}} \delta(t)v(\beta) = \sum_{t=0}^{n_{FH1}} \delta(t) \quad \text{and} \quad \sum_{t=0}^{n_{27}} \delta(t)v(\beta) = \sum_{t=0}^{n_{FH2}} \delta(t). \tag{4}$$

We are then able to compute $v(\beta)$ since we have the estimates of the utilities of life duration, which, using the notation of the experimental section, can also be expressed as:

$$\sum_{t=0}^{14} \delta(t)v(\beta) = q \sum_{t=0}^{14} \delta(t) \Leftrightarrow v(\beta) = q \quad \text{and}$$

$$\sum_{t=0}^{27} \delta(t)v(\beta) = q \sum_{t=0}^{27} \delta(t) \Leftrightarrow v(\beta) = q. \tag{5}$$

To summarize, if $v(\beta)$ varies for the two durations, then q will vary as well, and both generalized CPTO and the generalized QALY model will be rejected. It may be the case, on the other hand, that the power QALY model is rejected because $q = n_{FH}/n_\beta$ is different for the two durations, but that the generalized QALY model cannot be rejected because $q = \sum_{t=0}^{n_{FH}} \delta(t) / \sum_{t=0}^{n_\beta} \delta(t)$ may still hold depending on the precise form of the utility of life duration function. Our test, therefore, involved the comparison of $v(\beta)$ for the two durations according to both (2) and (3).

Second, in an alternative procedure, we fixed the duration in full health (n_{FH}) at 10 (*FH10*) and 22 (*FH22*) years, respectively, and asked for the number of years with back pain (n_β) that they considered equivalent (see Appendix B). The subsequent test of CPTO was performed along the same lines as specified earlier. The order of the questions was random.

Results

Fifty-six subjects (80%) were included in the analysis. The other 14 subjects were dropped from the sample because they had at least one inconsistent answer or had not understood the utility elicitation part. The median results did not change when including these subjects. The average age of the 56 included subjects was 21.8 years (sd = 2.99), and 36% were female.

In Table 2, we present some summary statistics concerning the uncorrected and corrected TTO values (these results were used to test the procedural invariance of TTO in Attema and Brouwer [27]). The difference between these values is around 0.05 (6%) for the *BP* questions and 0.13 (30%) for the *FH* questions (see last row of Table 2).

In order to test the conventional and generalized CPTO assumptions, we compared the small and the long duration

Table 2 TTO values

		BP14	BP27	FH10	FH22
	Uncorrected values				
	Mean	0.71	0.75	0.37	0.54
	Standard error	0.02	0.02	0.03	0.02
	Interquartile range	0.64–0.86	0.73–0.89	0.2–0.5	0.39–0.63
Difference between BP14 and BP27: 0.04 (6%, p < 0.01)	Range	0.07–1	0.07–1	0.1–0.94	0.22–0.94
Difference between corrected BP14 and corrected BP27: 0.05 (7%, p < 0.01)	Average number of years required/sacrificed	4.07	6.73	27.30	24.79
	Corrected values				
Difference between FH10 and FH22: 0.16 (44%. p < 0.01)	Mean	0.75	0.80	0.49	0.67
	Standard error	0.02	0.03	0.03	0.03
Difference between corrected FH10 and corrected FH22: 0.18 (37%, p < 0.01)	Interquartile range	0.69–0.88	0.69–0.94	0.35–0.64	0.52–0.82
	Range	0.09–1	0.13–1	0.09–0.94	0.29–0.98

for both elicitation procedures. For both uncorrected and corrected TTO values, CPTO was rejected, with the value being higher for longer than for shorter durations (paired t-tests, $p < 0.01$). This finding is in contrast with most of the aforementioned studies.

In the *FH* questions, our results seem to be caused partly by the large fraction of subjects (23%) that gave the same answer to *FH10* and *FH22*. In the *BP* questions, many subjects (34%) answered as if using a proportional heuristic, i.e. their answer to the second question was twice the amount of their answer to the first question. Because the input of the second question was somewhat lower than twice the amount of the first question ($27 < 2*14$), this resulted in a higher uncorrected TTO value for a longer duration for these subjects. We, therefore, repeated the analysis excluding these subjects, which still yielded a significant difference between the different durations. For *BP14* and *BP27*, CPTO was still rejected in the same direction ($p < 0.02$). For the *FH* questions, *FH22* also still yielded higher TTO values than *FH10*, both when excluding proportional heuristic subjects and when excluding subjects who gave the same answer to both questions ($p < 0.01$. For the *BP* questions, there were no subjects giving the same answers to both questions). As a result, for the alternative procedure, there was also a violation of CPTO in the opposite direction of most of the earlier found violations.

Summarizing, our results indicate that correcting for utility curvature and avoiding MET does not seem to be sufficient to restore the validity of the assumption of CPTO.

Conclusion

What can we infer from this study other than that we have added to the confusion regarding constant proportional trade-off? We believe some important observations need to be made.

First, the review of the literature shows that violations of CPTO are common. Though often the violation causes shorter durations to result in a decreased willingness to trade and, therefore, higher health state valuations, which may partly be due to failing to correct for utility of life duration, the opposite has also been shown. The reviewed studies differ in many respects, including the time horizon chosen and whether a correction for utility of life duration curvature has been applied. Not many studies did the latter. Of the four that did, three found no violation of CPTO, while one found that shorter durations resulted in higher valuations [26]. The fact that the latter study used relatively old patients (average age of 61) in their study may have influenced results, not only because of the way they view health problems, but also because of the fact that their subjective life expectancy may have been less than the projected ones. (See [28] on how this could bias the results.) Such differences between the studies make it difficult to derive general conclusions from the existing evidence.

The present study was clearly small and the sample consisted of students, hampering generalization. Still, we found a robust violation of CPTO for both uncorrected and corrected TTO values in our sample. Remarkably, this violation is in the opposite direction of most of the previously found violations of CPTO [17] and the only study correcting for utility of life duration curvature to find such a violation of CPTO.

After correcting for utility of life duration curvature, CPTO is still rejected, which indicates a more fundamental rejection of the QALY model. It seems that individuals do not trade off utility of life duration for health status at a constant rate, but instead at a rate that depends on the duration involved. For relatively long durations, like the ones used in our study, the amount of years traded is relatively low also after correction for utility of life duration curvature. Given this finding, the plausibility of relatively high TTO values for very short durations (who would trade off two days

to avoid low back pain when having only ten days left to live?) and the diverse violations of CPTO reported in the literature (which indeed must be related to the fact that TTO values vary strongly between studies as reported by Arnesen and Trommald [29]), it is interesting to hypothesize on the shape of this relationship between duration and trade-offs.

Given the importance of loss aversion in the TTO [3, 9], we hypothesize that a possible explanation for the variation in findings and, therefore, for a general relationship between health state duration and health state valuation in TTO is driven by this phenomenon. In a conventional TTO with a 'short' duration, loss aversion may relate especially to the amount of time left to live and stronger for smaller time horizons (durations). Loss aversion then causes subjects to be overly reluctant to give up life years, leading to relatively high TTO values. For 'long' durations, on the other hand, the absolute amount of years sacrificed may become dominant in the trade-off, i.e. the reference point of the subjects changes, with people being reluctant to trade off more than some absolute amount of time. Thus, the absolute amount of time remaining is most influential when the TTO uses short durations and the absolute amount of time sacrificed is most influential for longer durations. The result will be that individuals give up fewer years for short and long durations, and will be less driven by these considerations in between these two points, causing TTO values to be a U-shaped function of duration. The reviewed studies did not investigate short, intermediate and long durations together, and, hence, future research testing a wide range of durations using a within-subject design is called for to formally test this hypothesis.

For now, it seems that the (generalized) QALY model may be too simple, that the trade-off of life years is indeed not so constantly proportional and, therefore, that health state valuations may depend on the duration of these health states.

Appendix A: health state descriptions (translated from Dutch)

Card 1—Regular back pain

You have regular back pain. This has the following consequences for your functioning in daily life:

- You have no problems in walking about.
- You have no problems to wash or dress yourself.

- You have *some* problems with your usual activities.
- You have *moderate* pain or other discomfort.
- You are not anxious or depressed.

Card 2—Full health

You have no complaints and are in perfect health. This has the following consequences for your functioning in daily life:

- You have no problems in walking about.
- You have no problems to wash or dress yourself.
- You have no problems with your usual activities.
- You have no pain or other discomfort.
- You are not anxious or depressed.

Appendix B: instructions (translated from Dutch)

BP14 (BP27)

Imagine you have two options. In one option, you have 14 (27) years to live in an impaired health state, i.e. one with regular back pain (see the description), after which you will die painlessly. In the other option, you live for a smaller number of years, but in full health. Indicate how many years you would be willing to live in full health such that you are indifferent between these possibilities.

FH10 (FH22)

Imagine you have two options. In one option, you have 10 (22) years to live in full health, after which you will die painlessly. In the other option, you live for a larger number of years, but in an impaired health state, i.e. one with regular back pain (see the description). Indicate how many years you would be willing to live with regular back pain such that you are indifferent between these possibilities.

References

1. Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35*(11), 1095–1108.
2. Lamers, L. M., McDonnell, J., Stalmeier, P. F. M., Krabbe, P. F. M., & Busschbach, J. J. (2006). The Dutch tariff: Results and arguments for an effective design for national EQ-5D valuation studies. *Health Economics, 15*(10), 1121–1132.
3. Bleichrodt, H. (2002). A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Economics, 11*(5), 447–456.
4. Bleichrodt, H., & Johannesson, M. (1997). The validity of QALYs: An empirical test of constant proportional tradeoff and utility independence. *Medical Decision Making, 17*, 21–32.

5. Dolan, P., & Stalmeier, P. F. M. (2003). The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. *Journal of Health Economics, 22*(3), 445–458.

6. Van der Pol, M., & Roux, L. (2005). Time preference bias in time trade-off. *European Journal of Health Economics, 6*, 107–111.

7. Stiggelbout, A. M., Kiebert, G. M., Kievit, J., Leer, J. W., Habbema, J. D., & De Haes, J. C. (1995). The "utility" of the time trade-off method in cancer patients: feasibility and proportional trade-off. *Journal of Clinical Epidemiology, 48*(10), 1207–1214.

8. Sackett, D. L., & Torrance, G. W. (1978). The utility of different health states as perceived by the general public. *Journal of Chronic Disease, 31*, 697–704.

9. Bleichrodt, H., Pinto, J. L., & Abellan-Perpinan, J. M. (2003). A consistency test of the time trade-off. *Journal of Health Economics, 22*(6), 1037–1052.

10. Attema, A. E., Bleichrodt, H., & Wakker, P. P. (2010). *Measuring discounting and QALYs more easily and reliably, Working paper*. Rotterdam: Erasmus University Rotterdam.

11. Bleichrodt, H., & Gafni, A. (1996). Time preference, the discounted utility model and health. *Journal of Health Economics, 15*(1), 49–66.

12. Pliskin, J. S., Shepard, D., & Weinstein, M. C. (1980). Utility functions for life years and health status. *Operations Research, 28*, 206–224.

13. Doctor, J. N., & Miyamoto, J. M. (2003). Deriving quality-adjusted life years (QALYs) from constant proportional time tradeoff and risk posture conditions. *Journal of Mathematical Psychology, 47*(5), 557–567.

14. Osterdal, L. P. (2005). Axioms for health care resource allocation. *Journal of Health Economics, 24*, 679–702.

15. Attema, A. E., & Brouwer, W. B. F. (2009). The correction of TTO-scores for utility curvature using a risk-free utility elicitation method. *Journal of Health Economics, 28*(1), 234–243.

16. Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106*(4), 1039–1061.

17. Tsuchiya, A., & Dolan, P. (2005). The QALY model and individual preferences for health states and health profiles over time: A systematic review of the literature. *Medical Decision Making, 25*(4), 460–467.

18. Miyamoto, J. M., & Eraker, S. A. (1988). A multiplicative model of the utility of survival duration and health quality. *Journal of Experimental Psychology: General, 117*(1), 3–20.

19. Hall, J., Gerard, K., Salkeld, G., & Richardson, J. (1992). A cost utility analysis of mammography screening in Australia. *Social Science and Medicine, 34*(9), 993–1004.

20. Cook, J., Richardson, J., & Street, A. (1994). A cost utility analysis of treatment options for gallstone disease: methodological issues and results. *Health Economics, 3*(3), 157–168.

21. Stalmeier, P. F. M., Wakker, P. P., & Bezembinder, T. G. G. (1997). Preference reversals: Violations of unidimensional procedure invariance. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1196–1205.

22. Unic, I., Stalmeier, P. F., Verhoef, L. C., & van Daal, W. A. (1998). Assessment of the time-tradeoff values for prophylactic mastectomy of women with a suspected genetic predisposition to breast cancer. *Medical Decision Making, 18*(3), 268–277.

23. Kirsch, J., & McGuire, A. (2000). Establishing health state valuations for disease specific states: an example from heart disease. *Health Economics, 9*(2), 149–158.

24. Stalmeier, P. F. M., Chapman, G. B., de Boer, A. G. E. M., & van Lanschot, J. J. B. (2001). A fallacy of the multiplicative QALY model for low-quality weights in students and patients judging hypothetical health states. *International Journal of Technology Assessment in Health Care, 17*(4), 488–496.

25. Stalmeier, P. F., Bezembinder, T. G., & Unic, I. J. (1996). Proportional heuristics in time tradeoff and conjoint measurement. *Medical Decision Making, 16*(1), 36–44.

26. Martin, A. J., Glasziou, P. P., Simes, R. J., & Lumley, T. (2000). A comparison of standard gamble, time trade-off, and adjusted time trade-off scores. *International Journal of Technology Assessment in Health Care, 16*(1), 137–147.

27. Attema, A. E., & Brouwer, W. B. F. (2008). Can we fix it? Yes we can! But what? A new test of procedural invariance in TTO-measurement. *Health Economics, 17*(7), 877–885.

28. van Nooten, F., & Brouwer, W. B. F. (2004). The influence of subjective expectations about length and quality of life on time trade-off answers. *Health Economics, 13*(8), 819–823.

29. Arnesen, T., & Trommald, M. (2005). Are QALYs based on time trade-off comparable?—A systematic review of TTO methodologies. *Health Economics, 14*(1), 39–53.