Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Health Economics 22 (2003) 445–458
JOURNAL OF HEALTH ECONOMICS
www.elsevier.com/locate/econbase

The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic

Paul Dolan [a,b,*], Peep Stalmeier [c,d]

[a] *Department of Economics, Sheffield Health Economics Group, University of Sheffield, 30 Regene Street, Sheffield S1 4DA, UK*
[b] *Health Economics Research Programme, University of Oslo, Oslo, Norway*
[c] *Department of Medical Technology Assessment, University Medical Centre Nijmegen, Nijmegen, The Netherlands*
[d] *Joint Centre for Radiation Oncology Arnhem-Nijmegen (RADIAN), University Medical Centre Nijmegen, Nijmegen, The Netherlands*

Received 17 September 2002; received in revised form 21 November 2002; accepted 28 November 2002

Abstract

In order to calculate quality adjusted life years (QALYs) from time trade-off (TTO) responses, individual preferences are required to satisfy the *constant proportional time trade-off* (CPTTO) assumption. Respondents who use a simple proportional heuristic may appear to satisfy CPTTO but will in fact generate preference reversals for states that are associated with a maximal endurable time (MET). Using data from 91 respondents, the study reported here examines the extent to which valuations satisfy the CPTTO assumption and the extent to which they might be generated by the proportional heuristic. The results suggest that respondents are using a proportional heuristic that casts doubt on the validity of using the TTO method to calculate QALYs for health states that are associated with MET preferences.

JEL classification: I10

Keywords: Time trade-off; QALYS; Constant proportional time trade-off; Proportional heuristic; Maximal endurable time

* Corresponding author.
E-mail address: p.dolan@shef.ac.uk (P. Dolan).

0167-6296/03/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0167-6296(02)00120-0

1. Background

The quality adjusted life year (QALY) approach seeks to combine the value of changes in length and quality of life into a single composite measure. The main issues in the calculation of QALYs are concerned with how to describe and then value different states of health. To allow comparisons across different programmes, a health state descriptive system must allow for different dimensions of health to be combined to form an overall single index. A number of descriptive systems have been specifically designed for this purpose (Brazier et al., 1999), and the EQ-5D (see Fig. 1) is one of the most widely used of these (Brooks,

Mobility

1. No problems walking about
2. Some problems walking about
3. Confined to bed

Self-Care

1. No problems with self-care
2. Some problems washing or dressing self
3. Unable to wash or dress self

Usual Activities

1. No problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
2. Some problems with performing usual activities
3. Unable to perform usual activities

Pain/Discomfort

1. No pain or discomfort
2. Moderate pain or discomfort
3. Extreme pain or discomfort

Anxiety/Depression

1. Not anxious or depressed
2. Moderately anxious or depressed
3. Extremely anxious or depressed

Note: For convenience each composite health state has a five digit code number relating to the relevant level of each dimension, with the dimensions always listed in the order given above. Thus 21223 means:

2 Some problems walking about
1 No problems with self-care
2 Some problems with performing usual activities
2 Moderate pain or discomfort
3 Extremely anxious or depressed

Fig. 1. The EQ-5D descriptive system.

1996). Valuations for states of health can be elicited in a number of different ways although economists have tended to favour the standard gamble or the time trade-off (TTO) method (Dolan, 2000).

TTO valuations for a subset of EQ-5D health states have been elicited from over 3000 members of the UK general public (Dolan et al., 1996a). For each EQ-5D state in turn, respondents were asked to imagine that the state would last for 10 years and then they would die. They were then asked to think about how many years in full health, x, they would consider to be equivalent to living for 10 years in that EQ-5D state. The value for that state is then calculated as $x/10$. Values were also elicited for states rated as worse than dead but they are not discussed further here. A general population 'tariff' value for all 243 EQ-5D states has been estimated from direct observations on 42 of these states (Dolan, 1997). This tariff is increasingly being used in evaluative studies to calculate the number of QALYs gained from alternative interventions and is now recommended by the UK's National Institute for Clinical Excellence for Use in cost–utility studies.

2. Introduction

As a very simple example of how QALYs would be calculated using any tariff of health state values, consider an intervention that has known effects on health status and life expectancy. Assume also that health status does not vary over time. The QALY gain to an individual from this intervention will be

$$Q_1T_1 - Q_0T_0$$

where Q_1 and Q_0 represent pre- and post-intervention health status and T_1 and T_0 represent pre- and post-intervention life expectancy, respectively. A number of authors have shown, in various ways, that quite stringent assumptions have to be made for this algorithm to fully represent individual preferences (Pliskin et al., 1980; Miyamoto and Eraker, 1985; Bleichrodt et al., 1997).

At a general level, for the QALY algorithm to be a valid representation of the individual utility from changes in health, all of these models require that preference theory is satisfied; that is, they assume that individuals have clear well-defined preferences and are rational according to some set of axioms (such as those required in consumer theory, namely, completeness, reflexivity and transitivity). If preference theory is not satisfied, then the number of QALYs generated by an intervention cannot be assumed to represent utility, and so it may be possible that, under certain circumstances (such as when a health state becomes increasingly intolerable over time), less QALYs are preferred to more.

So that QALYs can be calculated from health state values generated by the TTO method, preferences are required to satisfy the constant proportional time trade-off (CPTTO) assumption. CPTTO entails that an individual is willing to sacrifice a constant proportion of her remaining life years in order to achieve a given improvement in her health, irrespective of the number of life years that remain. CPTTO may not be a very good representation of people's preferences in that the value of some less severe states may increase over time and the value of some more severe states may decrease over time, but it is required for the QALY model to hold when values have been elicited using the TTO method. CPTTO implies that

respondents in the EQ-5D study would have been willing to trade-off the same fraction of life years whether the states were described as lasting for 5 or 20 years instead of for 10 years. In this way, the same set of tariff values would have been generated irrespective of the specified duration.

The evidence relating to CPTTO is somewhat mixed. The early work in this area suggests that people trade-off a larger fraction of their remaining years of life as the number of these years increases (Sackett and Torrance, 1978; Pliskin et al., 1980; McNeil et al., 1981). More recent studies, however, have found that the largest trade-offs are for the shortest durations of life expectancy (Stalmeier et al., 1996, 2001; Unic et al., 1998). Encouragingly for the QALY model, there is also evidence that CPTTO might provide a reasonably good approximation of preferences. Stalmeier et al. (1996) found that the values for states lasting 10, 25 and 50 years were remarkably similar and Bleichrodt and Johannesson (1996) found that the mean values for states lasting 10 years and for states lasting 30 years were almost identical to one another.

If, on average, CPTTO does hold, it would only be applicable to those states for which more time is always preferred to less, and there is some evidence that this may not always be the case. For example, Sutherland et al. (1982) having found from a sample of 20 colleagues that the proportion preferring death to varying durations in each of five health states increased as the duration of the states increased, postulate that for some states there exists a *maximum endurable time* (MET) beyond which people do not wish to live. In other words, the value of those states becomes negative after some threshold. This concept has been reinforced by the results from a much larger general population study (Dolan, 1996). These results suggest that it might be inappropriate to calculate QALYs using TTO values of states for which there might come a point at which death would be preferred to any more time in those states, despite (positive) TTO values that suggest otherwise.

However, it is difficult to know precisely how to interpret the results from these studies since they did not explore the reasons behind people's responses. In a study designed to "get behind the numbers", Robinson et al. (1997) conducted a qualitative follow-up to the 'tariff study' which was designed to explore why so many more states were rated as worse than dead using the TTO method than when using a visual analogue scale (VAS). The conclusion that "the 10-year duration of the health state is more salient in the TTO than the VAS" suggests that MET preferences might exist. In other words, respondents might think about a short duration in the VAS task and rate the state as better than dead but contemplate 10 years in that state in the TTO task and consequently rate it as worse than dead.

For respondents with MET preferences, Stalmeier et al. (1997) show that *preference reversals* can occur. On average, respondents in their study were indifferent between 10 years with very frequent migraines and 4 years in full health, i.e. $[10, M] \sim [4, H]$. They were also indifferent, on average, between 20 years with very frequent migraines and 8 years in full health, i.e. $[20, M] \sim [8, H]$. If more years in full health are preferred to less, i.e. $[8, H] > [4, H]$, then this implies that $[20, M] > [10, M]$. However, 103 (out of 176) respondents exhibited MET preferences, i.e. $[20, M] < [10, M]$, and, of these, only 24 gave a lower number of equivalent healthy years to the longer duration. In other words, the majority of respondents who preferred 10 years of migraine to 20 years of migraine gave TTO responses that suggested the opposite preference ordering, and hence this results in a preference reversal. Such reversals falsify the assumptions underlying the QALY model,

so that using a value of 0.4 (i.e. 8/20) for migraine has no valid interpretation within the QALY model for those with preference reversals.

Preference reversals involving money gambles are now well documented in the economics literature (Grether and Plott, 1979; Tversky et al., 1990). Here, out of two gambles, respondents tend to place a higher certainty equivalent value on the higher pay-off/smaller probability gamble but choose to play out the smaller pay-off/higher probability one. Respondents in the Stalmeier et al. (1997) study would not have committed a similar reversal if their preferences had been consistent with CPTTO. Instead, the authors conclude that most respondents with a MET preference use a *proportional heuristic*, whereby they trade-off a constant proportion of their remaining life expectancy, not because they satisfy CPTTO but because they simply double their stated number of years in full health as the number of years in the poor health state doubles.

Since people's health state valuations are likely to be rather imprecise and constructed partly during the process of elicitation (Dolan, 2000), the proportional heuristic provides respondents with a simple 'rule of thumb' to help them generate responses to quite difficult valuation questions. For those who use the proportional heuristic in combination with a preference reversal, values derived using the QALY model clearly have no valid interpretation within that model. However, for those who use the proportional heuristic in the absence of a preference reversal, values derived using the QALY model may have a valid interpretation within the QALY model.

In the study reported here, we assess respondents' trade-offs when asked to value one of the EQ-5D health states using the TTO method for durations of 10 and 20 years. By directly asking respondents whether they would prefer to live for 10 or 20 years in that state, we are able to examine the extent to which their valuations satisfy the CPTTO assumption and the extent to which they might be generated by the proportional heuristic. If CPTTO holds, then $[10, \mathrm{P}] \sim [x, \mathrm{H}]$, $[20, \mathrm{P}] \sim [2x, \mathrm{H}]$ and, crucially, $[20, \mathrm{P}] > [10, \mathrm{P}]$, where P is the EQ-5D health state and x is the number of years in full health, H. We would have strong evidence of the proportional heuristic if the same set of TTO responses were generated (i.e. $[10, \mathrm{P}] \sim [x, \mathrm{H}]$ and $[20, \mathrm{P}] \sim [2x, \mathrm{H}]$) but if $[10, \mathrm{P}] > [20, \mathrm{P}]$. This would show that MET preferences exist but that these are not reflected in people's TTO responses. The study was designed specifically to provide qualitative insights into the reasons behind respondent preferences and to gain insights into why shorter or longer durations in the same health state might be preferred. This information would thus shed some light on the descriptive validity of the tariff for calculating QALYs.

3. Methods

There is no existing evidence about which (if any) of the EQ-5D states might be associated with MET preferences. In choosing the state for valuation in this study, we were looking for a state that some respondents would prefer to live in for 10 years rather than for 20 years. In other words, they would think that, at some point between these durations, the state would become intolerable. After discussion with our colleagues, the state 21223 was chosen for valuation in this study (see Appendix A). This state was not one of the 42 directly valued in the UK general population study but those states were chosen with a view to estimating

values for the remaining 200 states. State 21223 is an entirely plausible combination of dimensions and it has a relatively low tariff value of 0.186, so it is possible that some respondents would prefer to live in this state for 10 years rather than for 20 years.

The questionnaire comprised of a self-completion booklet containing six questions. In question 1, respondents were asked to describe their own health on the EQ-5D. In questions 2 and 3, they valued state 21223 for durations of 10 years (TTO_{10}) and 20 years (TTO_{20}) using a TTO protocol similar to the 'no-props' variant reported in Dolan et al. (1996b). In question 4, respondents were asked to write down whether they thought there ought to be any relationship between their responses to the two TTO questions and, if so, what it ought to be. In question 5, they were then asked to choose between living for 10 years in 21223 or for 20 years in the same state. In question 6, they were asked to write down the reasons for this choice.

Response categories for TTO_{10} were presented to respondents in the standard way; that is, the numbers of years in full health were presented in descending order. Response categories for TTO_{20} were presented in the reverse order such that the number of years in full health increased (see Appendix A). This reversed titration was chosen for the following reason. For those with MET preferences, a preference reversal (Stalmeier et al., 1997) occurs if the number of equivalent healthy years for TTO_{20} is greater than the number of equivalent healthy years for TTO_{10}. If TTO_{20} was to titrate downwards from 20 years in full health, then respondents may have been induced through anchoring to state a higher number of equivalent healthy years than their true preferences would suggest, thus increasing the likelihood of a preference reversal. Of course, respondents might just as easily anchor on 0 year in TTO_{20} but this would mean that they would state a relatively low number of equivalent healthy years. This would reduce the likelihood of preference reversals, and means that any reversals that remain would be robust in that they exist despite a framing effect that mitigates them.

The two qualitative questions (4 and 6) were structured to facilitate understanding of the written responses and to allow those responses to be coded more easily. Such focused questions are best suited to the self-completion format employed in this study. For the qualitative analysis, both authors went through the written responses and coded the responses to questions 4 and 6 according to the categories shown in Table 1. Any discrepancies between the two sets of codings were resolved through discussion.

The questionnaire was handed out to a group of undergraduate economics students in the UK and a group of undergraduate medical informatics students in The Netherlands. Both groups were given verbal instructions about how to complete the questionnaire and they could ask questions at any time about how to fill it in. The aim of this study was not to achieve a representative sample but rather to draw some general conclusions about the TTO preferences of British and Dutch respondents.

The Kolmogorov–Smirnov test was used to test for normality in the quantitative responses to TTO_{10} and TTO_{20}. Because there were only very weak departures from normality, parametric tests were used in the subsequent analyses. A MANOVA was used to test for significant differences between questions and across sub-groups based on self-reported health state, country of administration and qualitative response. Whether the frequency of MET preferences differed according to the qualitative responses was tested using Chi-square. Scatterplots are used to provide visual representations of the relationship between TTO_{10} and TTO_{20} by the various qualitative response categories.

Table 1
Coding of the qualitative responses to questions 4 and 6

Response category[a]	Definition of category
Question 4: "Do you think that there should be any relationship between the answer you gave to question 2 and the answer you gave to question 3? If you do, could you say something about what this relationship is?"[b]	
(1) –	Empty response
(2) Yes, proportional	"Yes" with explicit mention of a proportional relationship
(3) Yes	"Yes" without mention of a proportional relationship
(4) No	Explicit "no", meaning there is no relationship
(5) Longer is worse	Mention of "intolerable health state after long period"
Question 6: "Could you please give the reasons why you made the choice you did?" ('choice' refers to that offered in the 'maximal endurable time' (MET) question 5)[b]	
(0) No MET	No MET preference
(1) MET duration	MET preference and mention of duration
(2) MET other	MET preference and mention of things other than duration

[a] The authors coded these. The responses themselves were open-ended.
[b] This was the exact wording of the questions.

4. Results

There was a total sample of 91 respondents, 61 from the UK and 30 from The Netherlands. There were 51 male respondents and 40 female respondents. A Mann–Whitney *U*-test confirmed that there were no significant differences ($P < 0.05$) between the responses from the two countries, and that valuations did not differ by sex or self-reported health status. There were no missing data. The mean value for TTO_{10} was 0.47 with a standard deviation of 0.24. The mean and standard deviation for TTO_{20} were 0.43 and 0.25, respectively. The very similar means and standard deviations for TTO_{10} and TTO_{20} suggest that, as in the study by Stalmeier et al. (1997), proportionality holds at the aggregate level. The overall correlation between TTO_{10} and TTO_{20} is 0.78 suggesting that, despite considerable variation in values across respondents, the two sets of responses are related to one another at the individual level.

Table 2 shows the correlation between values for the two durations according to the five categories used to code responses to question 4. It can be seen that those who state that there

Table 2
Correlations between TTO_{10} and TTO_{20} for response categories for question 4 (Q4)

Response category	*N*	Correlation (Pearson's coefficient *r*)
(1) –	20	0.82
(2) Yes, proportional	18	0.87
(3) Yes	43	0.80
(4) No	7	0.33
(5) Longer is worse	3	−0.26
Entire sample	91	0.78

Table 3
TTO values by duration and response categories for question 4 (Q4)

Response category	*N*	10 years mean (S.D.)	20 years mean (S.D.)
(1) –	20	0.432 (0.211)	0.404 (0.217)
(2) Yes, proportional	18	0.558 (0.204)	0.546 (0.204)
(3) Yes	43	0.470 (0.239)	0.427 (0.274)
(4) No	7	0.425 (0.289)	0.425 (0.168)
(5) Longer is worse	3	0.200 (0.246)	0.025 (0.043)
Entire sample	91	0.467 (0.235)	0.432 (0.251)

should be no relationship between the two TTO valuations have very low correlations as compared to those who suggest that a relationship exists. The highest correlation is found for those who explicitly mention proportionality. These results lend some support to the validity of the qualitative responses. Table 3 shows the TTO values that were found to be slightly lower for the longer duration ($P = 0.04$). It can be seen that values differ more markedly by response categories ($P = 0.03$) where, interestingly, for categories 2–4, values become progressively lower as respondents become less explicit about the relationship between TTO_{10} and TTO_{20}.

Table 4 shows that 44 out of 91 respondents (48%) had MET preferences; that is, they preferred less time to more time in state 21223. As it turns out, this rate is quite similar to that reported by Stalmeier et al. (1997) and, using some of the same data, Stalmeier et al., 2001, after controlling for the difference in utilities between those data and ours. In Fig. 3 in Stalmeier et al. (2001), the rate of MET preferences is about 60% for people who give the migraine health state a value of 0.45, which is about the average value for 21223 in our study. Moreover, the preference reversal rate is 38/44 (86%) in this study, which is comparable to the rate of 79% and reported by Stalmeier et al. (1997).

Table 3 shows that only three respondents explicitly state that "longer is worse" in the question about the relationship between TTO_{10} and TTO_{20}, which suggests that respondents really do think differently about the two TTO valuation questions as compared to the question that asks them to choose between 10 and 20 years in 21223. There is little correlation between the five response categories for question 4 and whether or not MET preferences are present. However, when question 4 is recoded into two categories according to whether or not the relationship between TTO_{10} and TTO_{20} is seen as being proportional (i.e. category 2) or

Table 4
TTO values by duration and MET preferences

MET preference	*N*	10 years mean (S.D.)	20 years mean (S.D.)
No	47	0.555 (0.187)	0.554 (0.203)
Yes	44	0.373 (0.247)	0.302 (0.233)
Entire sample	91	0.467 (0.235)	0.432 (0.251)

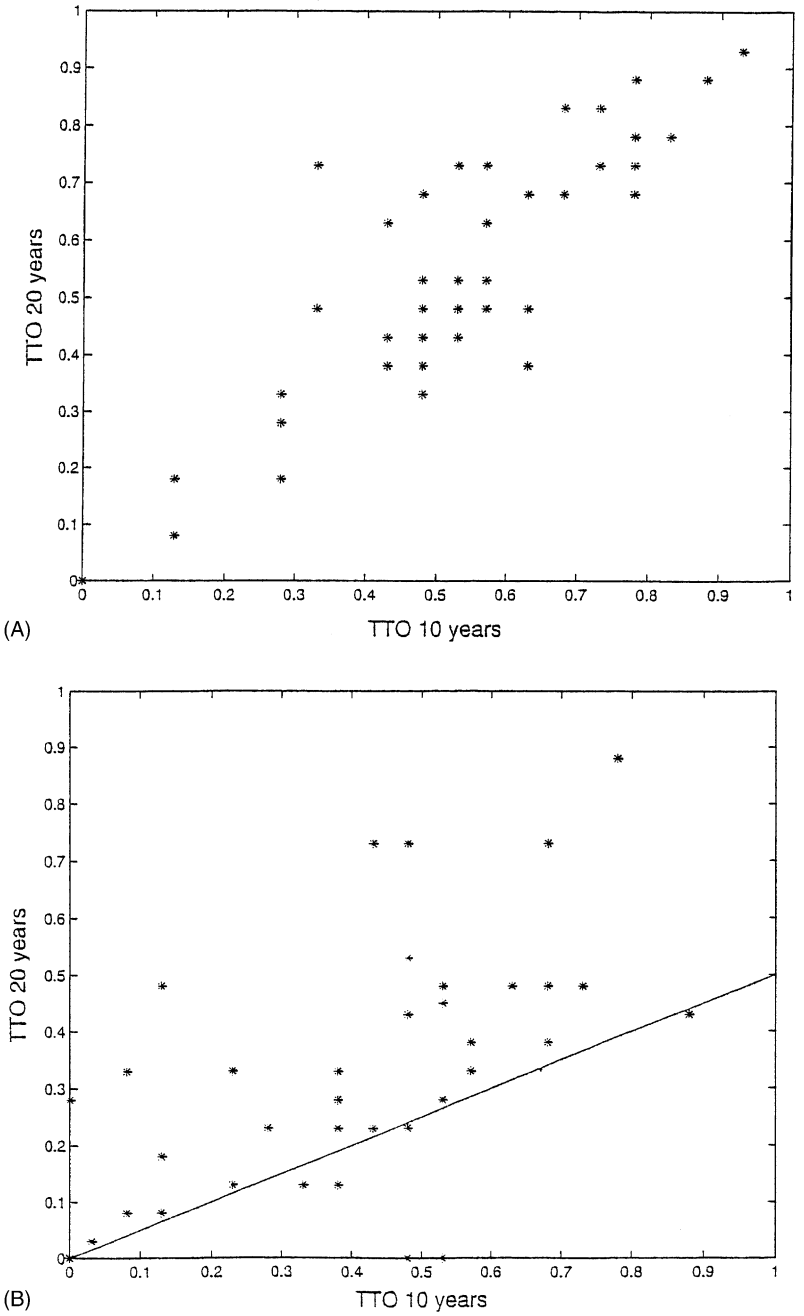

Fig. 2. Scatterplots of TTO_{10} vs. TTO_{20} for those without MET preferences who prefer the longer duration in health state 21223 (A), and those with MET preferences who prefer the shorter duration (B).

Table 5
TTO values by duration and response categories for question 6 (Q6)

Response category	*N*	10 years mean (S.D.)	20 years mean (S.D.)
No MET	47	0.555 (0.187)	0.554 (0.203)
MET duration	18	0.417 (0.229)	0.278 (0.204)
MET other	26	0.342 (0.258)	0.319 (0.243)
Entire sample	91	0.467 (0.235)	0.432 (0.251)

not (i.e. categories 1, 3–5), those who suggest that it ought to be proportional are less likely to have MET preferences than those who do not suggest such a relationship (4 out of 18 as compared to 40 out 73). This lends qualitative support to the results of earlier studies (such as Stalmeier et al., 1997) that suggest (but do not state explicitly) that CPTTO holds for those who prefer more time to less in a given health.

Fig. 2A shows a scatterplot of TTO_{10} with TTO_{20} for those respondents who prefer 20–10 years in state 21223 (i.e. those without MET preferences). Notice that many points lie on or near the 45° line, thus highlighting the proportional relationship between TTO_{10} and TTO_{20}. Fig. 2B shows the analogous scatterplot for those with MET preferences who prefer 10–20 years in state 21223. As would be expected, many more points lie below the 45° line, i.e. where $TTO_{10} > TTO_{20}$.

Table 4 shows quite clearly that it is those with MET preferences who are responsible for the lower mean value for TTO_{20} than for TTO_{10} ($P < 0.01$). Table 5, which shows the mean values for TTO_{10} and TTO_{20} according to the qualitative responses to question 6, suggests that it is a smaller subset of respondents still who are responsible for $TTO_{10} > TTO_{20}$, i.e. those with MET preferences who explicitly mention the duration of the state as the reason for their MET preference. The marked differences between the values of the two groups of respondents with MET preferences suggest that further qualitative research is needed into the precise reason(s) why an MET preference exists.

5. Discussion

One of the assumptions of the QALY algorithm is that the value of a health state is a linear function of the time spent in that state. For this assumption to hold for valuations elicited using the TTO method, it is necessary for respondents to trade-off a constant proportion of their remaining life expectancy, irrespective of how long or short that life expectancy is. The use in economic evaluations of the TTO-based tariff for the EQ-5D is premised on this assumption.

From the study reported in this paper, it would appear that CPTTO holds for EQ-5D state 21223 for durations of 10 and 20 years. Or, more accurately, CPTTO would seem to hold for those who prefer 20 years in state 21223 to 10 years in that state. Such respondents have higher values than those who prefer the shorter duration. Consequently, it would seem reasonable to conclude that the QALY model is a good approximation of the preferences

of those who prefer longer to shorter durations in various states of ill health. In turn, then, it would appear that the EQ-5D tariff is suitable for use in evaluating those interventions where the pre- and post-treatment health states are unlikely to be associated with a maximal endurable time (MET).

From the scatterplot in Fig. 2B, it appears that CPTTO could also hold for many of those 48% of respondents with MET preferences. Such respondents have lower values but are still on or near the 45° line. According to preference theory, however, CPTTO should not hold for these respondents. Since they strictly prefer 10 years in 21223 to 20 years in that state, all of those respondents with MET preferences should be found below the 22.5° line in Fig. 2B for the assumptions of the QALY model to hold. That is, the healthy years equivalent for the longer duration should be less than the healthy years equivalent for the shorter duration to satisfy the MET preference and hence avoid the preference reversal. It is striking how few respondents there are in this region. This suggests that respondents either simply did not understand the valuation exercise (which we would expect to have articulated to us during the elicitation process) or that preference theory does not hold for the majority of respondents with MET preferences.

Indeed, rather than satisfying the CPTTO assumption, these results suggest that respondents—perhaps even including those without MET preferences—are using a proportional heuristic when responding to TTO questions. That is, if the duration of the state doubles the number of equivalent healthy years approximately doubles. It just so happens that this heuristic is indistinguishable from the CPTTO assumption for those respondents without MET preferences. Whether respondents satisfy CPTTO or use a proportional heuristic may not be a serious problem for states where longer durations are nearly always going to be preferred to shorter durations. So, for example, for many of the EQ-5D health states that contain only levels one or two on each of the dimensions, it might be appropriate to use the set of tariff values to generate QALYs on the assumption that a longer life is preferred to a shorter one. However, for more severe states, such as those containing one or more dimensions at the third and most severe level on the EQ-5D, it might be misleading to use the tariff to calculate QALYs in that such states might be associated with MET preferences. Therefore, a general set of health state values should be used with caution in a cost–utility analysis where the health effects are expressed in terms of QALYs.

It is worth noting that this study is different from the ‘tariff study’ in a number of respects. An obvious difference relates to the samples—a general population sample of 3000 people was used to generate the tariff as compared to a sample of 91 students reported here. However, there is a large body of literature that suggests that violations of preference theory, such as that implied by preference reversals, are remarkably robust across respondent characteristics (e.g. see Camerer (1995) for a review of the literature relating to individual decision making under uncertainty). Therefore, we hypothesise that preference reversals would also have been observed in the tariff study if respondents had been asked questions to test for this. However, it would be advisable in future research to test for preference reversals in a general population sample, where it would be possible to test whether reversals are related to any background variables, such as age and experience of illness, which have been found to affect health state valuations.

A potentially more serious problem relates to the fact that respondents in this study were only allowed to consider the state to be *equivalent to*—rather than worse than—dead. In this way, some respondents may have been persuaded to rate the state as strictly better than dead. This may be particularly true when valuing the state for a 10-year duration since, assuming that respondents work their way down the page, 0 year in full health is the very last response category they see. For the 10-year duration, 21223 was considered to be at least as bad as death by 7% of respondents. The comparable figure for the 20-year duration was only 8%. Since the response categories were presented in the reverse order in this question (so as to minimise the observed violations of preference theory), this suggests that most respondents did consider this state to be better than dead in the context of this study. Thus, it is highly unlikely that the MET preferences found in 48% of our respondents are a consequence of 21223 being considered worse than death for both 10 and 20 years. Rather, it would seem that respondents exhibit diminishing marginal utility for state 21223, such that it eventually becomes worse than dead.

6. Concluding remarks

The results from this study suggest that respondents with MET preferences are using a proportional heuristic when responding to TTO questions. Even though our data do not prove this, respondents without MET preferences might also be using the same heuristic. Therefore, the CPTTO-like patterns of responses reported elsewhere may be little more than the proportional heuristic in disguise. All of this suggests that TTO values may not fully comply with the preference theory upon which their interpretation for use in economic evaluations is based.

Nevertheless, the TTO values that have been successfully elicited over the last couple of decades are still useful. If the proportional heuristic holds, the values can be used to rank health states and to provide information about the relative distances between them. Moreover, the number of QALYs estimated for profiles of health that contain relatively mild states (where MET preferences are less likely) provide a good approximation of the relative health benefits from alternative interventions. What the proportional heuristic does mean, however, is that the values may not fully comply with the properties required by the QALY model and could in some circumstances (e.g. for severe states that last for relatively long periods of time) lead to the wrong conclusions about the ordering of different profiles of health.

Appendix A. Question 3

Imagine that you would live for 20 years in the health state described under ‘Life A’ below, after which time you would die. Now imagine that there is a treatment which would shorten your life but would mean that you would live in the health state described under ‘Life B’ below. For each of the choices, please indicate whether you would prefer Life A or B.

Life A: Some problems in walking about; no problems with self-care; some problems with usual activities; moderate pain or discomfort extremely anxious or depressed (years)	Circle A or B		Life B: No problems in walking about; no problems with self-care; no problems with usual activities; no pain or discomfort not anxious or depressed (years)
20	A	B	0
20	A	B	1
20	A	B	2
20	A	B	3
20	A	B	4
20	A	B	5
20	A	B	6
20	A	B	7
20	A	B	8
20	A	B	9
20	A	B	10
20	A	B	11
20	A	B	12
20	A	B	13
20	A	B	14
20	A	B	15
20	A	B	16
20	A	B	17
20	A	B	18
20	A	B	19
20	A	B	20

References

Bleichrodt, H., Johannesson, M., 1996. The validity of QALYs: an experimental test of constant proportional trade-off and utility independence. Medical Decision Making 17, 21–32.

Bleichrodt, H., Wakker, P., Johannesson, M., 1997. Characterizing QALYs by risk neutrality. Journal of Risk and Uncertainty 15, 107–114.

Brazier, J.E., Deverill, M., Green, C., 1999. A review of the use of health status measures in economic evaluation. Journal of Health Services Research and Policy 4, 174–184.

Brooks, R., 1996. EuroQol: the current state of play. Health Policy 37, 53–72.

Camerer, C., 1995. Individual decision making. In: Kagel, J.H., Roth, A.E. (Eds.), Handbook of Experimental Economics. Princeton University Press, Princeton.

Dolan, P., 1996. Modelling valuations for health states: the effect of duration. Health Policy 38, 189–203.

Dolan, P., 1997. Modelling valuations for EuroQol health states. Medical Care 11, 1095–1108.

Dolan, P., 2000. The measurement of health-related quality-of-life for use in resource allocation decisions in health care. Handbook of Health Economics. North-Holland, Amsterdam, Chapter 32, pp. 1723–1760.

Dolan, P., Gudex, C., Kind, P., Williams, A., 1996a. The time trade-off method: results from a general population study. Health Economics 5, 141–154.

Dolan, P., Gudex, C., Kind, P., Williams, A., 1996b. Valuing health states: a comparison of methods. Journal of Health Economics 15, 209–231.

Grether, D., Plott, C.R., 1979. Economic theory of choice and the preference reversal phenomenon. American Economic Review 69, 623–638.

McNeil, B.J., Weichselbaum, R., Pauker, S.G., 1981. Speech and survival: trade-offs between quality and quantity of life in laryngeal cancer. The New England Journal of Medicine 305, 982–987.

Miyamoto, J.M., Eraker, S.A., 1985. Parameter estimates for a QALY utility model. Medical Decision Making 5, 191–213.

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

Robinson, A., Dolan, P., Williams, A., 1997. Valuing health status using VAS and TTO: what lies behind the numbers? Social Science and Medicine 45, 1289–1297.

Sackett, D.L., Torrance, G.W., 1978. The utility of different health states as perceived by the general public. Journal of Chronic Diseases 31, 697–704.

Stalmeier, P.F.M., Bezembinder, T.G.G., Unic, I.J., 1996. Proportional heuristics in time trade-off and conjoint measurement. Medical Decision Making 16, 36–44.

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.

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

Sutherland, H.J., Llewelyn-Thomas, H., Boyd, N.F., Till, J.E., 1982. Attitudes toward quality of survival. The concept of “maximum endurable time”. Medical Decision Making 2, 299–309.

Tversky, A., Slovic, P., Kahneman, D., 1990. The causes of preference reversal. American Economic Review 80, 204–217.

Unic, I., Stalmeier, P.F.M., Verhoef, L.C.G., van Daal, W.A.J., 1998. Assessment of the time trade-off values for prophylactic mastectomy of women with a suspected genetic predisposition to breast cancer. Medical Decision Making 18, 268–277.