The impact of losses in income due to ill health: does the EQ-5D reflect lost earnings?

Tilling, C and Krol, M and Tsuchiya, A and Brazier, J and van Exel, J and Brouwer, W

The University of Sheffield

2009

HEDS Discussion Paper 09/04


ScHARR
School of Health and
Related Research

The University Of Sheffield.

ScHARR

Health Economics and Decision Science Discussion Paper Series

No. 09/04

The impact of losses in income due to ill health: Does the EQ-5D reflect lost earnings?

Carl Tilling[1], Marieke Krol[2], Aki Tsuchiya[1],
John Brazier[1], Job van Exel[2], Werner Brouwer[2]

1. School of Health and Related Research, University of Sheffield
2. Institute for Medical Technology Assessment, Erasmus University Rotterdam

Corresponding author:
Carl Tilling
School of Health and Related Research
University of Sheffield
Regent Court
30 Regent Street
Sheffield
S1 4DA
UK
Email : c.tilling@sheffield.ac.uk

Introduction

In order to allocate scarce health care resources, the National Institute for Health and Clinical Excellence (NICE) promotes the use of economic evaluation to decide which health care technologies should be recommended for use in the NHS. The 'reference case' in the current NICE *Guide to the Methods of Technology Appraisal* (2008) does not include any indirect or productivity costs in the evaluation, since the perspective used for costs is that of the NHS and Personal Social Services, (the same applies to the previous *Guide*; NICE, 2004). At the same time, the *Guide* aims to capture "all health effects on individuals" on the outcomes side. Thus, if people take into account the impact of lost earnings in the health valuation exercise, then indirect costs would be included in the analyses, albeit implicitly. On the other hand, the House of Commons Health Select Committee has recently recommended that "wider benefits and costs […] be more fully incorporated into NICE's assessment" (Health Select Committee, 2007) so we could see future legislative changes that require productivity costs to be incorporated more explicitly in future economic evaluations for NICE.

Either way, two key questions in the context of UK health policy are: do the published preference indices for EQ-5D reflect the impact of lost earnings? Are we currently implicitly including indirect costs in our analyses? It is crucial to investigate whether or not individuals take into account any possible impact of lost income in health state valuation exercises. If respondents do consider income effects, and these considerations change valuations, then these effects would need to be excluded both under the current NICE reference case, or where productivity costs are included in the numerator to avoid double counting. This study adapts the study design used to generate population value sets for EQ-5D, as first used in the Measurement and Valuation of Health (MVH) Study (Dolan, 1997), and carries out valuations of hypothetical EQ-5D states using Time Trade Off (TTO) exercises through an online survey administered in the Netherlands. Furthermore, this study uses a number of different TTO questions to explore the impact of losses in income on the valuation of hypothetical health states, and to determine the relationship between income and health. To understand the effect that income considerations may have in health state valuation exercises it is necessary to understand the relative importance of health and income when valued both simultaneously and independently. For example, would the same loss of income be valued as worse when it is associated with worse health states? Specifically our objectives are to (a) examine whether EQ-5D health state values, obtained through online TTO, reflect losses in income due to ill health; (b) examine the impact of including specific ex-post instructions to consider, or not to consider, income changes when hypothetical EQ-5D states are valued, on the health state values; (c) examine how the above impact is distributed across the different dimensions of EQ-5D, and (d) explore the possible interactions between health and income in health state valuation.

Background

An important component of benefits in economic evaluations from the societal perspective is the gains in productivity resulting from getting sick individuals back into paid employment. Traditionally, improved productivity as a result of healthcare was included as a negative cost in the numerator of the Cost-Effectiveness ratio. This

was initially done through the human capital approach (Weisbrod, 1961; Rice and Cooper 1967). Under this approach lost production as a result of morbidity or mortality is valued by measuring time lost from work and multiplying this with the gross wage of the involved individual. The relevant period of time over which costs/savings are measured is the total period of time in which the person is unable to be productive compared to the alternative scenario. In the case of disability or mortality this can obviously amount to a considerable length of time (until the age of retirement).

An alternative approach to including productivity costs in monetary terms is the friction cost method (Koopmanschap and van Ineveld 1992; Koopmanschap and Rutten 1993; Koopmanschap et al 1995). This method takes account of involuntary unemployment and the possibility of replacement. When a worker leaves the workforce due to morbidity or mortality they can be replaced by a previously unemployed member of society. Therefore, although there are replacement costs associated with recruiting and training a new worker and productivity costs in the transition (friction) period, there are no long term production losses. Estimates of the friction cost and human capital methods do not differ significantly in the case of short term absence. However, in the case of long term morbidity and mortality the differences are, as expected, substantial.

The practice of valuing productivity costs in monetary terms, in the numerator of the Cost-Effectiveness ratio, was challenged by the controversial recommendations of the "Washington Panel" (Gold *et al.* 1996). They recommended measuring most of the productivity costs (viz. replacement costs included in the numerator) through quality of life measurement in the denominator of the C/E ratio in terms of health effects, using changes in income as a proxy for productivity costs. In other words, they assume that when people answer health state valuation questions (e.g. time trade off questions) they take into account the effect of ill health on their ability to work and hence on their income (even when the question is silent on the issue), so that the value set for measures such as EQ-5D already incorporate the impact of ill health on productivity. The Panel, therefore, argued that to include changes in productivity in the numerator is a form of double counting.

The Panel's recommendations received considerable criticism for both theoretical and empirical reasons. Theoretically, personal income is a poor proxy for productivity costs owing to the existence of private insurance and social security benefits (Brouwer et al,1997a). In addition, strictly speaking, there can be people who are productive, but not in paid work, whose productivity should be in a societal all-encompassing evaluation. Importantly, empirically, when the recommendations were published there was no evidence to support the Panel's key assumption, that health state valuation exercises evaluate not just the health related quality of life of hypothetical states, but also the impact of lost earnings due to ill health. Efforts have been made in recent years to investigate this assumption but these studies are generally characterised by small and unrepresentative samples and the results were inconclusive and inconsistent due to important differences in design (Tilling *et al.* 2009).

Eight studies have attempted to address people's considerations on income in health state valuation exercises. Four of these have evaluated hypothetical EQ-5D states (whilst the others have used specific conditions). The first to value EQ-5D states

(Krol et al, 2006) asked 185 members of the Dutch general population to value three states using a visual analogue scale (VAS), and found that without specific instructions, 36% of the respondents stated to have spontaneously included effects of income (determined through follow-up question). Krol *et al.* (2009) replicate the above study using TTO (210 respondents). They found that 64% of respondents included income effects without instructions on the matter. Krol et al. (2006) found that valuations were revised upwards for 2 of the 3 states when those that had included income effects were instructed not to. Krol et al. (2009) found no significant differences following instruction. Brouwer *et al.* (2008), ask 75 members of the Dutch general population to value EQ-5D states using VAS and found that 69% of respondents did not consider income effects. They found the incorporation of income effects to be insignificant in health state valuations. A recently published study by Davidson and Levin (2008) asked 200 Swedish students to complete TTO and VAS exercises. They found that 96% of respondents did not spontaneously consider income effects. They also found that explicit instruction on income losses led to lower valuations for one of the four states in the TTO valuations, and 2 of the four states in the VAS valuations.

Meltzer et al. (1999) asked 831 US patients to value blindness and back pain through TTO and found that less tan 25% of respondents spontaneously considered income effects. They also found that explicit information on income losses led to significantly different valuations for back pain. Sendi and Brouwer (2005) asked 20 Swiss health professionals to value multiple Sclerosis through VAS, finding that 40% of respondents spontaneously considered income effects, and these considerations led to significantly lower valuations. Myers et al. (2007) asked 181 US Undergraduate students to value carpel tunnel syndrome through standard gamble and found that those with explicit information on income losses gave significantly lower valuations. Finally, Richardson et al. (2008) asked 181 patients and general population to value visual impairment through TTO and found that 38% of respondents spontaneously included income effects. They found that this led to significant differences in valuations in some cases. More information on these studies can be found in a literature review by Tilling et al. (2008).

In the UK context, while the NICE *Guide* (2004; 2008) clearly states that the costs for a reference case analysis should not include any indirect costs, it remains silent on whether or not it expects health state outcome measures to include the impact of lost earnings. This leads to a potential inconsistency, since on the one hand the scope for including the impact of lost earnings via costs is restricted, on the other hand the same may already be included as part of the health effects on individuals. Thus, a key concern for users of the EQ-5D instrument would be whether or not the published population value sets already incorporate this loss in earnings due to ill health.

Methods

Background, Ranking and VAS

Data were gathered through an online self-complete questionnaire, presented in Dutch, in the Netherlands. Invitations were sent out to a subset of potential survey respondents in order to obtain a representative sample of 300 members of the Dutch general public. The data collection was performed by an online market research

company (Survey Sampling International; www.surveysampling.com). We used the TTO format as used in the MVH protocol (Dolan, 1997) and the time horizon was 10 years. The main difference is that our survey was online rather than face to face. Respondents were presented with on-screen visual aids to make the task as easy as possible.

All respondents were asked a number of background questions: age, sex, education, marital status and occupation. In addition, number of children, net own income and net household income were included to help us understand the effect dependents and own income have upon the propensity to include income effects. Furthermore, to ensure representativeness, ethnic origins and religion were included due to the diverse nature of Dutch society.

Following the background characteristics respondents were asked to describe their own health through the EQ-5D descriptive system.

Respondents were next asked to rank four hypothetical EQ-5D health states (see below for details), full health, dead and "your own health today". They were then asked to place the same seven states on a standard EQ-5D visual analogue scale (VAS).

The TTO exercises

Following the above preliminary exercises the main part of the study consisted of a number of different TTO questions, as outlined in table 1.

Three versions of the questionnaire were used, with allocation of respondents being determined randomly. The versions differed only in terms of the levels of income loss they faced in TTO's 3 and 4. Respondents first valued the four health states through TTO1 (the states were the same as they encountered in the VAS and ranking exercises). They were then asked if they had included income considerations in these valuations. In TTO2 respondents were given instructions to either include or exclude income effects depending on their response to the follow up to TTO1. In TTO3 respondents were given information about the specific level of income loss they would incur in the health state. In TTO4 respondents valued an income loss with health remaining constant at perfect health. In TTO5 respondents valued an income gain with health remaining constant in perfect health.

Table 1: The TTO exercises

TTO			Version		
			1	2	3
1	Standard MVH TTO question	*"You can live for 10 years in health state X or a shorter period of time in full health."*	4 states [a]	4 states	4 states
2	Repeat of TTO 1 with instruction to include or exclude income effects [b]		4 states	4 states	4 states
3	Respondents explicitly told how much income they will lose in the given health state	*"You can live for 10 years in health state X or you can live for a shorter period of time in full health. In state X your ability to work will be impaired and your current income will fall by 20% [or 40% or 60%]."*	4 states, 20% income loss	4 states, 40% income loss	4 states, 60% income loss
4	Trading time to avoid an income loss with health constant in perfect health	*"You can live for 10 years with 40% [or 60% or 80%] of your current income or you can live for a shorter period of time with your current income."*	20% income loss	40% income loss	60% income loss
5	Trading time for an income gain with health constant in perfect health	*"You can live for 10 years with your current income or you can live for a shorter period of time with an increase of 20% [or 40% or 60%] of your current income."*	20% income gain	40% income gain	60% income gain

Note: [a] The four EQ-5D states valued in all versions of TTO 1-3 were: 11112, 22211, 11222, 22322. [b] Determined by follow up to TTO 1.

Each respondent had a total of 14 TTO exercises to complete. This may be considered a large amount, however this is not uncommon (e.g. both the Dutch MVH study, Lamers et al. 2006, and the Japanese MVH study, Tsuchiya *et al* 2002, asked each respondent to value 17 different states). Given a sample size of 300, we will have 300 responses per state for TTO 1 and 2, and 100 responses per questionnaire version.

It is important to include the standard TTO question (TTO1) as a baseline against which the later TTO questions could be compared. Directly following TTO1 respondents were asked a number of follow up questions. They were asked if they had considered the effect the states would have on their ability to work, on their income, on their friends and relatives and on their leisure time. They were also asked if they had considered the implication that they only had 10 years left to live. Recent research has shown that respondents do not consider this (reduced life span) which perhaps suggests that they may not fully consider the implications of the given health states (van Nooten *et al.* in press). Finally, respondents were asked if they had private insurance that would cover any income losses. The social security system is rather generous in the Netherlands so it is likely that nearly all respondents will have some form of social insurance (except any non-EU citizens) but some may have additional private insurance.

TTO2 is an ex-post inclusion/exclusion question. The "ex-post inclusion" approach was used by Sendi and Brouwer (2005), while Krol. *et al* (2006) and Krol *et al.* (2008) used the "ex-post exclusion" approach. Therefore we will be able to compare our results with these studies and also further test the effect of explicit instructions.

TTO 3 provides specific information about income losses that will be associated with the given health state. Meltzer *et al.* (1999) also provide respondents with specific information. In version 0 respondents were given no guidance, in version 1 they were told disability payments would cover 60% of their income, and in version 2 they were told that there would be no disability payments (respondents randomly allocated to one of the three versions). Unfortunately, they ask respondents to value blindness and back pain so our results will not be comparable with theirs.

TTO 4 takes a new approach by asking individuals to value negative income effects in the absence of health effects. One concern with this is possible non-responses on moral grounds; people may feel that giving up life for money is unethical.

The Health States

As mentioned, four EQ-5D health states will be valued[1]:

11112 22211 11222 22322

We chose these health states in order to have variation in the severity of the health states as well as variation in levels of impairment of the different domains. This may

[1] The EQ-5D Descriptive system has 5 dimensions and 3 levels per dimension, giving a total of 243 health states. For example, 22322 describes the following state:– Some problems with walking about, some problems with washing and dressing, unable to perform usual activities, some pain or discomfort and moderate anxiety and depression.

especially be important for the 'usual activities' dimension since it perhaps is most closely related to the ability to work.

One potential problem is that we paired all health states with all levels of income loss in TTO 3, and some respondents may consider it unrealistic for state 11112 to cause a 60% loss in income.

Hypotheses and analysis

Data were converted to utility scores by dividing by ten the number of years in health state X equivalent to 10 years in full health. Therefore if 6 years in health state X is deemed equivalent to 10 years in full health then this response will be coded as 0.6 (6/10). No protocol for states worse than dead was included as we felt this would be too complicated for a self-complete questionnaire. A zero discount rate is assumed which is common though results in a slight downward bias in results (e.g. Attema and Brouwer 2009). Data analysis was performed in Stata version 9.

Using different TTO questions will allow us to test a number of *null hypotheses*:

1) *The majority of respondents, when there is no mention of income, will not take income considerations into account.*

Among the existing studies, 6 out of 7 studies (one did not test for spontaneous inclusion) found that 40% or less of respondents spontaneously included income effects. Only one existing study found that a majority of respondents spontaneously included income effects (64%, Krol *et al.* 2009). This hypothesis will be tested simply by observing responses to the follow up question to TTO 1 – "did you consider effects the state might have upon your ability to work and hence upon your income?" Additionally, a multi-variate probit regression will be used to determine how background characteristics affect the probability that an individual will consider income effects. The binary dependent variable will be whether or not income effects were taken into account.

2) *Valuations of those that do and do not spontaneously include income effects will not differ.*

Of the six existing studies to have tested this only two have found significant differences between the two groups, buut one (Sendi and Brouwer, 2005) had a very small sample size of 20, while Richardson *et al.* (2008) asked the follow up question approximately one month after the initial TTO exercise. This hypothesis will be tested by comparing the responses to TTO 1 of those that did and did not consider income effects. The hypothesis will be tested formally through a t-test. Additionally, four standard Ordinary Least Squares (OLS) regressions will be performed. Valuations of the four health states (in TTO1) will make up the four dependent variables. The independent variables will consist of a dummy variable for whether or not income effects were spontaneously included and a number of background characteristics.

3) (a) *Those that do not spontaneously include income effects in the standard TTO question will not alter their valuations when asked to repeat the exercise considering income effects.*

(b) *Similarly, those that do spontaneously consider income effects in the standard TTO question will not alter their valuations when asked to exclude income effects.*

Krol *et al.* (2006) and Krol *et al.* (2009) both asked respondents who spontaneously considered income effects to repeat the exercise excluding these effects. The first of these studies found that valuations were revised upwards for two of the three states, while the second study found no significant differences between the two groups. Sendi and Brouwer (2005) found that those that did not consider income, when asked to repeat the exercise including these effects, amended their valuations downwards (as expected).

This hypothesis will be tested by comparing responses to TTO 1 and TTO 2. The hypothesis will be tested formally through a t-test.

4) *Whether or not respondents think the given health states will affect their income will not be affected by background characteristics.*

This will be tested through four probit models, in which the dependent variables will be whether or not respondents thought each of the four states would reduce their income, and the explanatory variables will be background characteristics. If any of the variables are significant then the null hypothesis will be rejected.

5) *The valuations of the 4 health states in TTO3 will not differ depending on the level of income loss they are paired with.*

This will be tested through unpaired t-tests. If the valuations are significantly different then the null hypothesis will be rejected. Meltzer et al. (1999) found significant differences in valuations of back pain depending on the level of disability payments respondents were told they would receive.

6) *The values of TTO 3 can be fully explained by use of a linear additive model based on values of those that did not consider income effects (either spontaneously in TTO 1 or following instruction in TTO 2) and the values from TTO 4.*

In other words, if health state x is valued at 0.8 (i.e. a 0.2 decrement) and 20% income loss is valued at 0.8 (i.e. another 0.2 decrement), then health state x with income loss of 20% should be valued at 0.6 (i.e. a 0.4 decrement). However, if the relationship between health and income is not additive then the null will be rejected.

Results

Data are available from 321 members of the Dutch general public who participated in the online survey. Preliminary data examination showed that many respondents had been unwilling to trade any life years in a number of the 14 TTO exercises. Figure 1 illustrates the number of TTO exercises in which respondents were not prepared to trade time for improved health/income. This shows that 25% of respondents were unwilling to trade any time in any of the 14 TTO exercises. For some respondents this may be a genuine representation of preferences but we suspect that many of these respondents strategically chose not to trade. Respondents were selected from a database of individuals who have signed up to complete exercises of this nature. Therefore they may have deduced that the quickest way to complete the exercise is by choosing not to trade. The sooner they complete the exercise the sooner they are awarded a given amount of money to be donated to a charity of their choice and the chance to win a prize themselves. Van Nooten et al. (in press) also found numerous respondents opted not to trade in TTO exercises in their online questionnaire.

Figure 1 - Histogram showing the number of TTO's in which respondents were unwilling to trade

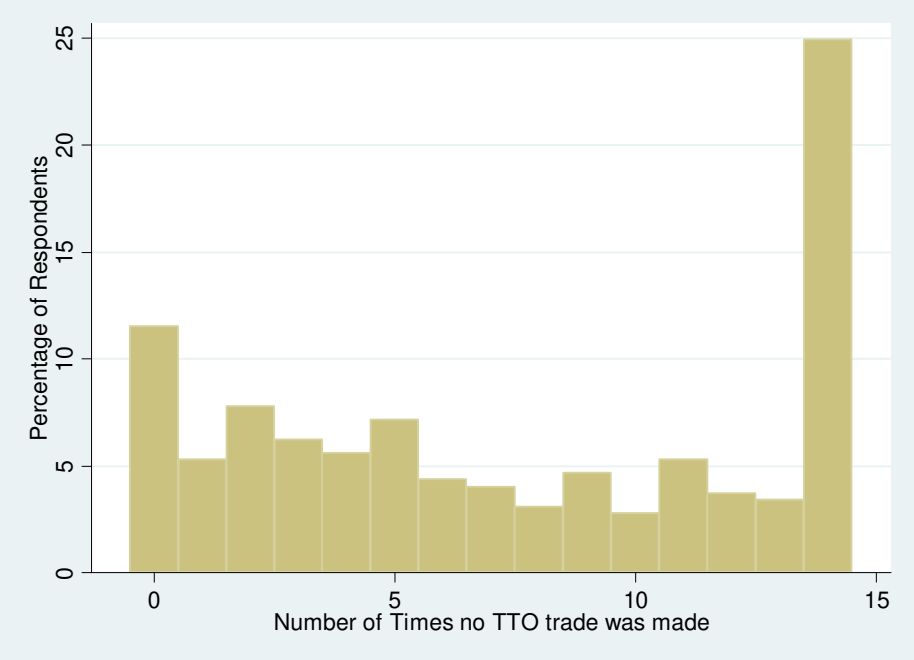

Table 2 shows the background characteristics firstly for the entire sample and then for those that have traded in at least one of the TTO's and those that have not traded at all (i.e. 'extreme' non-traders). The sample has slightly more males than females. All members of the sample were aged between 18 and 65 as we felt that people of these ages were most likely to be concerned about income. 42% of the sample were not employed and this is likely to affect the likelihood of considering income effects and the importance of these considerations. More than half of the sample had children, which is also likely to affect the likelihood of considering income effects as more people are dependent upon that income. Just under half of the sample are married and the mean VAS score for own health was 0.76. Of the entire sample 49% stated that they had spontaneously considered income effects.

Table 2 – Background Characteristics by Traders and Non-Traders

		All	Traders	Non-Traders	Chi^2 Test (p-values) Traders vs Non-Traders
Number of Respondents		321	241	80	
Gender	Male	51.0%	52.0%	54.0%	0.350
	Female	49.0%	48.0%	46.0%	
Age	Average (SD)	44(13.1)	43.19 (13.19)	46.6 (12.37)	
	18-35	29.0%	32.0%	21.0%	0.148
	36-50	32.0%	31.0%	33.0%	
	51-65	39.0%	37.0%	46.0%	
Educated beyond the minimum school leaving age	Yes	67.0%	66.0%	70.0%	0.507
	No	33.0%	34.0%	30.0%	
Educated to Degree Level	Yes	31.0%	32.0%	29.0%	0.592
	No	69.0%	68.0%	71.0%	
Employment	Employed	52.5%	53.5%	50.0%	0.874
	Self-Employed	5.5%	5.0%	7.5%	
	House Wife/Husband	13.0%	12.5%	15.0%	
	Pensioner	6.5%	7.0%	5.0%	
	Work Seeking	3.0%	3.0%	2.5%	
	Unable to Work	11.5%	10.0%	16.0%	
	Student	8.0%	9.0%	4.0%	
Net Own Monthly Income	<1000 Euros	39.0%	38.0%	41.0%	0.873
	1000 - 1499	22.0%	21.5%	24.0%	
	1500 - 1999	18.0%	19.0%	16.0%	
	>2000 Euros	21.0%	21.5%	18.0%	
Children	Yes	54.0%	49.5%	67.5%	0.005
	No	46.0%	50.5%	32.5%	
Religion	Protestant	17.0%	16.5%	19.0%	0.182
	Roman Catholic	26.5%	28.5%	20.0%	
	Atheist	49.5%	49.5%	50.0%	
	Other	7.0%	5.5%	11.0%	
Marital Status	Married	46.5%	42.5%	59.0%	0.118
	Single/Never Married	21.0%	22.5%	16.0%	
	Divorced	10.0%	12.0%	4.0%	
	Widowed	2.0%	2.0%	1.0%	
	Living Together	17.5%	18.0%	17.5%	
	Other	3.0%	3.0%	2.5%	
Mean Self-Reported Health on the EQ-VAS[2]		0.76	0.75	0.80	0.073
Spontaneously Included Income in TTO1	Yes	49.0%	42.5%	70.0%	0.000
	No	51.0%	57.5%	30.0%	

[2] Due to the exclusion of some meaningless valuations (see below text) the relevant sample sizes for this variable are: All (280), Traders (213), Non Traders (67).

Two variables were highly significantly correlated with whether or not respondents were prepared to trade in any of the TTO exercises: whether or not they had children and whether or not they spontaneously included income effects. Parents were more likely to be extreme non-traders than non-parents. This suggests that parents would rather live in a poor health state than die early and leave their children behind. Extreme non-traders were more likely to spontaneously consider income effects than traders. For the whole sample 49% spontaneously considered income effects, compared with 70% amongst the extreme non-traders. This suggests that either these non-traders do not feel the health state will affect their income, or they feel it will affect their income but this change in income does not affect their TTO valuation. The other possible explanation is that their responses are meaningless strategic non-trades. Self-reported health on the VAS was weakly correlated with whether or not respondents traded, with non-traders being in better health than traders.

The existence of more parents among the extreme non-traders does suggest that these may be meaningful preferences rather than strategic responses. However, the aim of our study is to compare changes in valuations depending upon income effects, not to generate health state valuations comparable with existing tariffs. Responses of non-traders will not help us achieve this aim, and instead may dilute the more meaningful responses of traders. We have chosen to exclude these extreme non-traders from our analysis which reduces the sample size from 321 to 241. Furthermore, 41 respondents gave negative VAS valuations of own health (13 of whom were extreme non-traders). It is very unlikely that someone in a state of health worse than dead would be able to complete an online questionnaire. Examination of these responses suggested that they were not meaningful, and were predominantly caused by very high valuations of dead. Comparison with their EQ-5D valuations showed that these respondents were generally in good health. These respondents are excluded from analysis involving VAS of own health (reducing sample size to 213), but included in all other analysis.

The top half of table 3 shows the results for the standard MVH TTO (1), firstly for the main sample (n=241) and then by their response to the follow up question of whether or not they spontaneously included income effects. Two sided t-tests directly compare the mean results of those who did and did not spontaneously include income effects. The bottom half of the table shows the results of TTO2 (ex-post/ex-ante). Respondents who spontaneously included income effects in TTO1 were instructed to exclude them. Respondents who did not spontaneously include income effects were instructed to include them. The first observation is that respondents consistently value state 22211 higher than state 11222 which suggests that they consider pain and depression to be worse than problems with mobility and self-care. We would expect the values for spontaneous inclusion to be lower than those for spontaneous exclusion (1 vs 2), however this is only the case for one of the four states, and in this case the t-test is insignificant. The t-test suggests that the differences in valuations are only weakly significant for the most severe state (22322), and in this case spontaneous exclusion gives a lower result which is contrary to expectations.

Table 3 – TTO Results from TTO1 and TTO2 both including and excluding income effects

		All (n=241)			(1) Spontaneously Included Income (n=102)			(2) Spontaneously Excluded Income (n=139)			T-test p-values. Including vs Excluding (1 vs 2)	T-test p-values Ex-Post Instruction (1vs 3)
	Health State	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**		
TTO 1 (MVH)	**11112**	0.92	1.00	0.18	0.93	1.00	0.16	0.91	1.00	0.19	0.270	0.056
	22211	0.86	0.97	0.21	0.85	0.95	0.22	0.86	0.98	0.21	0.698	0.029
	11222	0.82	0.90	0.22	0.84	0.90	0.21	0.81	0.90	0.23	0.289	0.598
	22322	0.68	0.73	0.28	0.72	0.80	0.26	0.65	0.70	0.29	0.051	0.618
		All (n=241)			**(3) Instructed to Exclude Income Effects (n=102)**			**(4) Explicitly Instructed to Include Income Effects (n=139)**			**T-test p-values. Including vs Excluding (3 vs 4)**	**T-test p-values Ex-Ante Instruction (2 vs 4)**
	Health State	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**		
TTO 2 (Ex-ante/ Ex-Post)	**11112**	0.92	1.00	0.17	0.95	1.00	0.12	0.90	1.00	0.20	0.007	0.242
	22211	0.85	0.91	0.21	0.89	0.94	0.17	0.83	0.90	0.23	0.022	0.004
	11222	0.81	0.90	0.22	0.85	0.90	0.17	0.78	0.87	0.24	0.011	0.037
	22322	0.67	0.70	0.28	0.73	0.80	0.25	0.63	0.66	0.29	0.008	0.34

Table 4 – OLS and Probit Regressions to show the effect of background characteristics on valuations in TTO1 and on the propensity to spontaneously include income effects (n=213)

	OLS				PROBIT
Variable	**11112 R^2=0.069**	**22211 R^2=0.107**	**11222 R^2=0.058**	**22322 R^2=0.097**	**Included Income Pseudo R^2= 0.057**
Intercept	0.868***	0.635***	0.640***	0.515***	
Included Income	0.018	-0.019	0.021	0.053	
Income>999euros per month	0.016	0.006	-0.013	-0.010	0.090
Gender (Male=1, Female=0)	-0.007	-0.003	-0.018	-0.005	0.034
Age	0.001	0.003*	0.003**	0.003*	0.002
Married=1, Other=0	-0.033	-0.040	-0.042	-0.093**	-0.064
Educated Beyond Minimum School Leaving Age	0.021	0.016	0.026	0.035	-0.052
Have a Degree	-0.015	0.012	-0.043	-0.062	-0.112
Working=1, Not Working=0	0.040	-0.002	-0.007	-0.005	0.021
Have Income Insurance	0.013	-0.001	-0.002	-0.009	0.203***
Have Children	0.031	0.088***	0.029	0.112**	-0.010
VAS Own Health	-0.005	0.115*	0.101	0.011	-0.248
Values presented are coefficients. Significance is shown as follows: * 10%, **5%, ***1%					

When those who spontaneously included income effects were asked to exclude these effects (1 vs 3) the valuations of all four health states went up at the aggregate level. These changes are only significant for the first two states, but the statistical significance is weak and the magnitude of the change is small. When those who did not spontaneously include income effects were instructed to exclude these effects (2 vs 4) the valuations of all four health states went down at the aggregate level. These changes were statistically significant for states 11222 and 22211. As expected the largest differences in valuations are between those that are explicitly instructed to include income effects and those that are explicitly instructed to exclude these effects in TTO 2 (3 vs 4).The valuations of all four health states are lower when respondents are instructed to include income effects. These differences are significant at the 5% level for two of the states and at the 1% level for the other two.

Table 4 shows the results of multivariate regression analysis. In the four columns of OLS results the dependent variables are the valuations of the four health states through the standard MVH TTO (1). The explanatory variables are background characteristics and whether or not respondents spontaneously included income effects. The results suggest that having children significantly increases valuations for two of the four states. Age has a weakly significant positive effect on valuations for states 22211 and 22322, and a more significant positive effect for state 11222. Being married leads to significant lower valuations for the worst state. Whether or not respondents spontaneously included income effects did not significantly affect valuations which supports the findings in table 2.

The final column of table 4 shows the results of a probit model in which the dependent variable is whether or not respondents spontaneously included income effects and the explanatory variables are once again background characteristics. Those with income insurance are more likely to spontaneously include income effects (significant at 1% level).

Table 5 shows how background characteristics affect the likelihood someone will think the health states will reduce their income. For the 4 states the percentage of respondents who thought their income would fall was 13%, 42.5%, 39% and 53.5% respectively. It is interesting to note that although state 11222 is valued lower than state 22211, more people think 22211 will affect their income. People obviously perceive moderate problems with mobility and self-care more likely to affect one's ability to work than pain, discomfort and anxiety and depression. For all but the mildest state Age has a significant negative impact on the likelihood of thinking the states will reduce income. For all but the mildest state, being in employment highly significantly increases the likelihood of thinking a state will reduce income. This is unsurprising given that the incomes of those not in work will not be affected if ill health hinders their ability to work. Having income insurance highly significantly reduces the likelihood of thinking the worst health state will reduce income, and weakly reduces the likelihood for state 11222.

Table 5 – Probit regression showing the effect background characteristics have on the likelihood of thinking a given health state will reduce income (n=213)

	PROBIT (Dependent Variable - Likelihood of thinking given health state will affect Income)			
Variable	**11112 Pseudo R2=0.050**	**22211 Pseudo R2=0.120**	**11222 Pseudo R2=0.136**	**22322 Pseudo R2= 0.230**
Income>999euros per month	-0.007	-0.007	0.062	0.111
Gender (Male=1, Female=0)	0.016	-0.054	-0.117	-0.123
Age	0.001	-0.009**	-0.011***	-0.010**
Married=1, Other=0	0.022*	0.011	0.002	-0.018
Educated Beyond Minimum School Leaving Age	-0.018	-0.018	0.058	0.076
Have a Degree	0.069	-0.016	0.041	0.006
Working=1, Not Working=0	-0.020	0.336***	0.217***	0.439***
Have Income Insurance	0.009	-0.098	-0.136*	-0.254***
Have Children	-0.046	-0.089	-0.070	-0.056
VAS Own Health	-0.041	-0.170	0.046	0.182
Values presented are coefficients. Significance is shown as follows: * 10%, **5%, ***1%				

Table 6 – Valuations of the four health states combined with the three different levels of income loss (TTO3)

	20% Income Loss (n=78)			T-test p-values: 20% vs 40%	40% Income Loss (n=80)			T-test p-values: 40% vs 60%	60% Income Loss (n=83)		
Health State	Mean	Median	SD		Mean	Median	SD		Mean	Median	SD
11112	0.89	1.00	0.19	0.052	0.81	0.98	0.29	0.529	0.78	0.90	0.27
22211	0.82	0.90	0.21	0.283	0.78	0.89	0.28	0.068	0.70	0.70	0.27
11222	0.77	0.82	0.23	0.469	0.74	0.83	0.29	0.366	0.70	0.75	0.29
22322	0.67	0.70	0.27	0.330	0.63	0.60	0.30	0.722	0.61	0.60	0.30

Table 6 shows the valuations of the four health states combined with the three different levels of income loss – 20%, 40% and 60% - that were given to respondents depending on which version of the questionnaire they received. The valuations of the four health states in all three versions of the questionnaire go from best to worst in the following order: 11112, 22211, 11222, and 22322 (this is the same ordering as in TTO1). This holds in all cases except one: for 60% income loss state 11222 is valued

higher than state 22211. The same states across versions are valued lower as the amount of income loss increases. This holds in all cases. However, the differences between the valuations for different levels of income loss are only significant in one case: 22211 with 40% loss vs 22211 with 60% income loss. The lack of significance in these tests appears to be due to the small sample sizes.

Table 7 firstly shows mean TTO valuations without income considerations. This was either spontaneously in TTO1 or following explicit instruction in TTO2. The table also shows mean TTO valuations of just income loss (TTO4). The values generated when these combinations of states and income levels were valued simultaneously, through TTO3 (see table 5) are presented as the actual values. We have also presented hypothetical values representing what the outcomes of the different combinations would be firstly assuming a model with no interactions (i.e. additive), and secondly assuming some degree of interaction (as specified using a multiplicative formulation). These hypothetical values were generated at the individual level. The additive values were generated by adding the disutilities of the two valuations together and then subtracting from 1 e.g. if the health state was valued at 0.8, and the income loss was valued at 0.8, then the additive value would be given by: 1-[(1-0.8) +(1-0.8)] = 0.6. The multiplicative value was simply generated by multiplying the two values together e.g. 0.8*0.8=0.64. Further work will explore other specifications for interactions (e.g. multilinear). Paired t-tests were performed to compare the hypothetical additive and multiplicative values with the actual values. Significance in these t-tests suggests that the given relationship (additive or multiplicative) is unlikely to represent the actual relationship between health and income.

We attempted to estimate the number of respondents that could be approximated (crudely) as additive or multiplicative for each combination of health and income. This was done by taking an average of each individuals' hypothetical additive and multiplicative values and then determining whether their actual value was higher or lower than this average. If it was higher we deemed them to fall approximately into the multiplicative category and if it was lower we deemed them to fall into the additive category.

The first observation from table 7 is that for all combinations the actual value is higher than both the additive and multiplicative values. The t-tests in table 6 comparing the additive and actual values show that there are at least weakly significant differences between the two in 11 out of the 12 combinations. This would suggest that the relationship between income and health is unlikely to be purely additive. The t-tests between the multiplicative and actual values are at least weakly significant in 3 of the 12 combinations. This suggests that the relationship between health and income is closer to multiplicative than additive. In reality the relationship between health and income may be approximated by a multiplicative function plus a constant.

Table 6 – Comparisons of actual values through TTO3 (health state with explicit level of income loss) with hypothetical Additive and Multiplicative values generated through combining valuations without income with valuations of just income loss

Health State	Mean TTO valuation without income considerations[a]		Mean Income Loss Value for 20% (n=78)	Number of Respondents	Mean Income Loss Value for 40% (n=80)	Number of Respondents	Mean Income Loss Value for 60% (n=83)	Number of Respondents
		Mean TTO value for income loss only (TTO4)	**0.901**		**0.819**		**0.755**	
11112	**0.928**	Additive:	0.834*	55	0.762*	49	0.663***	43
		Multiplicative:	0.855	23	0.778	31	0.710	40
		Actual:	0.888		0.812		0.784	
22211	**0.874**	Additive:	0.772	53	0.701**	42	0.626**	44
		Multiplicative:	0.800	25	0.727*	38	0.678	39
		Actual:	0.817		0.775		0.695	
11222	**0.827**	Additive:	0.719*	44	0.674**	45	0.563***	38
		Multiplicative:	0.751	34	0.705	35	0.636**	45
		Actual:	0.774		0.744		0.703	
22322	**0.682**	Additive:	0.579***	32	0.507***	37	0.434***	37
		Multiplicative:	0.625*	46	0.567	43	0.546	46
		Actual:	0.674		0.629		0.613	

Paired t-tests were performed to compare the additive and multiplicative values with the actual values for each combination of health state and income loss.
The significance of these tests is shown as follows: * 10%, ** 5%, ***1%.

a: taken from TTO1 of those who did not include income spontaneously, and TTO2 of those who did (n=241)

Discussion and Conclusions

Our results show that (for the whole sample) 49% of respondents claimed to spontaneously include income effects. This is lower than one of the two studies using TTO valuation of EQ-5D health states (Krol et al. 2008), which produced a value of 64%, but higher than the other (Davidson and Levin, 2008), which found that 6% of respondents spontaneously included income effects. It is possible that respondents may have considered these effects for some states but not others. However, we could only ask respondents whether they had taken income effects into account after valuing all 4 health states in order to avoid contaminating the exercise.

The findings support those of all three existing studies valuing EQ-5D states (Krol et al 2006, 2008, Brouwer et al. in press): that spontaneous inclusion of income effects does not significantly affect health state valuations at the aggregate level. This suggests that previous studies using either the human capital or friction cost methods to value productivity costs in the numerator of the C/E ratio have not double counted these costs. Similarly, from the current NICE perspective, the results suggest that economic evaluations not explicitly including productivity costs have not done so implicitly through the health state valuation exercise either.

The results do contradict the findings of Krol et al. (2009), but support the findings of Krol et al. (2006) by finding that explicit instruction does lead to statistically significant differences in valuations in some cases, particularly when comparing results from explicit inclusion and explicit exclusion. It is worth noting that we are not able to confirm or dispute the finding of these studies with regards to ex-ante instructions (that they do not statistically significantly affect valuations). In light of the fact that spontaneous inclusion/exclusion seems to be insignificant the role of explicit instruction may be redundant. If there is a desire to include productivity costs in the numerator explicitly instructing respondents to exclude income effects may bias valuations downwards (imagine telling someone not to think about a pink elephant). If future research shows that explicit inclusion indeed changes valuations, this may potentially offer a way to include productivity costs (partly) through the denominator. Nevertheless, there are strong arguments that incorporating productivity costs through the numerator represents the more accurate and certain option (Brouwer *et al.* 1997a,b, Brouwer *et al.* 2005, Meltzer *et al.* 1999).

The results suggest that older members of the sample were significantly less likely to think a given state would reduce their income. This cannot be explained by retirement as only 7% of the sample are retired. Employed people are more likely to think a given health state will reduce income. Therefore, given that only 52.5% of our sample were employed, we can not rule out the possibility that spontaneous inclusion of income effects may have caused significant differences in valuations if our sample had contained a greater number of employed persons. It also suggests that previous studies using student samples (Myers *et al.* 2007, Davidson and Levin 2008) may be flawed.

The results attempting to explore the relationship between health and income, when valued separately and simultaneously, are interesting. The consistency of the results across the 12 different combinations of health and income suggest that the creation of an interaction term between health and income is entirely possible. Whether this

could lead to a method to include income effects through general population valuation rather than through monetrary calculation, remains questionable. Explicit instruction may lead to adjusted valuations but this is shrouded in uncertainty. An important, and thus far unmentioned point is that income effects are a poor proxy for productivity costs. Income insurance may reduce the loss to the individual valuing the given health state, but it does not reduce the loss to society. There is a growing pressure on NICE to incorporate wider societal effects, most notably productivity costs. If they are to do so, inclusion in the numerator of the cost-effectiveness ratio may represent the most credible option. Explicit inclusion in the denominator by capturing productivity costs in the health outcome measure causes numerous problems and offers no noticeable benefits. Without explicit instructions, the effects of income considerations in health state valuations appear to be negligible.

Some weaknesses of this study need to be noted. The use of an online self-complete survey may not be appropriate for a large number of different TTO's, as suggested by the number of non-traders. This study needs to be replicated using an interview method of administration (as used to generate commonly used value sets), which would allow continual guidance and explanation and also enable qualitative feedback to be gathered, which may enable researchers to further understand the thought processes of respondents. Furthermore, no research in this area has been carried out in the U.K. Factors such as different social security systems can lead to significantly different results between countries. Research is needed in the U.K. to see if these results hold.

The power of this study is weak. Assuming standard deviations in TTO valuations of 0.16 (the lowest SD in table 2) and alpha of 0.05, we can detect a difference of 0.1 with power 0.998. However, assuming standard deviations of 0.29 (the highest SD in table 2) we can only detect a difference of 0.1 with power 0.753. Future studies need to be appropriately powered which may be difficult if the interview method of administration is used.

This study did not have a protocol for states worse than dead. We felt that since respondents completed the tasks independently and without guidance, it may become too complicated and time consuming to include a protocol for states worse than dead. Given that the worst health state (22322) has a value on the Dutch tariff of 0.092 (Lamers *et al.*2006) we were concerned that a significant proportion of respondents may value this state as worse than dead. In fact, in TTO3 with the highest income loss level of 60% (which should elicit the lowest values) only 7, 4, 5 and 7 responses were zero for the four health states respectively. However, if this study was to be repeated in the UK using interview method of administration and the same four health states it may be worth including a protocol for states worse than dead.

We plan to do further analysis using this data. Panel regression analysis can be used to include valuations of all four health states in the regressions. This would obviously increase the sample sizes in the regressions. We have not used the ranking and VAS results. These could be compared with the TTO results as an internal consistency test. Furthermore it would be useful to see if TTO extreme non-traders also gave states similar values in the VAS and Ranking exercises. As mentioned, we plan to analyse the income gain and income loss responses to see if there are any systematic differences between the two. Additionally, just as there was a follow up to TTO1

asking if respondents had considered income effects, there was also a question asking if they had considered leisure. While one may argue that the QALY without leisure becomes a hollow concept evidence has shown that not all respondents include it and inclusion can lead to different valuations (Sendi and Brouwer, 2005). There were additional questions on income, most notably partner's income. It would be useful to link the responses to these questions to factors such as whether they thought the states would reduce their income, whether they spontaneously included income effects and whether this changed their valuation. Finally, a feedback question asked if respondents found the scenarios hard to imagine. This may offer a further explanation for non-traders.

References

Attema AE, Brouwer WBF. The correction fo utility scores for utility curvature using a risk free utility elicitation method. *Journal of Health Economics* 2009; **28** (1): 234-243

Brouwer WBF, Koopmanschap MA, Rutten, FHH. Productivity Costs Measurement through quality of life? A response to the recommendation of the Washington Panel. *Health Economics* 1997a; **6**: 253-259.

Brouwer WBF, Koopmanschap MA, Rutten, FHH. Productivity Costs in cost-effectiveness analysis: numerator or denominator: a further discussion. *Health Economics* 1997 b; 6: 511-514

Brouwer WBF, Meerding WJ, Lamers L, Severens H. Productivity and health related quality of life – an exploration. *PharmacoEconomics* 2005; **23:** 209-18

Brouwer WBF, Grootenboer S, Sendi P. The incorporation of income and leisure in health state valuations when the measure is silent: an empirical inquiry into the sound of silence. Medical Decision Making, in press.

Davidson T, Levin LA. Do individuals consider expected income when valuing health states? *International Journal of Technology Assessment in Health Care* 2008; 24(4): 488-494

Dolan P. Modelling Valuations for EuroQol Health States. *Medical Care* 1997; 35(11): 1095 - 1108

Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. Oxford: Oxford University Press, 1996.

House of Commons Health Select Committee. (2007). *First Report of Sessions 2007-8 on the National Institute for Health and Clinical Excellence*.
http://www.publications.parliament.uk/pa/cm200708/cmselect/
cmhealth/27/2702.htm [accessed 26 May 2008]

Koopmanschap MA, van Ineveld B. Towards a new approach for estimating indirect costs of disease. *Social Science and Medicine* 1992; **34**: 1005-10

Koopmanschap M, Rutten FHH. Indirect costs in economic studies. *Pharmacoeconomics*1993; **4**: 446-54

Koopmanschap MA, Rutten FHH, van Ineveld B, Van Roijen L. The friction cost method of measuring the indirect costs of disease. *Journal of Health Economics* 1995; **14**: 171-89

Krol M, Brouwer WBF, Sendi P. Productivity costs in health-state valuations: does explicit instruction matter? *Pharmacoeconomics* 2006; **24:** 401-414

Krol M, Sendi P, Brouwer WBF. Breaking the silence: Exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises. *Value in health* 2009; **12** (10): 172-180..

Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Busschbach JJV. The Dutch Tariff: Results and arguments for an effective design for national EQ-5D valuations studies. *Health Economics* 2006; **15**(10): 1121-1132.

Meltzer D, Johannesson M. Inconsistencies in the "Societal Perspective" on Costs of the Panel on Cost-Effectiveness in Health and Medicine. *Medical Decision Making* 1999; **19**: 371-377

Meltzer D, Weckerle CE, Chang LM. Do people consider financial effects in answering quality of life questions? *Medical Decision Making* 1999; **19**: 517

Myers J, McCabe S, Gohmann S. Quality-of-Life Assessment When There Is a Loss of Income. *Medical Decision Making* 2007; **27:** 27-33

NICE (2004) *Guide to the Methods of Technology Appraisal* (reference N0515)

NICE (2008) *Guide to the Methods of Technology Appraisal* (reference N1618)

Rice D, Cooper, B. The economic value of human life. *American Journal of Public Health* 1967; **57**: 1954-66

Richardson J, Peacock S, Lezzi A. Do quality-adjusted life years take account of lost income? Evidence from an Australian survey. *European Journal of Health Economics,* Forthcoming 2008.

Sendi P, Brouwer WBF. Is silence golden? A test of the incorporation of the effects of ill-health on income and leisure in health state valuations. *Health Economics* 2005; **14:** 643-647.

Tilling C, Krol M, Tsuchiya A, Brazier J, Brouwer WBF. *In or Out? Income Losses in Health State Valuation: A Review*. Manuscript, 2008.

Van Nooten FE, Koolman X, Brouwer WBF. Thirty Down, Ten to Go? Health Economics, in press.

Weisbrod BA. The valuation of human capital. *Journal of Political Economy*1961; **69** (5): 425-436