

# Values for health states with different life durations

Luciana Scalone · Peep F. M. Stalmeier ·  
Silvano Milani · Paul F. M. Krabbe

Received: 15 May 2013 / Accepted: 10 September 2014 / Published online: 30 September 2014  
© Springer-Verlag Berlin Heidelberg 2014

## Abstract

**Background** Recent research suggests that the value of health states may depend on their duration, implying a more complex relationship between quality and quantity of life than the standard linear relationship assumed in the quality-adjusted life years (QALY) model.

**Objective** This study models how duration affects the value of health states, using discrete choice data.

**Methods** A study was conducted comprising health profiles including various durations. Health states were described with the EQ-5D. Duration was introduced as a sixth domain with six levels (1, 5, 10, 15, 30, 50 years). Sixty pairs of health profiles were selected, applying a Bayesian approach. Two hundred eight university students self-completed the computerized response tasks. Data were analyzed with a conditional logistic regression model.

**Results** There is clear evidence for non-linear values for duration, that is, preferences for duration are more accurately described by a logarithmic function.

**Conclusions** In discrete choice experiments using health states with different durations, these data present the first evidence for decreasing marginal utility for duration. Our results suggest that refinement of the standard QALY framework may be needed for states with a long duration.

**Keywords** Values · Health states · Duration · Quality-adjusted life years · Discrete choice analysis · Time preferences

## Introduction

HRQoL values for health states can be derived indirectly with scaling models from judgmental tasks such as paired comparisons, discrete choices, and ranking. Judgments can also be elicited directly with specific valuation techniques (e.g., time trade-off, visual analogue scale). Frequently, the values for health states obtained by one of the measurement approaches mentioned above are multiplied by a certain amount of time (years) duration, to compute quality-adjusted life years (QALYs) [1]. Basically, the QALY is obtained by adjusting the amount of years lived in a health state by the value associated with this health state. The standard QALY model assumes that health-state values are independent of the duration of the corresponding health state [2, 3]. This assumption implies that in the QALY model, for health states valued as better than dead, such states with a longer duration would always be preferred to the same states with shorter durations. Earlier research indicated that this basic assumption may not hold [4–7], suggesting that the value assigned to health states can depend on their duration.

---

L. Scalone (✉)

Department of Clinical Medicine and Prevention, Faculty of Medicine, Research Centre on Public Health, University of Milano-Bicocca, Villa Serena, Via Pergolesi 33, 20052 Monza, Italy  
e-mail: luciana.scalone@unimib.it

P. F. M. Stalmeier

Department for Health Evidence, Radboud University Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

S. Milani

Department of Occupational Medicine, Institute of Medical Statistics and Biometrics “GA Maccacaro”, Faculty of Medicine, University of Milano, Cascina Rosa, Via Vanzetti 5, 20133 Milano, Italy

P. F. M. Krabbe

Department of Epidemiology, University Medical Center Groningen, University of Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands

The present study is a first step in the development of models for discrete choice experiments, using health profiles that vary simultaneously in health and duration. Our aim is to model the value of health states at various durations of life. Unlike the standard QALY model, the models proposed in this paper estimate how health and duration are valued for durations longer than the standard 10 years. We discuss the possible consequences of our findings for the standard QALY model and options for future health-valuation studies.

## Methods

### Measurement framework

We adopted the technique of discrete choice (DC) analysis. With this technique respondents are presented with scenarios, in which two or more alternative options, here two health profiles with varying duration, are described. In each scenario, the respondents are assumed to choose the option providing the highest benefit or value. From the responses, it is possible to determine the importance of each attribute relative to the other attributes included in the scenario, and the willingness to give up the benefits of one attribute to gain the benefits of another attribute. Also, it is possible to assess the value for the different health profiles, as a function of their attribute levels.

Over the past 15 years, DC analysis has been used to elicit preferences for healthcare interventions and to value health benefits and patient experience factors [8]. More recently, DC modeling has also been considered as a technique for health-state valuation [9–16].

### Construction of health profiles

We constructed health states based on the EQ-5D health-state system. The EQ-5D descriptive system comprises five domains (mobility, self-care, usual activities, pain/discomfort, anxiety/depression), each one with three possible levels (no problems, some problems, severe problems). Overall, 243 ( $3^5$ ) different health states can be described with this system, with 11111 representing the best and 33333 the worst health state [17].

Duration was included as a sixth attribute in the profiles (Fig. 1). To achieve sufficient differentiation for duration and to cover a wide range, six different levels were chosen: 1, 5, 10, 15, 30, and 50 years followed by death. In particular, “1 year” was chosen because it was recognized as the minimum duration feasible and comprehensible in this study; “10 years” is used in conventional time trade-off (TTO) valuations from which QALYs are calculated;

“50 years” is about the maximum life expectancy of people aged 20–25 years, who participated in this study. To obtain intermediate estimates and to make comparisons between these durations easier for the respondents, we added the attribute levels of 5, 10, and 15 years.

A fractional factorial experiment was designed to obtain a total of 60 different pairs of health profiles (Table 1). This design allowed the estimation of models with both main effects and several interaction terms. The 60 pairs were selected using a Bayesian efficient discrete choice experimental design applied to the EQ-5D descriptive system [19].

To arrive at an acceptable cognitive burden, each participant received a subset of nine pairs. The nine pairs were randomly selected from the 60 pairs using a bounded randomization procedure to ensure that all pairs were valued a similar number of times, allowing a maximum difference of three valuations per pair.

This study was conducted as part of a larger project within the EuroQoL Group, in which the main objective was to explore the application of DC analysis to derive values for health states described with the EQ-5D instrument, and to compare these with values obtained using the TTO approach [14].

The Bayesian approach was adopted in the main study to improve the efficiency of the experimental design, allowing for more precise estimates. In particular we applied an iterative procedure (nested Monte Carlo simulation) with a computer algorithm [18]. Using this procedure, 2,000 possible fractional experimental designs were randomly selected from the full design. These 2,000 designs were compared on their D-errors, which were computed on the basis of expected values of the main effects model, and the most efficient design was selected. More details about the algorithm and procedure used can be found in the paper by Stolk et al. [14]. As priors for the main effects estimates, we included the weighted average of estimates obtained from previous studies using the TTO technique [9–16].

Because no prior information for duration was available, levels for duration were included manually in the final design: levels for duration were balanced among the pairs (level balance), and were always different between the health profiles described within each scenario, following the minimum overlap criterion [22] (Table 1).

### Subjects

A convenience sample of university students from the Erasmus University in Rotterdam, the Netherlands, was recruited. Practical reasons justified the choice of involving university students, including it being realistic for respondents to consider scenarios that lasted as long as 50 years.

Imagine that you could be in one health state, A or B, for the period of time below specified. During that time your health state would remain unchanged, then you will die.

**(a)**

- No problems in walking about
- No problems to wash and dress my-self
- Some problems to perform my usual activities
- Having some pain or discomfort
- Not anxious or depressed
  
- Living in this state 5 years before dying

**(b)**

- Some problems in walking about
- Unable to wash and dress my-self
- Unable to perform my usual activities
- Having some pain or discomfort
- Moderately anxious or depressed
  
- Living in this state 30 years before dying

Which condition do you consider better, A or B?

**Fig. 1** Example of a scenario containing two health profiles

**Table 1** Final set of 60 pairs of EQ-5D health states integrated with life durations

Pair	Option 1		Option 2		Pair	Option 1		Option 2	
	Health state	Duration (years)	Health state	Duration (years)		Health state	Duration (years)	Health state	Duration (years)
1	21231	5	22323	10	31	13211	10	21233	30
2	23223	30	31113	5	32	33311	30	22133	1
3	11112	1	12221	5	33	32112	5	23312	15
4	33322	30	23312	10	34	21112	15	22111	10
5	22331	1	23233	30	35	32211	15	13333	50
6	32133	30	22312	15	36	13131	50	13113	5
7	33123	5	22233	50	37	22313	1	23231	10
8	23212	50	32121	1	38	31313	50	32231	10
9	32322	5	33131	15	39	12123	15	33321	30
10	11231	50	32111	30	40	22311	15	32123	30
11	33222	15	11312	50	41	11133	1	21123	15
12	13122	15	21212	10	42	31311	50	21313	1
13	22221	10	13212	50	43	21212	10	32213	30
14	22312	10	11212	1	44	11121	5	22112	1
15	22132	1	12321	15	45	13313	30	31221	50
16	12332	10	31333	15	46	21321	1	12111	15
17	22333	10	33332	1	47	33323	50	23122	5
18	31222	5	12112	10	48	11223	15	32321	50
19	31131	50	13111	5	49	23313	15	32222	10
20	12233	15	13132	50	50	31323	1	22321	5
21	31131	30	12121	50	51	33113	30	32332	5
22	33131	5	21323	30	52	22131	1	21212	50
23	33122	50	31132	10	53	23222	30	31113	5
24	11133	1	32211	5	54	12222	15	33121	30
25	12231	50	21121	1	55	31132	30	21333	1
26	12312	1	13131	30	56	12213	10	31232	50
27	21111	30	11311	10	57	23312	10	13123	5
28	11223	1	12313	15	58	21211	10	32313	30
29	13231	15	31231	5	59	31133	10	21331	1
30	31123	30	12212	50	60	13321	30	13231	50

## Procedure

Students were gathered into a group, given online exercises to complete, and assisted by a researcher if needed. For their participation in the study the students were offered €20. No ethical approval was required for this study.

## Modeling

The results of three models are presented. The first model mimics the QALY model. The second model has many more parameters, to arrive at a flexible description of the preferences for duration. The third model is similar to the first model but uses a logarithmic function of duration to describe non-linear preferences for duration. Conditional logistic regression models were applied to analyze the DC data [23], using STATA software (v.11.0, routine clogit). These models belong to the group of probabilistic choice models and are embedded in random utility theory [24]. From these models, the value assigned to an option and its attribute level can be estimated with a linear, additive function of the attribute levels included in the health profiles.

We started to analyze the data with model 1, which is intended to reflect the standard QALY model as closely as possible:

$$\begin{aligned}
 V = & \beta_0 + \beta_1 X_{MO2} + \beta_2 X_{MO3} + \beta_3 X_{SC2} + \beta_4 X_{SC3} + \beta_5 X_{UA2} \\
 & + \beta_6 X_{UA3} + \beta_7 X_{PD2} + \beta_8 X_{PD3} + \beta_9 X_{AD2} + \beta_{10} X_{AD3} \\
 & + \beta_{11} X_{years} + \beta_{12} (X_{MO2} \times X_{years}) + \beta_{13} (X_{MO3} \times X_{years}) \\
 & + \beta_{14} (X_{SC2} \times X_{years}) + \beta_{15} (X_{SC3} \times X_{years}) \\
 & + \beta_{16} (X_{UA2} \times X_{years}) + \beta_{17} (X_{UA3} \times X_{years}) \\
 & + \beta_{18} (X_{PD2} \times X_{years}) + \beta_{19} (X_{PD3} \times X_{years}) \\
 & + \beta_{20} (X_{AD2} \times X_{years}) + \beta_{21} (X_{AD3} \times X_{years})
 \end{aligned} \tag{1}$$

In model 1,  $V$  represents the value assigned to each health profile. The  $X$  terms represent the different attribute levels. The EQ-5D attribute levels are included as dummy variables. Each EQ-5D domain has three levels (no problems, some problems, and severe problems), therefore two dummy variables are included per EQ-5D domain, using “no problems” as a reference. Hence, for each EQ-5D domain, we estimated the value assigned to having some problems (labeled as MO2 for mobility, SC2 for self-care, UA2 for usual activities, PD2 for pain/discomfort, AD2 for anxiety/depression) and to having severe problems (labeled as MO3 for mobility, SC3 for self-care, UA3 for usual activities, PD3 for pain/discomfort, AD3 for anxiety/depression), relative to having no problem. Duration is

included in the model as a continuous linear variable to estimate the mean value assigned to each year of duration ranging from 1 to 50 years. As regards the regression coefficients ( $\beta$ ), the constant  $\beta_0$  was initially included to check for the presence of systematic effects on choices, such as a tendency to always choose the same left health profile [25];  $\beta_{1-11}$  represent the weights estimated for the main effects of the attribute levels, while  $\beta_{12-21}$  represent the effects of the interactions between the attribute levels and duration.

Model 1 is a multiplicative model since it includes first-order interaction terms between each EQ-5D attribute level and duration. Though we would have preferred to leave out the main effects in model 1, in order to mimic the multiplicative QALY model as closely as possible, this was not possible for statistical reasons. Main effects must remain in a multiplicative model in order to avoid biases in the estimates of the main effect and interaction coefficients.

To understand how duration is modeled for different health states, we created a second model in which duration was included as a categorical variable using the six levels presented to the respondents. In order to obtain a model able to estimate interactions between every EQ-5D attributes and duration, the duration levels were collapsed from 6 to 4, and then combined with the EQ-5D attribute levels (model 2). Finally, we estimated models in which several continuous non-linear functions of duration were tested, such as square root or logarithmic functions. The most efficient model (model 3) is presented. The statistical efficiency of the models was investigated using (1) goodness of fit tests, measuring the correspondence of the model with the observations, and (2) the parsimony criterion, which focuses on the minimum number of variables necessary to obtain reliable estimates. In particular, we calculated the pseudo (McFadden)  $R^2$  adjusted for the number of parameters included in the model, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC).

The Wald test was applied to test whether the  $\beta$ s were significantly different from 0 at a two-tailed  $p$  value  $<0.05$ .

We checked how many respondents showed ‘lexicographic’ preferences; i.e., whether they always chose the option with the longer time duration. Additional analyses were done for the models while excluding such respondents.

From the estimates obtained from the selected models, trends were calculated (shown as curves) for a sample of 15 possible EQ-5D health states: best health (11111), 13 mild, moderate and severe states (11113, 11131, 11133, 11312, 13311, 31311, 12333, 23232, 32211, 32223, 32313, 31333, 33323), and the worst state (33333).

To scale the values, the health profile (11111, 1 year) was chosen as the reference profile. In other words,  $V$  equals 0 for the health profile (11111, 1 year) because the reference categories are one for the EQ-5D domain levels and 1 year for duration.  $V < 0$  is therefore interpreted as less preferred than (11111, 1 year), and  $V > 0$  is interpreted as more preferred than living (11111, 1 year). Thus, values obtained with the proposed models do not refer to values anchored between 0, for dead, and 1, for full health.

## Results

### Subjects

Between June and July 2008, choice data were obtained from 208 students, out of the 209 enrolled in the study. Data from one respondent were erroneously not saved. Accordingly, a total of 1,872 observations (choices) were obtained. The participants were aged from 17 to 39 years (mean  $\pm$  SD = 22.7  $\pm$  3.5), 30.3 % were males. Twelve percent of the participants were Bachelor, while 88 % were Master students. Each pair of health profiles was answered by a minimum of 21 to a maximum of 40 respondents.

### Model selection

Sixteen respondents (7.7 %) showed lexicographic preferences for the health profiles, always preferring the longer time durations. The remaining respondents chose a health profile with a shorter time duration at least once. Excluding the respondents with dominant preferences from the analyses did not alter the estimated parameters. Therefore, the results obtained from the full study sample are shown and discussed.

In the models analyzed, the constant  $\beta_0$  was not statistically significant, meaning that there was no systematic preference for the right or left health profile. Therefore,  $\beta_0$  was excluded from the models below.

As regards the model in which duration is included with categorical dummy variables, model 2 (Table 2), some duration levels were collapsed to obtain 4 levels: 1 year (reference level), 5 and 10 years collapsed, 15 and 30 years collapsed, and 50 years.

Estimates from models 1, 2 and 3 are shown in Table 2. In the three models, a number of estimates for the EQ-5D attribute levels are positive, and in model 1 and 3 some of these are statistically significant. Main effects estimates for duration are positive and significant in all three models. Estimates of interaction terms are negative in most cases and, in models 1 and 3, are often statistically significant.

### Relationships between quality and quantity of life on individuals' preferences

Figure 2a, b and c were generated to show the values for each health state at the different time durations. Model 1, in which duration was linear (similar to the standard QALY model), produced straight diverging lines (see Fig. 2a). The estimates from model 2, with different weights for the 4 collapsed levels of duration, showed non-linear trends for each health state (Fig. 2b). Some states were represented by increasing curves, for instance the state 11111. Some states, e.g. 23232, were initially increasing, and decreasing for longer durations. State 33333 is represented by decreasing curves. Turning to the logarithmic model 3 in Fig. 2c, curves with a decreasing slope for longer durations, i.e., with a decreasing marginal utility (or disutility) were found for states better and worse than dead. Although there are clear differences between Figs. 2b and c, the more flexible model 2 partly resembles the logarithmic model 3, supporting the logarithmic model.

The performance values in Table 2 suggest that values are non-linearly related to duration. This is shown by the poorer performance of model 1 compared to model 3, both having 21 parameters. Model 3 outperforms model 1 on all goodness of fit tests. Specifically, the AIC value for model 3 (AIC = 1,990.84) is much better than that of model 1 (AIC = 2,047.99).

## Discussion

With the present discrete choice (DC) study we developed a model for values assigned by individuals to health states with various durations. The model gives insights on the likely trend of these values for life durations lived up to 50 years, which corresponds to the life expectancy of the participants. The most important finding is the clear evidence for non-linear values for duration. In particular, values for duration were best summarized using a logarithmic function. This finding is novel in the context of DC experiments comparing health states with various durations.

Recently, Bansback and colleagues [16] published a discrete choice experiment which included a combination of the EQ-5D domain levels and a life-years attribute, in order to compare its performance with values obtained with the time trade-off method. They found that discrete choices are promising as a stand-alone method for producing values amenable for QALY calculations. The authors included in their design durations up to 10 years, and modeled duration as linear. In contrast, we included longer durations up to 50 years and found evidence for non-linear values for duration.

**Table 2** Results of model 1 (including duration as a linear variable), model 2 (in which duration is included as a categorical variable) and model 3 (including a logarithmic function for duration)

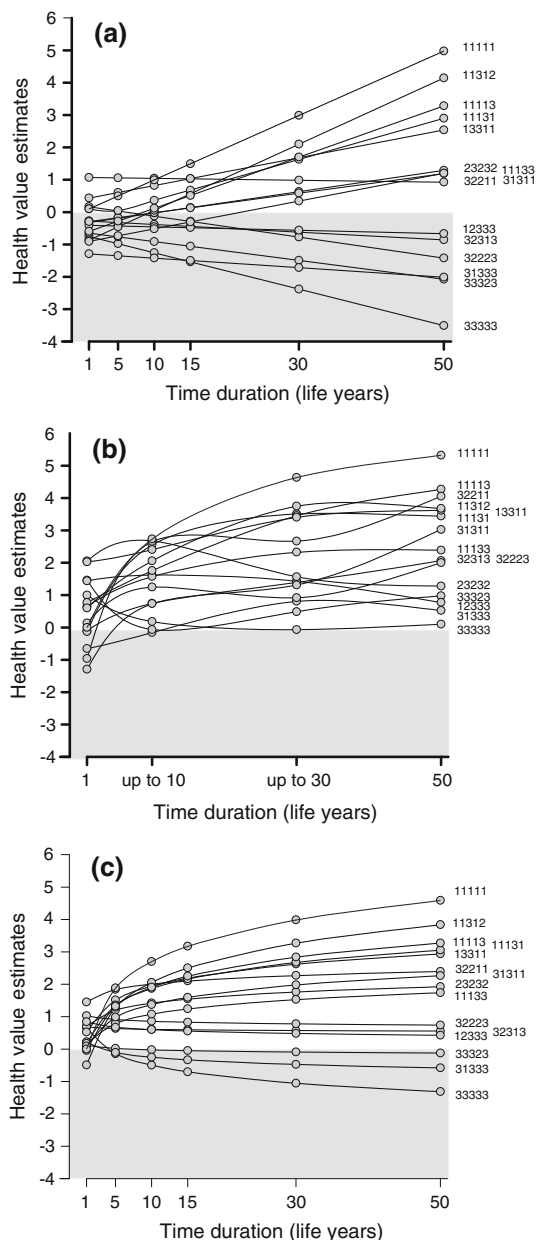
Description of independent variables	Model 1		Model 2		Model 3	
	Independent variables <sup>a</sup>	$\beta$ ( <i>p</i> )	Independent variables <sup>a</sup>	$\beta$ ( <i>p</i> )	Independent variables <sup>a</sup>	$\beta$ ( <i>p</i> )
Mobility	$X_{MO2}$	-0.351 (0.051)	$X_{MO2}$	0.028 (0.955)	$X_{MO2}$	-0.157 (0.592)
	$X_{MO3}$	-0.134 (0.428)	$X_{MO3}$	-1.665 (0.105)	$X_{MO3}$	0.016 (0.952)
	$X_{SC2}$	0.738 (0.000)	$X_{SC2}$	0.427 (0.278)	$X_{SC2}$	0.637 (0.013)
Self-care	$X_{SC3}$	0.581 (0.001)	$X_{SC3}$	1.654 (0.136)	$X_{SC3}$	0.448 (0.180)
	$X_{UA2}$	0.469 (0.011)	$X_{UA2}$	0.275 (0.681)	$X_{UA2}$	0.797 (0.016)
	$X_{UA3}$	-0.190 (0.257)	$X_{UA3}$	0.373 (0.341)	$X_{UA3}$	0.076 (0.761)
Pain/discomfort	$X_{PD2}$	-0.199 (0.191)	$X_{PD2}$	0.331 (0.713)	$X_{PD2}$	-0.381 (0.181)
	$X_{PD3}$	-0.273 (0.125)	$X_{PD3}$	-0.130 (0.820)	$X_{PD3}$	0.173 (0.469)
	$X_{AD2}$	-0.779 (0.000)	$X_{AD2}$	0.225 (0.808)	$X_{AD2}$	-0.568 (0.037)
Anxiety/depression	$X_{AD3}$	-0.675 (0.000)	$X_{AD3}$	0.767 (0.135)	$X_{AD3}$	-0.043 (0.859)
			$X_{5/10\text{ years}}$	2.735 (0.010)		
			$X_{15/30\text{ years}}$	4.636 (0.000)		
Duration	$X_{\text{years}}$	0.100 (0.000)	$X_{50\text{ years}}$	5.321 (0.000)	$N\log(X_{\text{years}})$	1.172 (0.000)
			$X_{MO2} \times X_{5/10\text{ years}}$	-0.497 (0.380)		
			$X_{MO2} \times X_{15/30\text{ years}}$	-0.299 (0.705)		
			$X_{MO2} \times X_{50\text{ years}}$	-1.060 (0.281)		
			$X_{MO3} \times X_{5/10\text{ years}}$	0.606 (0.568)		
			$X_{MO3} \times X_{15/30\text{ years}}$	0.500 (0.617)		
			$X_{MO3} \times X_{50\text{ years}}$	1.077 (0.350)		
			$X_{SC2} \times X_{5/10\text{ years}}$	0.293 (0.529)		
			$X_{SC2} \times X_{15/30\text{ years}}$	-0.968 (0.042)		
			$X_{SC2} \times X_{50\text{ years}}$	-0.768 (0.335)		
			$X_{SC3} \times X_{5/10\text{ years}}$	-1.314 (0.227)		

**Table 2** continued

Description of independent variables	Model 1		Model 2		Model 3	
	Independent variables <sup>a</sup>	$\beta$ ( <i>p</i> )	Independent variables <sup>a</sup>	$\beta$ ( <i>p</i> )	Independent variables <sup>a</sup>	$\beta$ ( <i>p</i> )
Interactions	$X_{MO2} \times X_{years}$	0.020 (0.059)	$X_{SC3} \times X_{15/30\ years}$	-2.533 (0.042)	$X_{MO2} \times MLog(X_{years})$	-0.003 (0.977)
	$X_{MO3} \times X_{years}$	-0.043 (0.000)	$X_{SC3} \times X_{50years}$	-2.080 (0.072)	$X_{MO3} \times MLog(X_{years})$	-0.363 (0.000)
	$X_{SC2} \times X_{years}$	-0.033 (0.000)	$X_{UA2} \times X_{5/10\ years}$	-0.015 (0.000)	$X_{SC2} \times MLog(X_{years})$	-0.264 (0.008)
	$X_{SC3} \times X_{years}$	-0.041 (0.000)	$X_{UA2} \times X_{15/30\ years}$	-0.041 (0.000)	$X_{SC3} \times MLog(X_{years})$	-0.302 (0.007)
	$X_{UA2} \times X_{years}$	-0.026 (0.000)	$X_{UA2} \times X_{50years}$	-0.026 (0.000)	$X_{UA2} \times MLog(X_{years})$	-0.304 (0.008)
	$X_{UA3} \times X_{years}$	-0.015 (0.061)	$X_{UA3} \times X_{5/10\ years}$	-0.015 (0.061)	$X_{UA3} \times MLog(X_{years})$	-0.253 (0.007)
	$X_{PD2} \times X_{years}$	-0.009 (0.236)	$X_{UA3} \times X_{15/30\ years}$	-0.009 (0.236)	$X_{PD2} \times MLog(X_{years})$	0.010 (0.924)
	$X_{PD3} \times X_{years}$	-0.036 (0.000)	$X_{UA3} \times X_{50years}$	-0.036 (0.000)	$X_{PD3} \times MLog(X_{years})$	-0.436 (0.000)
	$X_{AD2} \times X_{years}$	0.018 (0.012)	$X_{PD2} \times X_{5/10\ years}$	0.018 (0.012)	$X_{AD2} \times MLog(X_{years})$	0.188 (0.079)
	$X_{AD3} \times X_{years}$	-0.020 (0.004)	$X_{PD2} \times X_{15/30\ years}$	-0.020 (0.004)	$X_{AD3} \times MLog(X_{years})$	-0.325 (0.000)
			$X_{PD2} \times X_{50years}$			
			$X_{PD3} \times X_{5/10\ years}$			
			$X_{PD3} \times X_{15/30\ years}$			
			$X_{PD3} \times X_{50years}$			
		$X_{AD2} \times X_{5/10\ years}$				
		$X_{AD2} \times X_{15/30\ years}$				
		$X_{AD2} \times X_{50years}$				
		$X_{AD3} \times X_{5/10\ years}$				
		$X_{AD3} \times X_{15/30\ years}$				
		$X_{AD3} \times X_{50years}$				
Log-likelihood (LL)	1,002.997			956.8739		969.4204
Adjusted pseudo $R^{2b}$	0.219			0.246		0.245
AIC <sup>b</sup>	2,047.99			1,999.75		1,990.84
BIC <sup>b</sup>	2,164.22			2,237.74		2,097.07
No. of respondents	208			208		208
No. of choices	1,872			1,872		1,872

<sup>a</sup> In model 1 the domain levels are organized so that preferences for some problems (e.g.,  $X_{MO2}$ ) or for severe problems (e.g.,  $X_{MO3}$ ) are estimated against preferences for having no problems (reference level). In model 2, duration is collapsed into 4 levels

<sup>b</sup> A higher adjusted pseudo  $R^2$  shows a better model fit, while lower AIC and BIC show a better compromise between parsimony and model fit



**Fig. 2** Value estimates from model 1 (a), from model 2 (b) and from model 3 (c).  $V$  estimated for the state 11111 lived for 1 year is 0 (reference health profile).  $V < 0$  is interpreted as less preferred than living in the state 11111 for 1 year (gray area),  $V > 0$  is interpreted as more preferred than living in the state 11111 for 1 year (white area)

We observed multiplicative terms between duration and the EQ-5D levels, which is in accordance with the multiplication of quality and duration in the QALY model. This agrees with findings by other researchers [4–7, 16] that the value of a health profile depends on the time spent in a certain health state. However, in the standard QALY model, as applied in decision models, the value for different states is estimated on durations of up to 10 years under the assumption of linearity. Then, QALYs are

calculated using a linear extrapolation to longer durations. Instead, our results suggest that the standard QALY model may need adjustment at least when the value for health states with longer durations are calculated. Furthermore, our results suggest that non-linear trends might be present also for durations shorter than 10 years.

The results of our study can also be placed in the perspective of the maximum endurable time (MET) phenomenon [6, 26–28]. A study [29] examined the subjects' preferences for health states of different time durations compared with immediate death. It was shown that some health states, especially moderate and severe ones, are considered better than death for a short duration, but worse than death for longer durations. Similar patterns appeared also to be present in our data.

We certainly do not claim to be the first to establish the non-linear relationship for time using DC in general. Such results have been found previously in the field of intertemporal choice [30], and in the health domain by comparing health states with different time delays [31]. However, our results are novel in the context of DC experiments which compare health states with various durations, aiming to arrive at tariffs for health economical evaluations.

A limitation of the present study is that the design of the experiment that included duration could have been more efficient. Furthermore, the small sample size could have reduced opportunities for finding significant main effects for the attributes other than time. The sample consisted of university students only, who represent a highly selected part of the general population. The likely consequence of these limitations is that results are less precise, yielding estimates with unexpected sign or low statistical significance. On the other hand, an advantage of the sample used was the possibility of investigating values for durations up to 50 years. Additional studies should be designed and conducted to obtain more reliable results, applicable to a wider population.

To conclude, this study gives clear evidence that values for health states are non-linearly related with duration. These results were obtained in a discrete choice experiment, in which health profiles were presented with durations up to 50 years. This research suggests that refinement of the standard QALY framework may be needed and further research needs to be done to obtain more precise results and investigate them in a broader population.

**Acknowledgments** This work was partially based on our inspiring discussions and reflections during several meetings of the EuroQol Valuation Task Force. Therefore, we wish to express our thanks in particular to Frank de Charro, Nancy Devlin, Ben van Hout, Paul Kind, and David Parkin. A special thank you to Elly Stolk and Mark Oppe for their precious contributions to the study design and data collection. We thank also Eva de Vries, Anke Naaborgh, and Mark Huijsman for their assistance in the data collection, and Roberta



Ciampichini for her assistance in data management. Everyone who contributed significantly to the work has been listed in the acknowledgements. This research was made possible by a grant provided by the EuroQoL Group. It was presented at the ISPOR 13th Annual European Congress in Prague, November 2010, and was a candidate for receiving the best poster research presentation award.

**Conflict of interest** The authors of this manuscript have no conflict of interest to declare.

## References

- Fanshel, S., Bush, J.W.: A health-status index and its application to health-services outcomes. *Oper. Res.* **18**, 1021–1066 (1970)
- Sackett, D.L., Torrance, G.W.: The utility of different health states as perceived by the general public. *J. Chronic. Dis.* **31**, 697–704 (1978)
- Pliskin, J.S., Shepard, D.S., Weinstein, M.C.: Utility functions for life years and health status. *Oper. Res.* **28**, 206–224 (1980)
- Stalmeier, P.F.M., Wakker, P.P., Bezembinder, T.G.G.: Preference reversals: violations of unidimensional procedure invariance. *J. Experimental Psychology-Human Percept. Perform.* **23**(4), 1196–1205 (1997)
- Krabbe, P.F.M., Bonsel, G.J.: Sequence effects; health profiles, and the QALY model. In search of realistic modeling. *Med. Decis. Mak.* **18**, 178–188 (1998)
- Dolan, P., Stalmeier, P.: The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. *J. Health Economics* **22**(3), 445–458 (2003)
- Spencer, A.: A test of the QALY model when health varies over time. *Soc. Sci. Med.* **57**, 1697–1706 (2003)
- Ryan, M., Gerard, K., Amaya-Amaya, M. (eds.) *Using Discrete Choice Experiments to Value Health and Health Care*. Springer, Berlin (2008)
- Hakim, Z., Pathak, D.S.: Modelling the Euroqol data: a comparison of discrete choice conjoint and conditional preference modelling. *Health. Econ.* **8**, 103–116 (1999)
- Szeinbach, S.L., Barnes, J.H., McGhan, W.F., et al.: Using conjoint analysis to evaluate health state preferences. *Drug. Inf. J.* **33**, 849–858 (1999)
- McKenzie, L., Cairns, J., Osman, L.: Symptom-based outcome measures for asthma: the use of discrete choice methods to assess patient preferences. *Health. Policy.* **57**, 193–204 (2001)
- McCabe, C., Brazier, J., Gilks, P., et al.: Using rank data to estimate health state utility models. *J. Health. Econ.* **25**, 418–431 (2006)
- Ratcliffe, J., Brazier, J., Tsuchiya, A., et al.: Using DCE and ranking data to estimate cardinal values for health states for deriving a preference-based single index from the sexual quality of life questionnaire. *Health. Econ.* **18**, 1261–1276 (2009)
- Stolk, E., Oppe, M., Scalone, L., Krabbe, P.F.M.: Discrete choice modeling for the quantification of health states: the case of the EQ-5D. *Value in Health.* **13**(8), 1005–1013 (2010)
- Brazier, J., Rowen, D., Yang, Y., Tsuchiya, A.: Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. *Eur. J. Health. Econ.* **13**(5), 575–587 (2012)
- Bansback, N., Brazier, J., Tsuchiya, A., Anis, A.: Using a discrete choice experiment to estimate health state utility values. *J. Health. Econ.* **31**(1), 306–318 (2012)
- Rabin, R., de Charro, F.: EQ-5D: a measure of health status from the EuroQol Group. *Ann. Med.* **33**, 337–343 (2001)
- Rose, J.M., Bliemer, M.C.J.: Stated preference experimental design strategies. In: Hensher, D.A., Button, K. (eds.) *Transport modeling* (2nd ed.). Handbooks in transport, 1st edn. Elsevier, Oxford (2007)
- Shaw, J.W., Johnson, J.A., Coons, S.J.: US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med. Care.* **43**, 203–220 (2005)
- Lamers, L.M., McDonnell, J., Stalmeier, P.F.M., et al.: The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health. Econ.* **15**, 1121–1132 (2006)
- Dolan, P.: Modeling valuations for the EuroQoL health states. *Med. Care.* **35**, 1095–1108 (1997)
- Huber, W., Zwerina, K.: The importance of utility balance in efficient choice designs. *J. Marketing Research.* **33**, 307–317 (1996)
- Louviere, J.J., Hensher, D.A., Swait, J.D.: *Stated choice methods: analysis and application*. Cambridge University Press, Cambridge (2000)
- Manski, C.: The structure of random utility models. *Theor. Decis.* **8**, 229–254 (1977)
- Train, K.E.: *Discrete choice methods with simulation*. Cambridge University Press, Cambridge (2003)
- Sutherland, H.J., Llewellyn-Thomas, H., Boyd, N.F., Till, J.E.: Attitudes towards quality of survival: the concept of maximum endurable time. *Med. Decis. Making.* **2**, 299–309 (1982)
- Stalmeier, P.F.M., Chapman, G.B., de Boer, A.G.E.M., et al.: A fallacy of the multiplicative QALY model for low quality weights in students and patients judging hypothetical health states. *Int. J. Technol. Assess. Health. Care.* **17**, 488–496 (2001)
- Robinson, A., Spencer, A.: Exploring challenges to TTO utilities: valuing states worse than dead. *Health. Econ.* **15**, 393–402 (2006)
- Stalmeier, P.F.M., Lamers, L.M., Busschbach, J.J.V., Krabbe, P.F.M.: On the assessment of preferences for health and duration maximal endurable time and better than dead preferences. *Med. Care.* **45**, 835–841 (2007)
- Loewenstein, G.F., Prelec, D.: Preferences for sequences of outcomes. *Psychol. Rev.* **100**(1), 91–108 (1993)
- van der Pol, M., Cairns, J.: Estimating time preferences for health using discrete choice experiments. *Soc. Sci. Med.* **52**, 1459–1470 (2001)