VALUE IN HEALTH 19 (2016) 343–352

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Preference-Based Assessment

Dutch Tariff for the Five-Level Version of EQ-5D

Matthijs M. Versteegh, PhD[1,*], *Karin M. Vermeulen, PhD*[2], *Silvia M. A. A. Evers, PhD*[3,4], *G. Ardine de Wit, PhD*[5,6], *Rilana Prenger, PhD*[7], *Elly A. Stolk, PhD*[8]

[1]*Institute for Medical Technology Assessment, Erasmus University of Rotterdam, Rotterdam, the Netherlands;* [2]*Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands;* [3]*CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, the Netherlands;* [4]*Trimbos Institute, Netherlands Institute for Mental Health and Addiction, Utrecht, the Netherlands;* [5]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands;* [6]*National Institute of Public Health and the Environment, Bilthoven, the Netherlands;* [7]*Faculty of Behavioural, Management and Social Science, University of Twente, Enschede, the Netherlands;* [8]*Institute of Health Policy and Management/Institute for Medical Technology Assessment, Erasmus University of Rotterdam, Rotterdam, the Netherlands*

ABSTRACT

Background: In 2009, a new version of the EuroQol five-dimensional questionnaire (EQ-5D) was introduced with five rather than three answer levels per dimension. This instrument is known as the EQ-5D-5L. To make the EQ-5D-5L suitable for use in economic evaluations, societal values need to be attached to all 3125 health states. **Objectives:** To derive a Dutch tariff for the EQ-5D-5L. **Methods:** Health state values were elicited during face-to-face interviews in a general population sample stratified for age, sex, and education, using composite time trade-off (cTTO) and a discrete choice experiment (DCE). Data were modeled using ordinary least squares and tobit regression (for cTTO) and a multinomial conditional logit model (for DCE). Model performance was evaluated on the basis of internal consistency, parsimony, goodness of fit, handling of left-censored values, and theoretical considerations. **Results:** A representative sample (N = 1003) of the Dutch population participated in the valuation study. Data of 979 and 992 respondents were included in the analysis of the cTTO and the DCE, respectively. The cTTO data were left-censored at −1. The tobit model was considered the preferred model for the tariff on the basis of its handling of the censored nature of the data, which was confirmed through comparison with the DCE data. The predicted values for the EQ-5D-5L ranged from −0.446 to 1. **Conclusions:** This study established a Dutch tariff for the EQ-5D-5L on the basis of cTTO. The values represent the preferences of the Dutch population. The tariff can be used to estimate the impact of health care interventions on quality of life, for example, in context of economic evaluations.

Keywords: utility measurement, discrete choice experiment, EQ-5D-5L, time trade-off.

Introduction

In 2009, the EuroQol Research Foundation introduced a new descriptive system for the measurement of health, the five-level EuroQol five-dimensional questionnaire (EQ-5D-5L) [1]. Similar to the previous version of the EQ-5D, the EQ-5D-5L measures health-related quality of life on five dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The EQ-5D-5L can describe 3125 (5^5) unique health states because each dimension has five answer categories (levels)—no problems, some problems, moderate problems, severe problems, and extreme problems/unable to—compared with the three levels in the previous EQ-5D-3L version. To make the EQ-5D-5L suitable for use in economic evaluations, the health states need to be valued with a preference-elicitation method. This article reports how Dutch values for the EQ-5D-5L were collected and subsequently modeled.

The EQ-5D-5L has been introduced in response to perceived limitations of the EQ-5D-3L. Although the EQ-5D is a preferred generic utility measure in the United Kingdom [2], users have expressed concern over the crude three-level structure of the EQ-5D and there is evidence of ceiling effects in patient populations [3–5]. Measuring health problem intensity with just three levels restricts the instrument's potential to detect small differences in health and to evaluate changes in health-related quality of life of patients with mild conditions. To improve the discriminatory

Conflict of interest: M. Versteegh and E. Stolk are members of the EuroQoL Research Foundation.

* *Address correspondence to*: Matthijs Versteegh, Institute for Medical Technology Assessment (iMTA), Erasmus University of Rotterdam, Rotterdam, the Netherlands.

E-mail: versteegh@imta.eur.nl.

http://dx.doi.org/10.1016/j.jval.2016.01.003

potential of the EQ-5D, the EQ-5D-5L has been developed with an increased number of answer levels from three to five. Available evidence on the comparative performance of the EQ-5D and the EQ-5D-5L suggest that the EQ-5D-5L is a valid, reliable, and useful improvement over the previous EQ-5D [1,6,7]. The EQ-5D-5L has less ceiling effects and greater discriminative ability with potentially more power to detect differences between groups compared with the EQ-5D [8,9]. At present there are 123 different language versions of the EQ-5D available. In the development of the EQ-5D-5L, the next step is to generate national value sets that make the instrument suitable for use in economic evaluations.

In the past, the outcomes of valuation studies for the EQ-5D in different countries lacked comparability because of differences in the research protocols that were used [10]. To increase inter-country comparability of both valuation studies and outcome studies, and to increase the likelihood that observed differences between the EQ-5D-5L values collected in different countries reflect population preference differences rather than method heterogeneity, a protocol has been developed for valuation studies of the EQ-5D-5L [11]. This protocol is the result of several empirical studies [12–21] and describes the requirements for data collection. The protocol does not prescribe a preferred modeling approach because this depends on the characteristics of the data that are obtained and cannot be defined a priori.

The protocol includes two health state valuation tasks: a time trade-off (TTO) and a discrete choice experiment (DCE). The TTO is a widely used method for health state valuation that has been extensively used for EQ-5D valuation studies in the past. It is an interview-based method for health state valuation that may be difficult for respondents because of the iterative nature of the tasks and its questions concerning life and death [22]. Also, it often results in censored data because the properties of the TTO design define the lowest measurable value [23]. In the protocol, the TTO task is accompanied by a DCE task that measures preferences for the various health states on a latent scale. The TTO and the DCE may be complementary in the sense that both intend to measure the same construct (quality of life in a particular state of health) and that both procedures may come with idiosyncratic strengths and weaknesses. The use of multiple methods may then enhance the understanding of people's preferences for EQ-5D-5L health states.

This article reports on the EQ-5D-5L valuation study that was conducted in the Netherlands and the subsequent modeling strategies for estimating a "tariff," that is, an algorithm that can be used to attach values to all 3125 health states for use in economic evaluations.

Methods

Respondents

Interviews took place in five cities located in different parts of the Netherlands: Utrecht (central), Rotterdam (west), Maastricht (south), Enschede (east), and Groningen (north). A stratified sampling approach was used in which three strata were a priori defined for age (deciles, with being at least 18 years old as the eligibility criterion), sex, and education (eight levels including "unknown") on the basis of the distribution in the Netherlands as recorded by Statistics Netherlands (Centraal Bureau voor de Statistiek [CBS]). The education stratum was recoded for this publication to be internationally comprehensible in "higher," "lower," and "middle." Respondents were then randomly drawn from the panel until the strata quote was met. Respondents were recruited from these cities and their surrounding areas to achieve sufficient geographical spread. The interviews in each city took place in different months to allow selective recruiting during the progress of the study to maximize the samples' representativeness at the cost of having the most difficult to sample strata being recruited from specific parts of the Netherlands. Respondents were sampled from a commercial panel, and they received a financial incentive of €20 and an additional €7.50 as travel reimbursement. The sample was selected to represent the Dutch population in 2012 in terms of the distribution of age, sex, and education as recorded by the CBS. This survey aimed to elicit preferences of the general population about the severity of states of health and does therefore not fall under the scope of the Medical Research Involving Human Subjects Act in the Netherlands, exempting it from ethical review.

Health State Valuation Tasks

Two types of stated preference methods were applied in the study: the composite TTO (cTTO) and a DCE (without duration), embedded in a digital aid and accompanied by a face-to-face interviewer and an interview script. The study by Janssen et al. [17] contains a detailed description of the cTTO method. The concept of cTTO for the valuation of health states considered better than dead is identical to that of "classic" TTO that has been used in most EQ-5D-3L valuation studies: a TTO score indicates the amount of time in full health x that is considered equivalent, after a series of choice-based iterations, to a period of time t in an impaired state of health. The value of t has been set at 10 years in EQ-5D valuation studies, and the health state value is defined as x/t. Respondents who indicated that they consider the health state under valuation so poor that they would rather die immediately than have to live t years in the health state were switched to a lead-time TTO task. In the lead-time TTO task, the health state under valuation still lasted for t years, but it was preceded by a period of time l in full health. Respondents then are able to trade in their lead time l to express negative values. In this study, l was set at 10 and thus the ratio of lead time to disease time was 1:1. This ratio defines -1 as the lowest attainable value, which is computed with $(x - 10)/10$. Hence, the lowest attainable value is -1, at $x = 0$. The smallest tradable unit was 6 months, or 0.05, when expressed as health state value.

The DCE task presented respondents with two different EQ-5D-5L health states in which the levels, but not the order of the attributes, differed as experimented with in two previous studies [20,21]. In the present study, the DCE-derived values were estimated on a latent scale and not on full health (utility = 1) and death (utility = 0). The DCE task in this study can therefore not be used independently to estimate health state values. In this study, the DCE data were used to identify appropriate cTTO modeling techniques.

Health State Descriptions and Experimental Design

The EQ-5D-5L descriptive system (Dutch translation) was used in bullet-point format to describe the health states that were presented in the cTTO and DCE tasks. The EQ-5D-5L dimensions are mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension of health has five levels of severity, ranging from no problems (1) to extreme problems/unable to (5). Each health state is identified by a five-digit number that contains the severity levels per dimension; the first digit represents the severity level in the mobility dimension, the second in the self-care dimension, and so on. Hence, state 55555 refers to a state with the highest level of problems on each dimension.

For cTTO, 86 health states were included in the study. Health state 55555 (included in all 10 blocks) and the five mild states (21111, 12111, 11211, 11121, and 11112, each included in two blocks) were selected a priori for the design. The mild health states were purposefully included because it was expected that direct observations were required to statistically distinguish minor impairments from full health. The mild and the 55555 health states were supplemented with 80 additional states selected using the design strategy described in the study by Oppe and van Hout [24]. This design for TTO was generated by randomly selecting 80 health states from the full fractional design minus the six health states mentioned earlier. Next, an expected health state value was assigned to each of the 86 states in the design on the basis of priors. Subsequently, a regression model was estimated on the data set to calculate predicted values for each health state. The strategy was "looped" until a design was identified with small differences between prior and predicted values. The 80 states were distributed over 10 blocks in such a way that the full utility range was more or less covered within a block while all blocks also would have the same mean utility. In the final design, every block included eight states unique to that block, supplemented with state 55555 and one of the mild states. The DCE task involved forced choices between two health states described by the EQ-5D dimensions. One hundred ninety-six pairs were included in the DCE experimental design generated by Oppe and van Hout [24]. These pairs were distributed over 28 blocks, which resulted in seven pairs per person for the DCE task. The blocks were balanced in terms of their severity, which was calculated as the sum of the level scores on all dimensions. Block assignment, question order, and (in DCE) the left-right position of health states in the choice tasks were all randomized.

Sample Size Calculation

In accordance with the protocol for EQ-5D-5L valuation studies, the target sample size for this study was set at n = 1000. This sample size was proposed to balance statistical requirements for the modeling of the cTTO data and the feasibility of data collection. There were three key considerations. First, a previous study had identified that about 100 observations per health state result in standard error of mean health state utilities between 0.01 (for mild states) and 0.06 (for poor states) [17], which was considered sufficient precision. Second, the regression model that would be fitted to the data would have to allow for at least 21 parameters (i.e., four dummy variables for each of the five dimensions of health and a constant). Third, there was a limit to the number of health states an individual could value in one and the same task was set at 10 per person. Design considerations translated into a general design based on 10 blocks with at least eight unique health states. Aiming for 100 observations per health state leads to a sample size of 1000. This number corresponds to a standard multivariate regression sample size calculation of n = 1064 with a power of 0.9, five predictor variables (21 including dummies), and a small effect size of $f^2 = 0.02$ (assuming an R^2 of 0.5 and 0.51) in STATA (StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP) using the "powerreg" function.

Data Collection Process

Data were collected in the fall of 2012 using computer-assisted personal interviews. The respondents used the computer to complete all tasks. An interviewer was present to serve as host, explain all the tasks, and guide the respondents through the interview. Hereto, the EuroQol standard protocol for EQ-5D-5L valuation studies [11] has been embedded in a digital aid called the EuroQol valuation technology (EQ-VT) and accompanying materials (interviewer manual and training materials) that were designed to standardize interviewer behavior and promote quality. We used version 1.0 of the EQ-VT. Visual representation of the EQ-VT is described elsewhere [11].

The structure of the interview was as follows. First, the interviewer welcomed the respondent and explained the purpose of the research. Second, respondents recorded their own health state on the EQ-5D-5L and the EQ-5D visual analogue scale and answered background questions regarding age, sex, and experience with illness. Third, the interviewer and the respondent discussed how to interpret and carry out the cTTO task, using the hypothetical state "being in a wheelchair" as example. When respondents indicated that they understood the task, they valued 10 EQ-5D-5L health states. Fourth, respondents received instructions on how to carry out the DCE task and subsequently they completed seven choice sets. Finally, the respondents were given an opportunity to leave their comments, if any, about the study and were thanked for their cooperation.

In each of the five cities a local team of interviewers was trained. In total, 21 interviewers with an academic education (either enrolled in or having completed a Bachelor of Science degree, including PhD students, postdocs, and senior researchers) and some previous knowledge of EQ-5D were involved in this study. To guarantee equivalent task understanding, procedures, and interaction with respondents for all interviewers, the interviewers were trained in a day-long training session by E.S. and M.V. In this training session the interviewers received the word-for-word interview script with screenshots of the software. They were also introduced to the software and were made to perform several practice interviews under the supervision of E.S. and M.V. (three interviews demonstrated in front of the class, two during breakout sessions, and at least one practice session unsupervised at home with friends or family including uploading of the data to the repository to allow a check of the data). Each team was supervised by a local lead investigator who held at least a PhD and had experience with conducting valuation studies and by the principal investigator (E.S.). Data collection was monitored and screened for quality by the principal investigators, and the interviewers were reminded to follow the protocol on a regular basis. Close attention was paid to the distribution of values and suspect response patterns indicating "task short-cutting."

Exclusion Criteria

The cTTO data from respondents were excluded when the task was not finished (which resulted in not uploading the data to the repository) or when interviewers had indicated to the principal investigator that the respondent clearly was not able to understand the task or when a respondent gave the same value to all health states in the cTTO task. These exclusions were based on the argument that these respondents were unable to discriminate severity levels of health states using the cTTO task. Because this argumentation pertains only to the cTTO task, the DCE data of these respondents were not excluded.

Modeling cTTO and DCE

cTTO values were modeled using main effects models that included a constant and 20 main effects derived from the EQ-5D-5L descriptive system, using ordinary least squares (OLS) and tobit models. The constant is interpreted to reflect the utility decrement associated with any deviation from full health. Random effects were included to account for the panel structure in the data. The basic equation for the random-effects OLS regression with random intercept is given in Equation 1:

$$Y_{it} = \beta 0_i + MO_{it}\beta_{MO} + SC_{it}\beta_{SC} + UA_{it}\beta_{UA} + PD_{it}\beta_{PD} + AD_{it}\beta_{AD} + \varepsilon_{it} \quad (1)$$

where Y_{it} refers to the TTO values as dependent variables. The terms MO, SC, UA, PD, and AD are five dummy-coded regressors, respectively, for mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, each representing the four levels beyond "no problems" of the five dimensions of the EQ-5D-5L. In the equation,

$$MO_{it}\beta_{MO} = MO1_{it}\beta_{1m} + MO2_{it}\beta_{2m} + MO3_{it}\beta_{3m} + MO4_{it}\beta_{4m} + MO5_{it}\beta_{5m},$$

which is similar for SC, UA, PD, and AD, leading to a total of 20 regressors plus the constant. ε is the error term, i indicates the respondent, and t accounts for the panel structure of the data set (because there are 10 cTTO questions per respondent).

The tobit model assumes a latent variable Y^*_{it} underlying the observed Y_{it} cTTO values. This matches well with the censored cTTO data, which by nature of the applied cTTO task were left-censored at −1. The tobit model accounts for this censored nature of the data by estimating the latent variable Y^*_{it}, which can take on predicted preference values extrapolated beyond the range of the observed values. This is a favorable model characteristic because observed preference values were censored by the cTTO methodology at −1, whereas latent preferences of respondents might include valuations lower than −1. A likelihood function is used to adjust the parameter estimates for the probability of Y_{it} being above the censoring value. Hence, in the tobit model, the observed value Y_{it} has the following properties when the censoring value is −1:

$$Y_{it} = \begin{cases} Y^*_{it} \text{ if } Y^*_{it} > -1 \\ -1 \text{ if } Y^*_{it} \leq -1 \end{cases} \quad (2)$$

The equation for Y^*_{it} is linear and similar to Equations 1 and 2, with the parameter vectors estimated with the tobit likelihood function.

The DCE data were modeled under random utility using the conditional logit model. The model included the same 20 dummy parameters as used in the cTTO model, reflecting utility decrements associated with levels 2, 3, 4, and 5 for each of the five domains: MO, SC, UA, PD, and AD. Equation 3 shows the regression equation, where j is the choice alternative in choice sets.

$$U_{ij} = MO_{js}\beta_1 + SC_{js}\beta_2 + UA_{js}\beta_3 + PD_{js}\beta_4 + AD_{js}\beta_5 + \varepsilon_{js} \quad (3)$$

Hence, besides the DCE model, three potential tariff models were estimated: model 1, the random-effects linear regression described in Equation 2; model 2, the tobit model; and model 3, the most parsimonious version of the best-fitting model (i.e., model 1 or model 2). Model performance criteria are mentioned in the "Model Selection" section of this article. Regression analyses were performed in STATA 12.0 (StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP).

Sensitivity Analysis

A sensitivity analysis was conducted to explore how the presence of severely inconsistent responders impacted on the modeling of DCE and cTTO results. Hereto, all cTTO responses were removed of respondents who 1) valued state 55555 higher than any other state and the difference was equal to or larger than 0.5 on the utility scale and 2) valued the mild state (>0.5) lower than any other state that it logically dominated. A pair of cTTO responses was defined as logically inconsistent when the observed values of two states, state A and state B, contradicted the logical ordering of health states. That is, if state A was better on at least one dimension and no worse on other dimensions compared with state B, then state A should have logically received a higher value. If state B received a lower value instead, the response was then defined as logically inconsistent. Considering, however, that many inconsistencies may be the result of random error, the "seriousness" of the inconsistencies was evaluated by the size of utility difference between two states. During inspection of the data, it was identified that particular concern is warranted with utility differences of more than 0.5. Inconsistencies involving utility differences of more than 0.5 typically involved situations in which one or more states were valued as worse than dead, whereas state 55555 was not, or when the mild states had received a low value (e.g., utility = 0) that seemed to be mistakenly provided. This kind of inconsistency may be prevented if interviewers would pay closer attention to consistency of answers at the sorting question, or may be corrected if respondents were provided with the opportunity to review their responses and take the wrong ones out, if any. In contrast, random error will always occur and is typically not considered a sufficient reason for exclusion. For this reason, the sensitivity analysis excluded only the subset of inconsistent responses defined earlier.

DCE responses were considered problematic when respondent's responses followed a pattern indicating use of simplifying heuristics (i.e., when choosing between option a and option b, display the following response pattern: aaaaaaa, bbbbbbb, abababa, bababab). Regressions were rerun to assess the impact of removing DCE data that followed one of these patterns.

Model Selection

Models were compared regarding logical consistency, significance of the parameters, and predictive performance. A model was considered logically consistent when the coefficient preserved the severity ordering of the levels in each dimension. Predictive performance was analyzed by comparing predicted and observed values of cTTO using mean absolute error (MAE). It must, however, be noted that tobit extrapolates modeled values purposefully beyond the range of observed values, which may cause MAE to be an insufficient criterion for model selection in itself because the MAEs of tobit models will be higher when there is significant censoring in the data. To choose between OLS and tobit models, agreement with DCE results was explored. This was assessed by comparing the mean absolute difference, referred to as "DCE fit," between DCE values for all 3125 health states and values generated by the OLS and tobit models. DCE values are uncensored and on a latent scale and hence it is hypothesized that the tobit-predicted values, which are adjusted for censoring, are more similar to DCE-predicted values than to OLS-predicted values. The final model would be subjected to monotonicity constraints, if necessary.

Dutch EQ-5D-5L Reference Values

Reference values for the Dutch general population were calculated by multiplying the EQ-5D scores of the respondents selected for the model (N = 979) with the coefficients of the preferred regression model, that is, the new Dutch tariff. The sample was stratified on age, sex, and education, and it is for this stratification that the sample is representative.

Comparison of EQ-5D-3L Values with EQ-5D-5L Values

The EQ-5D-3L and EQ-5D-5L value sets were compared using the distribution of attainable values of both instruments as well as the distribution of observed values in a data set containing EQ-5D-3L and EQ-5D-5L responses of 3476 respondents. The health state values were generated on the basis of the Dutch EQ-5D-3L tariff [25], and the new tariff was estimated for the EQ-5D-5L. Kernel density graphs were produced to compare the distribution of attainable and observed values, or, in other words, to compare the theoretical and practically relevant evaluation spaces.

The first graph was created using a data set that contained all possible health states of both instruments, valued using the Dutch EQ-5D-3L tariff [25] and the new tariff estimated here for the EQ-5D-5L. The second graph was based on a data set including responses of 3476 respondents to whom EQ-5D-5L and EQ-5D-3L had been coadministered. The data were collected in six countries (Denmark, England, Italy, the Netherlands, Poland, and Scotland) and included patients with a range of diseases that guaranteed a good spread of observations on the dimensions of the EQ-5D. The data and their collection process are described elsewhere [26], with the only difference being that for this Dutch tariff study the data from students in Poland (N = 443) were excluded from the total data set (N = 3919).

Results

Characteristics of the Sample

In total, 1003 respondents attended the interview, and for 999 respondents recruitment characteristics were available. No data, however, were obtained from 11 respondents because of different reasons: the task was completed but the data were lost because of technical problems (N = 2); respondents could not start because of technical problems (N = 1); respondents could not start because of the absence of an interviewer (N = 2); respondents were not willing to participate after being informed about the subject (N = 3); or because of other/unknown reasons (N = 3). Some responses (N = 3) were removed from the cTTO on the basis of an interviewer's decision that the respondent did not understand or refused to complete the task. Furthermore, eight respondents did not trade any time and two gave the same value to all health states in the cTTO task but not in the DCE task. In both instances, the cTTO response suggests that preferences concerning the trading of time were independent of health state severity; their cTTO responses were excluded. For DCE, no additional respondents were excluded. Accordingly, the cTTO and DCE data sets contained responses of 979 (1003 − 24) and 992 (1003 − 11) respondents, respectively.

Table 1 – Respondents' characteristics.

Characteristics	Count	%	Dutch statistics (%)
Sampling characteristics			
Age (y)			
18 and 19	26	2.6	3.1
20–30	160	16.0	15.7
30–40	137	13.7	15.8
40–50	206	20.6	19.7
50–60	189	18.9	17.7
60–70	160	16.0	14.8
70–80	113	11.3	8.8
80 and older	8	0.8	4.4
Sex			
Male	491	49.1	49.3
Female	508	50.9	50.7
Education			
Lower education	385	38.4*	44.0
Middle education	322	32.1	27.5
Higher education	292	29.1	27.6
Unknown	4	0.4	0.9
Other characteristics			
Nationality			
Dutch	866	86.7	79.1
First- or second-generation immigrant	133	13.3	20.9
Marital status (married)	417	41.7	46.9
Income (€)			
<15,000	250	25.0	40.5
15,000–30,000	398	39.8	33.0
30,000–60,000	286	28.6	14.6
>60,000	65	6.5	11.9

*Significant difference at $\alpha = 0.05$.

The characteristics of the recruited sample (N = 999) are presented in Table 1. The sample is representative of the Dutch population for sex, age (except for age category 80–90 [t = −12.9; $P < 0.00$]), and level of education as recorded by the CBS in 2012. In the aggregated categories, there is a slight but significant underrepresentation of lower educated respondents (t = −3.56; $P < 0.00$), but this is the cause of aggregating smaller non-significant differences between the eight underlying education categories. The multicenter recruitment strategy resulted in a high level of geographical dispersion across the Netherlands, but with some clustering around research centers (Fig. 1) and an underrepresentation of inhabitants of the Noord-Holland province. The mean self-reported EQ-5D visual analogue scale score was 80.6. Health problems were most frequently reported in the pain dimension (49%) and least frequently in self-care (4%).

Data Characteristics

The observed cTTO values (N = 979) are presented in Table 2. Observed values ranged from −0.309 for state 55555 to 0.927 for state 11122. States 11112 and 11121 were valued at 0.907 and 0.915, respectively. The mean observed value was negative for 8 out of the 86 states that were included in the design. A clustering of values was observed at 1, 0, −1, and 0.5 in data pooled over respondents and health states (Fig. 2), reflecting strong agreement across respondents on the valuation of the mild and poor states. Many respondents were not willing to trade off any life-years to avoid mild health problems, whereas the poor health states frequently yielded values of −1 (this was the bottom value for 345 respondents) or 0 (bottom value for 122 respondents). The values −1 and 0 were used more than once by 221 and 55 respondents, respectively. Logical inconsistencies involving state 55555 and the mild state occurred frequently in 22% and 11% of respondents, respectively. Inconsistencies involving the mild states often traced back to small utility differences that most likely reflected random uncertainty about the values of the mild state and the one dominating it. Nevertheless, 87 respondents (8.8%) valued state 55555 minimally 0.5 higher than another health state in their set. These respondents seemed inconsistent in their decision to enter the worse-than-dead task. In the DCE data, 23 respondents (2.3%) answered following a specific pattern (aaaaaaa, bbbbbbb, abababa, bababab) but were not excluded from the analyses.

Discrete Choice Experiment

The DCE model results are presented in Table 3. The model that included 20 dummy variables contained only significant parameters but with inconsistent ordering of levels 2 and 3 in three dimensions (mobility, usual activities, and pain/discomfort).

Modeling cTTO

Regression models on cTTO data are presented in Table 4. The OLS model 1 outperformed the tobit model 2 in terms of reduced prediction errors as measured by MAE. The largest difference between parameter estimates of the OLS model and the tobit model was 0.029 for pain/discomfort level 5, with an average absolute difference of 0.008 for all parameters combined. The base-case models (models 1 and 2) had significant parameters ($P < 0.05$) but inconsistent ordering of levels 4 and 5 in the usual activities and self-care dimensions for both the OLS and tobit models. Tobit predictions for the 3125 health state values of EQ-5D-5L deviate less from DCE values for these health states as indicated by DCE fit. The relationship between DCE and tobit values is graphically presented in Figure 3.

With 1084 left-censored observations, the tobit model is theoretically preferred over the OLS model because of its

Fig. 1 – Place of residence of participating respondents in The Netherlands. (Color version of figure available online).

capability to extrapolate values beyond the censoring value of −1 inherent to the cTTO design. Combined with the other selection criteria, the tobit model was chosen over the OLS model to serve as the basis for the EQ-5D tariff. The most parsimonious version of this model, with constraints to deal with illogical ordering of parameters, is presented in Table 4 as model 3. Using this model as EQ-5D tariff leads to a value of −0.446 for health state 55555 and the highest value of 0.918 for health state 21111.

Sensitivity Analysis

The sensitivity analysis suggested that exclusion of respondents in whom data quality issues as described in the "Methods" section were identified had only trivial effects on the coefficients obtained in the DCE model. In the cTTO tobit model, the effects were slightly larger: the value for state 55555 decreased by about 0.02 utility points after exclusion of inconsistent responses and the values for the mild states were about 0.02 higher.

Reference Values for the Dutch General Population

Out of the 979 respondents in the respondent recruitment data, data were missing for the age and sex categories for 9 and 2 respondents, respectively. Table 5 presents the Dutch general population values by sex and age categories on the basis of the constraint tobit model.

Comparison of EQ-5D-3L and EQ-5D-5L Values

Figure 4A depicts the distribution of the 243 attainable values in the EQ-5D-3L and the 3125 attainable values in the EQ-5D-5L. This graph shows that the two instruments roughly cover the same evaluation space, but compared with the EQ-5D-5L, the EQ-5D-3L contains relatively few health states that result in values between 0.5 and 0.75. Figure 4B depicts the results of comparing the results of 3476 respondents filling out both the EQ-5D-5L and the EQ-5D-3L. In this sample, 118 of the 243 EQ-5D-3L states were reported versus 584 of the 3125 EQ-5D-5L states. The kernel density plot suggests that the EQ-5D-5L allows for more observations, and hence the potential to distinguish subgroups, in mild conditions (0.8–1) and in moderately severe conditions (0.3–0.5) using the Dutch tariff.

Discussion

The objective of this study was to derive a Dutch tariff for the EQ-5D-5L. The EQ-5D-5L is a new version of the EQ-5D, with an increased number of answer levels from three to five per dimension. Accordingly, the number of health states is increased from 243 to 3125. This study collected DCE and cTTO responses in a face-to-face setting among 1003 respondents (of which 979 were included for the final tariff) and modeled these results to estimate values for all health states of the EQ-5D-5L. A

Table 2 – Observed mean cTTO values and SDs (N = 979).

State	Mean ± SD	State	Mean ± SD
11112	0.907 ± 0.160	31524	0.283 ± 0.581
11121	0.915 ± 0.194	31525	0.310 ± 0.646
11122	0.927 ± 0.120	32314	0.428 ± 0.554
11211	0.925 ± 0.163	32443	0.147 ± 0.626
11212	0.886 ± 0.241	33253	0.270 ± 0.578
11221	0.859 ± 0.318	34155	−0.100 ± 0.637
11235	0.377 ± 0.608	34232	0.607 ± 0.498
11414	0.384 ± 0.572	34244	0.015 ± 0.644
11421	0.637 ± 0.438	34515	0.088 ± 0.646
11425	0.228 ± 0.597	35143	0.317 ± 0.545
12111	0.920 ± 0.197	35245	−0.048 ± 0.620
12112	0.837 ± 0.338	35311	0.648 ± 0.513
12121	0.847 ± 0.230	35332	0.577 ± 0.487
12244	0.229 ± 0.659	42115	0.314 ± 0.588
12334	0.374 ± 0.571	42321	0.692 ± 0.414
12344	0.204 ± 0.650	43315	0.209 ± 0.610
12513	0.625 ± 0.485	43514	0.153 ± 0.632
12514	0.385 ± 0.543	43542	0.065 ± 0.618
12543	0.115 ± 0.639	43555	−0.116 ± 0.669
13122	0.808 ± 0.273	44125	0.127 ± 0.650
13224	0.478 ± 0.595	44345	−0.178 ± 0.656
13313	0.754 ± 0.338	44553	−0.112 ± 0.686
14113	0.653 ± 0.430	45133	0.394 ± 0.571
14554	−0.144 ± 0.681	45144	0.026 ± 0.575
15151	0.204 ± 0.651	45233	0.326 ± 0.634
21111	0.928 ± 0.148	45413	0.345 ± 0.572
21112	0.879 ± 0.186	51152	0.200 ± 0.587
21315	0.401 ± 0.536	51451	0.089 ± 0.604
21334	0.410 ± 0.530	52215	0.187 ± 0.621
21345	0.097 ± 0.672	52335	0.122 ± 0.694
21444	0.128 ± 0.621	52431	0.468 ± 0.500
22434	0.305 ± 0.534	52455	−0.135 ± 0.662
23152	0.265 ± 0.649	53221	0.605 ± 0.520
23242	0.380 ± 0.554	53243	0.177 ± 0.630
23514	0.292 ± 0.611	53244	0.049 ± 0.605
24342	0.241 ± 0.575	53412	0.466 ± 0.537
24443	0.055 ± 0.598	54153	0.039 ± 0.648
24445	−0.143 ± 0.644	54231	0.532 ± 0.521
24553	0.120 ± 0.564	54342	0.063 ± 0.644
25122	0.598 ± 0.512	55225	0.082 ± 0.615
25222	0.602 ± 0.501	55233	0.271 ± 0.648
25331	0.597 ± 0.501	55424	−0.055 ± 0.663
31514	0.335 ± 0.570	55555	−0.309 ± 0.595

cTTO, composite time trade-off.

constrained random-effects tobit model was estimated on the cTTO data to derive the values. This model was favored over the OLS model on the basis of the nature of the data and the agreement with the DCE data. The largest impact of choosing the tobit model over the OLS model was in the pain dimension, with a difference of 0.029 in coefficients, whereas the average absolute difference between all parameters of the tobit and OLS models was 0.008. The coefficients obtained in the tobit model followed the monotonic structure of the EQ-5D-5L instrument, except for level 5 of the self-care and usual activities dimensions, which were constrained to be equal to level 4. The resulting tariff produced values that ranged from −0.446 for state 55555 to 1 for state 11111.

When compared with values produced by the EQ-5D-3L, the range of attainable values by the EQ-5D-5L is slightly larger. The increase in observations in the range between 0.8 and 1 and between 0.3 and 0.5 in 3476 patients who completed both the EQ-5D-3L and the EQ-5D-5L suggests that the new EQ-5D-5L is better equipped to discriminate between patients with minor complaints and moderate/severe complaints than its predecessor. The EQ-5D-5L aimed to improve the sensitivity to mild health problems and small changes in health status by expanding the number of answer categories from three to five. The valuation study reported here was able to assign different utility values

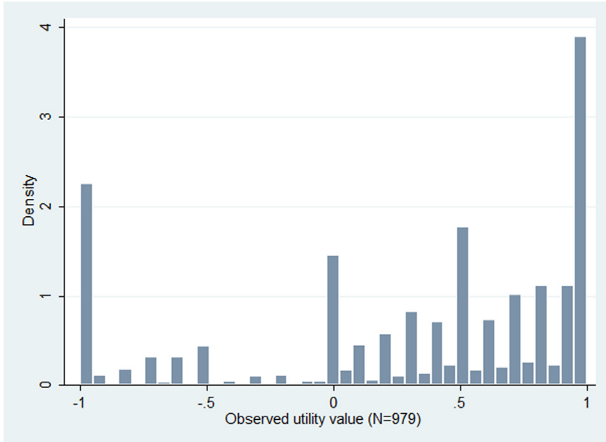

Fig. 2 – Observed utility values in the cTTO study. (Color version of figure available online).

Table 3 – DCE model.

EQ-5D	β	SE	P
mo2	−0.428	0.058	0.000
mo3	−0.411	0.076	0.000
mo4	−1.117	0.080	0.000
mo5	−1.282	0.088	0.000
sc2	−0.188	0.068	0.006
sc3	−0.296	0.073	0.000
sc4	−0.831	0.081	0.000
sc5	−0.857	0.078	0.000
ua2	−0.353	0.062	0.000
ua3	−0.209	0.072	0.004
ua4	−1.037	0.076	0.000
ua5	−1.093	0.082	0.000
pd2	−0.400	0.067	0.000
pd3	−0.349	0.070	0.000
pd4	−1.648	0.083	0.000
pd5	−2.372	0.102	0.000
ad2	−0.310	0.072	0.000
ad3	−0.572	0.073	0.000
ad4	−1.650	0.099	0.000
ad5	−2.292	0.111	0.000
#insignificant			0
#illogically ordered			3
Pseudo R^2			0.33
AIC			6455.7
BIC			6606.4

AIC, Akaike information criterion; ad, anxiety/depression; BIC, Bayesian information criterion; DCE, discrete choice experiment; EQ-5D, EuroQol five-dimensional questionnaire; mo, mobility; pd, pain/discomfort; sc, self-care; SE, standard error; ua, usual activities.

Table 4 – OLS and tobit models.

EQ-5D	Model 1: random-effect OLS			Model 2: random-effect tobit			Model 3: tobit with constraints	
	β	SE	P	*β*	SE	P	*β*	SE
mo2	−0.037	0.014	0.007	−0.032	0.016	0.041	−0.035	0.016
mo3	−0.061	0.015	0.000	−0.056	0.016	0.001	−0.057	0.016
mo4	−0.164	0.016	0.000	−0.166	0.018	0.000	−0.166	0.018
mo5	−0.192	0.015	0.000	−0.202	0.017	0.000	−0.203	0.016
sc2	−0.038	0.013	0.004	−0.039	0.016	0.012	−0.038	0.016
sc3	−0.064	0.015	0.000	−0.064	0.018	0.000	−0.061	0.017
sc4	−0.169	0.018	0.000	−0.180	0.018	0.000	−0.168	0.014
sc5	−0.150	0.015	0.000	−0.165	0.016	0.000	−0.168	0.014
ua2	−0.039	0.012	0.001	−0.040	0.016	0.013	−0.039	0.016
ua3	−0.085	0.016	0.000	−0.090	0.017	0.000	−0.087	0.017
ua4	−0.198	0.014	0.000	−0.207	0.017	0.000	−0.192	0.014
ua5	−0.167	0.014	0.000	−0.181	0.016	0.000	−0.192	0.014
pd2	−0.064	0.012	0.000	−0.064	0.015	0.000	−0.066	0.015
pd3	−0.087	0.014	0.000	−0.089	0.018	0.000	−0.092	0.018
pd4	−0.334	0.015	0.000	−0.353	0.016	0.000	−0.360	0.015
pd5	−0.390	0.017	0.000	−0.420	0.017	0.000	−0.415	0.017
ad2	−0.073	0.013	0.000	−0.073	0.017	0.000	−0.070	0.017
ad3	−0.146	0.016	0.000	−0.146	0.019	0.000	−0.145	0.019
ad4	−0.346	0.017	0.000	−0.360	0.017	0.000	−0.356	0.017
ad5	−0.401	0.017	0.000	−0.425	0.016	0.000	−0.421	0.016
Constant	0.953	0.012	0.000	0.956	0.022	0.000	0.953	0.022
#insignificant	0			0			0	
#illogically	2			2			0	
MAE	0.044			0.053			0.053	
DCE fit	2.851			2.824			2.825	

ad, anxiety/depression; DCE, discrete choice experiment; EQ-5D, EuroQol five-dimensional questionnaire; MAE, mean absolute error; mo, mobility; OLS, ordinary last squares; pd, pain/discomfort; sc, self-care; SE, standard error; ua, usual activities.

to the mild health states, indicating that respondents on average distinguished between these states. Compared with the EQ-5D-3L, a smaller gap was observed between perfect health (state 11111) and the next best state with tariff value 0.897 in the EQ-5D-3L (for state 11211, "some problems" in usual activity) and 0.918 in the EQ-5D-5L (for state 21111, "slight problems" in mobility). Although the new Dutch tariff is able to assign different utility values to different mild health states, the sensitivity of the instrument also depends on how patients respond to the Dutch descriptive system.

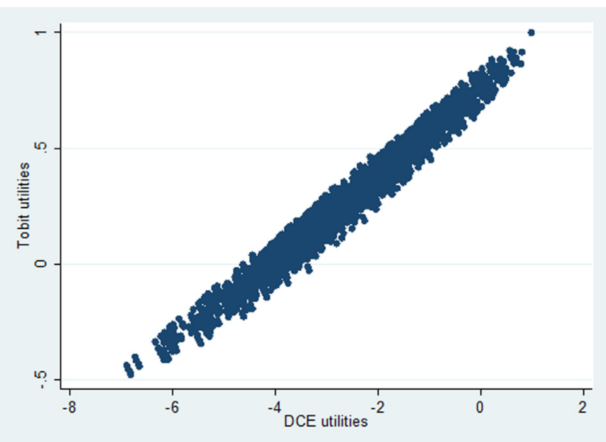

Fig. 3 – Comparison of DCE and Tobit derived utilities. (Color version of figure available online).

Similarity in values between the EQ-5D-3L and EQ-5D-5L versions was not necessarily expected because the wording of the instruments differs. For example, level 2 in the EQ-5D-3L was described as having "some problems" on a given dimension, whereas level 3 in the EQ-5D-5L was described as having "moderate problems." The worst mobility dimension was described as "confined to bed" in the level 3 version versus "unable to walk" in the level 5 version. The EQ-5D-3L and

Table 5 – Dutch general population EQ-5D-5L reference values.

Characteristics	Mean ± SD	Min.	Max.	N
Age (y)				
<20	0.958 ± 0.07	0.743	1	26
20 through	0.908 ± 0.146	0.031	1	158
30 through	0.903 ± 0.134	0.141	1	134
40 through	0.85 ± 0.196	−0.16	1	202
50 through	0.857 ± 0.183	−0.137	1	186
60 through	0.839 ± 0.179	−0.003	1	158
70 and high	0.852 ± 0.148	0.335	1	106
Sex				
Men	0.881 ± 0.172	−0.012	1	480
Women	0.858 ± 0.168	−0.16	1	497
Average	0.869 ± 0.170	−0.16	1	979

EQ-5D-5L, EuroQol five-dimensional questionnaire five-level; Max., maximum; Min., minimum.

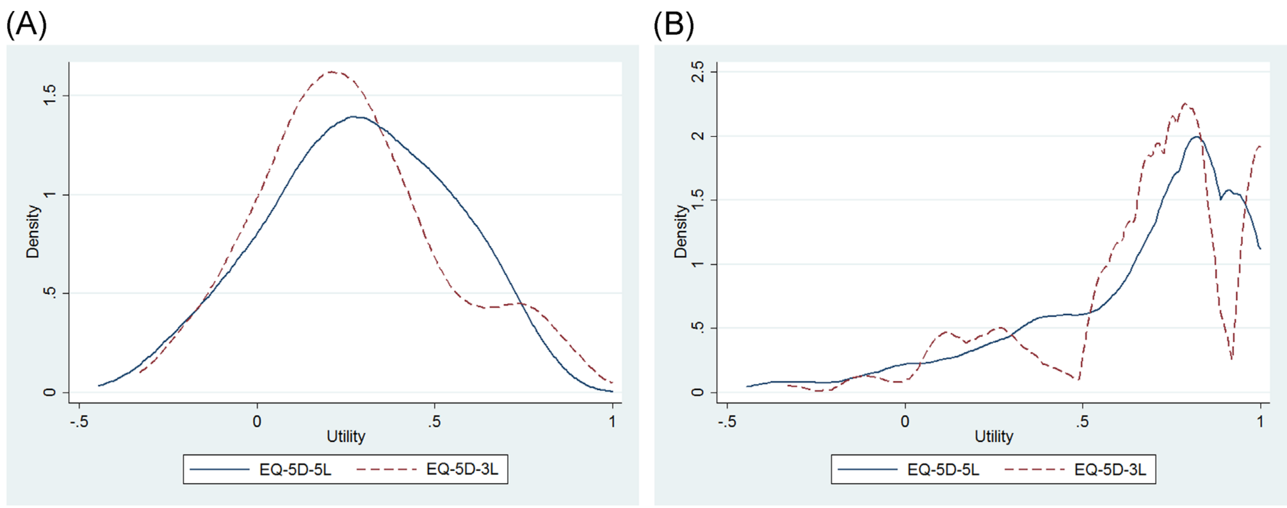

Fig. 4 – (A) Kernel density plot of all possible EQ-5D-3L and EQ-5D-5L values. (B) Kernel density plot of utility values of respondents who filled-out both EQ-5D-5L and EQ-5D-3L (N=3476). (Color version of figure available online).

EQ-5D-5L version values for the worst health state also differed: −0.329 (for EQ-5D-3L state 33333) and −0.446 (for EQ-5D-5L state 55555). Moreover, the tasks for valuing worse-than-dead states were conceptually different.

This study differs from several existing EQ-5D-3L tariffs because it does not correct for diminishing marginal utility. Many published tariffs for the EQ-5D-3L, including the Dutch one [25], include interaction terms that capture that having (severe) problems in multiple domains may lessen the impact of additional health deterioration. Examples are the n3 term used among others by Dolan [27] or the D1 term by Shaw et al. [28]. Models with these interaction terms fitted the observed TTO data better in those studies. This study refrained from including such interaction terms in the final model for two reasons. First, the design and sample size were optimized for main effects. Second, including interaction terms may result in a better fit to the observed data, but increases the risk of misprediction of values for health that were not included in the valuation study as demonstrated for the EQ-5D-3L elsewhere [29]. Because it was decided not to include interaction terms in the model, alternative strategies were explored to deal with logical inconsistencies in the parameter estimates. Both the cTTO and DCE models contained several parameter estimates that were illogically ordered given the monotonic structure of the EQ-5D-5L, albeit in different locations. In the final cTTO model, the affected parameters (self-care and usual activities level 5) were subjected to the constraint that a level 5 had to receive the same coefficient as a level 4. Because the observed differences between levels 4 and 5 in the affected dimensions were small and not statistically significant both in the cTTO and in the DCE, the loss of information following the constraints was very limited.

A limitation inherent to the cTTO task is the censoring of data. In this study, the DCE data were used as a yardstick to decide whether to apply the OLS or the tobit model on the cTTO data. Strong agreement was observed between the cTTO and the DCE model parameters in general, supporting the hypothesis that the two methods measure the same construct. To improve understanding of the validity conditions for models that integrate DCE and TTO data (referred to as hybrid models), further investigation of the agreement between the methods is warranted, for example, with regard to their ability to study preference heterogeneity.

The present study was performed in the “first wave” of several EQ-5D-5L valuation studies, and is hence among the first to apply the newly developed protocol. Research into the data from these first-wave studies identified some general issues: clustering of values at certain round numbers and inconsistencies [30]. For example, in each of these studies at least 20% of respondents valued one or more health states as being worse than 55555. Similarly, about 10% of respondents gave lower values to very mild health states than they did to more severe and logically worse states. There may be several explanations for these inconsistencies in responses. The clustering of values could correspond to respondents stating indifference early in the present TTO routing (“I would trade about 5 years”), whereas if properly motivated they may have given a more precise response (“I would trade exactly 4 years”) to follow-up questions. Furthermore, interviewers could differ in their engagement and ability to support respondents in performing the task well.

In the period between data collection and the submission of this article, a research program was developed aimed at testing proposed modifications to the EQ-VT that might increase data quality [30]. On the basis of that research, a new version of the EQ-VT has been developed that includes additional warm-up questions to the cTTO task, introduces pop-ups asking for confirmation of a response, and confronts respondents with their own values/inconsistencies, with the opportunity to make changes. Although these improvements are welcome, inconsistent respondents identified in the Dutch first-wave study showed only minor effects in the sensitivity analysis of this study. Therefore, it seems that the tariff established here represents Dutch views on health state severity levels well. Further research is required to investigate how these Dutch preferences are affected by, for example, income or experience with illness, and the data collected for this study can be used to address such questions.

Conclusions

This study established a Dutch tariff for the EQ-5D-5L on the basis of cTTO. The values represent the views of the Dutch population about the EQ-5D-5L health states. This value set may be used to compute utilities for use in calculating quality-adjusted life-years for Dutch health technology assessments and economic evaluations. Additional research is required to assess the responsiveness over time and the discriminative properties of the Dutch tariff in patient samples.

Source of financial support: This research was funded by the Netherlands Organisation for Health Research and Development (ZonMW; grant no. 152002044) and a grant from the EuroQoL Research Foundation.

REFERENCES

[1] Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20:1727–36.
[2] National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013. Available from: ⟨http://publications.nice.org.uk/pmg9⟩. [Accessed June 11, 2015].
[3] Brazier JE, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ 2004;13:873–84.
[4] Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. Value Health 2006;9:262–71.
[5] Versteegh M, Leunis A, Uyl-de Groot C, Stolk E. Condition-specific preference-based measures: benefit or burden? Value Health 2012;15:504–13.
[6] Janssen M, Birnie E, Haagsma J, Bonsel G. Comparing the standard EQ-5D three-level system with a five-level version. Value Health 2008;11:275–84.
[7] Janssen M, Birnie E, Bonsel G. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. Qual Life Res 2008;17:463–73.
[8] Pickard A, De Leon M, Kohlmann T, et al. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Med Care 2007;45:259–63.
[9] Pickard AS, Kohlmann T, Janssen MF, et al. Evaluating equivalency between response systems: application of the Rasch model to a 3-level and 5-level EQ-5D. Med Care 2007;45:812–9.
[10] Knies S, Evers S, Candel M, et al. Utilities of the EQ-5D: transferable or not? Pharmacoeconomics 2009;27:767–79.
[11] Oppe M, Devlin NJ, van Hout B, et al. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. Value Health 2014;17:445–53.
[12] Devlin N, Krabbe PM. The development of new research methods for the valuation of EQ-5D-5L. Eur J Health Econ 2013;14:1–3.
[13] Shah K, Lloyd A, Oppe M, Devlin N. One-to-one versus group setting for conducting computer-assisted TTO studies: findings from pilot studies in England and the Netherlands. Eur J Health Econ 2013;14:65–73.
[14] Luo N, Li M, Stolk E, Devlin N. The effects of lead time and visual aids in TTO valuation: a study of the EQ-VT framework. Eur J Health Econ 2013;14:15–24.
[15] Versteegh M, Attema A, Oppe M, et al. Time to tweak the TTO: results from a comparison of alternative specifications of the TTO. Eur J Health Econ 2013;14:43–51.
[16] Ramos-Goni J, Rivero-Arias O, Errea M, et al. Dealing with the health state "dead" when using discrete choice experiments to obtain values for EQ-5D-5L health states. Eur J Health Econ 2013;14:33–42.
[17] Janssen M, Oppe M, Versteegh M, Stolk E. Introducing the composite time trade-off: a test of feasibility and face validity. Eur J Health Econ 2013;14:5–13.
[18] Attema A, Edelaar-Peeters Y, Versteegh M, Stolk E. Time trade-off: one methodology, different methods. Eur J Health Econ 2013;14:53–64.
[19] Augustovski F, Rey-Ares L, Irazola V, et al. Lead versus lag-time trade-off variants: does it make any difference? Eur J Health Econ 2013;14:25–31.
[20] Stolk E, Oppe M, Scalone L, Krabbe P. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. Value Health 2010;13:1005–13.
[21] Krabbe P, Devlin N, Stolk E, et al. Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-state values. Med Care 2014;52:935–43.
[22] Brazier J. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford, UK: Oxford University Press, 2007.
[23] Tilling C, Devlin N, Tsuchiya A, Buckingham K. Protocols for time tradeoff valuations of health states worse than dead: a literature review. Med Decis Making 2010;30:610–9.
[24] Oppe M, van Hout B. The optimal hybrid: experimental design and modeling of a combination of TTO and DCE. Presented at: 27th Scientific Plenary Meeting of the EuroQol Group-Proceedings. EuroQol Group Executive Office, Rotterdam, the Netherlands, 2010: [Accessed June 11, 2015].
[25] Lamers LM, McDonnell J, Stalmeier PFM, et al. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ 2006;15:1121–32.
[26] Janssen M, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res 2013;22:1717–27.
[27] Dolan P. Modeling valuations for the EuroQol health states. Med Care 1997;35:1095–108.
[28] Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005;43:203–20.
[29] Bonsel GJ, Oppe M, Janssen MF. Unexpected large misspecification effects of health profile selection and interaction analysis to obtain a value function from unsaturated valuation datasets, using the standard EuroQol approach. Presented at: EuroQoL Proceedings Plenary Meeting. Stockholm, Sweden, September 25-26, 2014.
[30] Shah K, Rand-Hendriksen K, Ramos-Goni J, et al. Improving the quality of data collected in EQ-5D-5L valuation studies: a summary of the EQ-VT research methodology programme. Presented at: 31st EuroQol Scientific Plenary Meeting. Stockholm, Sweden, September 25-26, 2014.