VALUE IN HEALTH 19 (2016) 648–654

Available online at www.sciencedirect.com

ScienceDirect

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

Preference-Based Assessments

Comparison of Value Set Based on DCE and/or TTO Data: Scoring for EQ-5D-5L Health States in Japan

Takeru Shiroiwa, MPH, PhD[1,], Shunya Ikeda, PhD, MD[2], Shinichi Noto, PhD[3], Ataru Igarashi, PhD[4], Takashi Fukuda, PhD[1], Shinya Saito, PhD, MD[5], Kojiro Shimozuma, PhD, MD[6]*

[1]*Department of Health and Welfare Services, National Institute of Public Health, Wako, Japan;* [2]*School of Pharmacy, International University of Health and Welfare, Otawara, Japan;* [3]*Department of Health Sciences, Niigata University of Health and Welfare, Niigata, Japan;* [4]*Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo, Japan;* [5]*Graduate School of Health Sciences, Okayama University, Okayama, Japan;* [6]*Department of Biomedical Sciences, College of Life Sciences, Ritsumeikan University, Kusatsu, Japan*

ABSTRACT

Background: The valuation study of the five-level version of the EuroQol five-dimensional questionnaire (EQ-5D-5L) involved composite time trade-off (cTTO) and a discrete choice experiment (DCE). The DCE scores must be anchored to the quality-of-life scale from 0 (death) to 1 (full health). Nevertheless, the characteristics of the statistical methods used for converting the EQ-5D-5L DCE results by using TTO information are not yet clearly known. **Objectives:** To present the Japanese DCE value set of the EQ-5D-5L and compare three methods for converting latent DCE values. **Methods:** The survey sampled the general population at five locations in Japan. 1098 respondents were stratified by age and sex. To obtain and compare the value sets of the EQ-5D-5L, the cTTO and DCE data were analyzed by a linear mixed model and conditional logit, respectively. The DCE scores were converted to the quality-of-life scale by anchoring to the worst state using cTTO, mapping DCE onto cTTO, and a hybrid model. **Results:** The data from 1026 respondents were analyzed. All the coefficients in the cTTO and DCE value sets were consistent throughout all the analyses. Compared with the cTTO algorithm, the mapping and hybrid methods yielded very similar scoring coefficients. The hybrid model results, however, produced a lower root mean square error and fewer health states with errors exceeding 0.05 than did the other models. The DCE anchored to the worst state overestimated the cTTO scores of almost all the health states. **Conclusions:** Japanese value sets based on DCE were demonstrated. On comparing the observed cTTO scores, we found that the hybrid model was slightly superior to the simpler methods, including the TTO model.

Keywords: discrete choice experiment (DCE), EQ-5D-5L, preference, time trade-off (TTO).

Introduction

Quality-adjusted life-years (QALYs) are calculated from quality-of-life (QOL) scores obtained by a preference-based measure such as the EuroQol five-dimensional questionnaire (EQ-5D) [1]. In QALY calculation, QOL needs to be measured on a standardized scale anchored to 0 (death) and 1 (full health). In a preference-based measure, the responses are based on patients' own evaluations of their health states. The QOL score is calculated with a predetermined value set that reflects the preferences of the general population.

The Japanese version of the three-level EQ-5D (EQ-5D-3L) was completed in 1997 and was certified as an official version by the EuroQol Group [2]. It is based on an interview survey conducted in the general population using the time trade-off (TTO) method [3]. The EQ-5D-3L is the most popular measure for calculating QALY [4]. It, however, distinguishes only three levels for each dimension. Therefore, it is thought to lack sufficient sensitivity; moreover, it may be skewed by the ceiling effect, in which respondents tend to choose the first level of each item even if their health state is imperfect [5]. To resolve these problems, the EuroQol Group has increased the number of levels for each dimension from three to five.

During the development of the five-level version of the EQ-5D (EQ-5D-5L), linguistic expressions were somewhat modified to improve consistency and clarify their meaning. For example, in the mobility dimension, the term "I am confined to bed" was changed to "I am unable to walk about," which indicates a better

* Address correspondence to: Takeru Shiroiwa, Department of Health and Welfare Services, National Institute of Public Health, 2-3-6 Minami, Wako, Saitama 351-0197, Japan.

E-mail: t.shiroiwa@gmail.com.

http://dx.doi.org/10.1016/j.jval.2016.03.1834

health state in Japanese. In addition, some unchanged terms in the English version were altered in the Japanese-translated version after consultation with the EuroQol Group. In the self-care dimension, the term "wash myself" was translated as "wash my face" in the Japanese EQ-5D-3L, but was alternatively translated as "wash my body" in the Japanese EQ-5D-5L. Therefore, the QOL score in the EQ-5D-5L cannot simply be derived from the EQ-5D-3L by a cross-walking or mapping algorithm [5].

To create a scoring algorithm for the Japanese version of the EQ-5D-5L, we conducted a survey using both composite TTO (cTTO) [6] and discrete choice experiments (DCEs) [7–9] following the procedure determined by the EuroQOL Group. Although we have already reported a Japanese scoring algorithm of the EQ-5D-5L based on cTTO [10], the DCE-based scoring method must also be considered. In a DCE, the respondents are asked to select their preferred health state from two alternatives, providing a simpler and more understandable assessment than the cTTO. The DCE provides only ordinal preference information, whereas the TTO results in a cardinal score. Therefore, before building a scoring algorithm for the EQ-5D-5L, the DCE scores must be anchored to the QOL scale from 0 (death) to 1 (full health). In some studies (e.g., Viney et al. [11]), "death" as an alternative and "duration of life" as an attribute of the DCE card were included to obtain QOL scale coefficients. The EQ-5D-5L valuation study, however, applied another method in which the DCE coefficients were converted using cTTO information. Rowen et al. [12] suggested some methods for the conversion; these methods use the results of a TTO or a standard gamble (SG) in addition to a DCE, although few studies have reported the application of such methods to the EQ-5D-5L data for the construction of a value set. The characteristics of these methods remain uncertain, especially for the EQ-5D-5L. The objective of this study was to 1) estimate the DCE value set based on the Japanese EQ-5D-5L valuation study and 2) compare the value set based on TTO and three methods for converting latent DCE scores to the QOL scale.

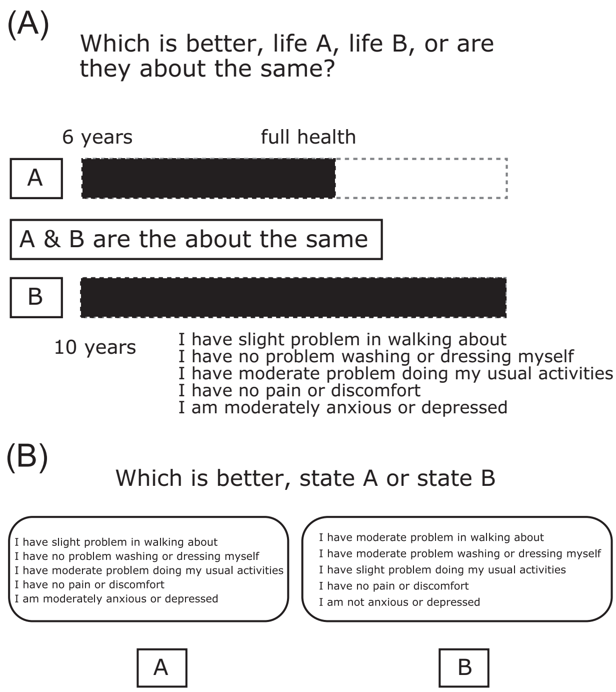

Fig. 1 – (A) Example of cTTO task. (B) Example of DCE task. cTTO, composite time trade-off; DCE, discrete choice experiment.

Methods

Subjects

The computer-based survey was conducted in the general population at five locations in Japan (Tokyo, Okayama, Nagoya, Osaka, and Niigata). Respondents were recruited by a research company (ANTERIO Inc.) that sampled 1098 respondents throughout the country (i.e., approximately 200 respondents at each location). The sample number was not determined on the basis of any rigid statistical consideration. Respondents were stratified by sex and age group in each location to collect the same number in each cell. At their local survey centers, they were interviewed in a computer-based (i.e., in a one-investigator, one-respondent) setting [13] over intervals that were between 30 and 60 minutes.

Before administering the survey, all the investigators at each location received training for approximately half a day. The authors, who had been directly trained by the EuroQol Group, also participated in this training. The trained investigators explained the survey procedures to the respondents following the survey manual. To ensure the quality and consistency of the investigators, the number of investigators was limited to approximately five at each location.

Study Methods

Each respondent was subjected to a cTTO study, followed by a DCE survey. In the cTTO phase (Fig. 1A), the "worse than dead (WTD)" health states were evaluated differently from the "better than dead" health states. A previous EQ-5D valuation study assessed the WTD health state by a Measurement and Valuation of Health (MVH) approach [14,15]. This approach, however, has been criticized because it admits TTO scores below −1. Therefore, to measure the WTD health state [16,17], the cTTO uses lead-time TTO [18–20], which limits the minimum TTO score to −1.

In the cTTO phase, 86 health states were selected from 3125 (=5^5) health states described by the EQ-5D-5L. Each respondent was asked to value 10 of these 86 health states. Before the cTTO survey, respondents were shown 11 health state cards, 10 for the presented health states and the remaining 1 for "death." To facilitate their understanding, the participants were asked to order the cards from better to worse health status. Next, the cTTO was conducted using a program (EQ valuation technology [EQ-VT]) created by the EuroQol Group. Together with the EuroQol Group, we had already translated the description, the manual, and other EQ-VT documentation from its original English version into Japanese. The respondents selected their answers without interference by the investigators.

From the 86 health states, 10 blocks (each containing 10 health states) were constructed and randomly allocated one block to each respondent. The 10 health states within a block were presented in random order. The state "55555" was included in all blocks and 5 health states were contained in two blocks; the other 80 health states appeared in only one block. Having completed the cTTO task, the participants started the DCE survey, in which they were requested to choose a preferred health state from a pair of options (Fig. 1B). This task was repeated 7 times for each respondent. The respondents were randomly allocated 1 of the 28 blocks, where each block included 7 pairs of health states (196 pairs in total), on the basis of a Bayesian efficient design [21]. Some health states shown in the DCE task were not the same as those presented in the cTTO task, and others were the same.

All these questions were presented on a personal computer preinstalled with the EQ-VT, and questions were automatically displayed by the algorithm until the score is determined. In the cTTO phase, when the respondent was asked about the period of life covered by a health state, the response was facilitated by a visual representation of the question (a bar graduated from 0 to 10 on which respondents marked a particular number of years) [22]. In the DCE phase, respondents were simultaneously presented with two boxed health states on the computer screen, and were requested to choose health state A or B.

Statistical Analysis

All answers were electronically recorded and posted online to the EuroQol Group Data Center. The quality control process of the EuroQol Group then checked the sent data, assessing the time of the investigator's explanation, time taken to answer the questions, the answer patterns (e.g., whether a participant had selected the same choice for all questions), and other quality factors. All or some of the answers obtained by "ineligible" investigators were excluded by the EuroQol Group, and the remaining answers were analyzed.

cTTO data

When respondents equated 10 years of life with the better than dead health state to x years of life with full health, the QOL score was calculated as x/10. Conversely, when y years of life with full health was equated to "life with full health for 10 years followed by life with a WTD health state for 10 years," the QOL score was calculated as $y/10 - 1$. The data were analyzed by a linear mixed model with "1 - QOL score" (disutility) as the response variable. To account for the intrarespondent correlation [23], a constant term and dummy variables representing the levels of the five dimensions ($5 \times [5 - 1] = 20$) were treated as fixed effects and the respondents were treated as random effects. The linear mixed model is expressed as $\mathbf{1} - \mathbf{Q} = \mathbf{X}\boldsymbol{\beta} + \varepsilon$, where $\mathbf{Q}$ is a vector of cTTO scores, $\mathbf{X}$ is a matrix of dummy variables, and ε is the error term.

DCE data

The DCE data were analyzed by a conditional logit model, using the same 20 dummy variables as the cTTO model. This analysis extracts the latent coefficients for EQ-5D-5L scoring. The DCE "dis-score," defined as the sum of latent DCE coefficients for each health state, must then be converted to the QOL scale. Rowen et al. [12] suggested four methods for converting DCE scores to the scale anchored to 0 (death) and 1 (full health): model 0, anchoring by the coefficient for "dead" [7,11,24]; model 1, anchoring to the worst cTTO state; model 2, mapping the DCE onto cTTO [25]; and model 3, a hybrid model [26]. Nevertheless, anchoring by the coefficient for "dead" was inapplicable to our data because "death" was not one of the presented health states.

Model 1 (anchoring to the worst cTTO state) multiplies the DCE dis-score by a constant $\gamma = (1 - \text{mean worst score of cTTO [55555]})/\text{worst DCE dis-score [55555]}$. In the present study, the worst DCE dis-score was estimated by the latent coefficients in the conditional logit model. DCE-based QOL scores can be calculated as $1 - \gamma DCE_i$, where DCE_i is the predicted DCE dis-score in the ith health state. The converted DCE-based QOL scores range from the mean worst cTTO score (−0.019 in our survey) to 1.

Model 2 (the mapping method) first calculates the mean cTTO scores of the 86 health states presented to respondents and then calculates the DCE dis-scores of the same 86 states. Mapping predicts the QOL scores by statistically relating DCE dis-score to cTTO disutility. To obtain the QOL score, the DCE dis-scores were converted by a mapping function f(•), namely, $1 - TTO_i = f(DCE_i) + \varepsilon_i$, where TTO_i is the observed mean cTTO score and DCEi is the predicted DCE dis-score for the ith health state ($1 \leq i \leq 86$). The mapping was implemented by a simple linear function $f(x) = ax + b$. In this case, the coefficient of each attribute was calculated by multiplying the latent DCE coefficients by the estimated parameter a, and the intercept was −b.

Model 3 (the hybrid model) is a novel method that simultaneously analyzes both cTTO and DCE data. This method adopts two statistical models: the likelihood-based approach and the

Table 1 – Basic characteristics of the analysis set.

Characteristic	N	%
Sex		
Male	515	50.2
Female	511	49.8
Age (y)		
20s	203	19.8
30s	204	19.9
40s	206	20.1
50s	206	20.1
≥60s	207	20.2
Location		
Tokyo	194	18.9
Okayama	147	14.3
Nagoya	210	20.5
Osaka	235	22.9
Niigata	240	23.4
Marital status		
Single	285	27.8
Married	655	63.8
Widowed	56	5.5
Divorced	30	2.9
Educational background		
Junior high school	33	3.2
High school	415	40.4
Vocational school	131	12.8
Junior college	116	11.3
University	318	31
Graduate school	13	1.3
Type of employment		
Full-time	387	37.7
Full-time (nonregular)	70	6.8
Part-time	181	17.6
Self-employment	100	9.7
Housewife	166	16.2
Retired	50	4.9
Student	70	6.8
Others	2	0.2
Personal income (expressed in ¥1 million)		
<1	310	30.2
1–<2	138	13.5
2–<4	266	25.9
4–<6	187	18.2
6–<10	100	9.7
10–<15	19	1.9
15–<20	3	0.3
≥20	3	0.3
Household income (expressed in ¥1 million)		
<1	34	3.3
1–<2	38	3.7
2–<4	175	17.1
4–<6	283	27.6
6–<10	318	31
10–<15	134	13.1
15–<20	27	2.6
≥20	17	1.7

¥, Japanese yen.

Table 2 – Estimates of TTO- and DCE-based analyses.

Item	Level	TTO	Latent DCE coefficient	DCE model 1 (anchored to worst state)	DCE model 2 (mapping)	DCE model 3 (hybrid)
Intercept		−0.0609	–	–	−0.0943	−0.0616
MO						
	2	−0.0639	−0.6109	−0.0761	−0.0745	−0.0654
	3	−0.1126	−0.7900	−0.0984	−0.0963	−0.1125
	4	−0.1790	−1.2833	−0.1598	−0.1564	−0.1792
	5	−0.2429	−1.9778	−0.2463	−0.2411	−0.2397
SC						
	2	−0.0436	−0.3729	−0.0464	−0.0455	−0.0380
	3	−0.0767	−0.4073	−0.0507	−0.0497	−0.0701
	4	−0.1243	−0.8480	−0.1056	−0.1034	−0.1175
	5	−0.1597	−1.2742	−0.1587	−0.1553	−0.1606
UA						
	2	−0.0504	−0.4713	−0.0587	−0.0575	−0.0572
	3	−0.0911	−0.5081	−0.0633	−0.0619	−0.0918
	4	−0.1479	−1.0550	−0.1314	−0.1286	−0.1551
	5	−0.1748	−1.4974	−0.1865	−0.1825	−0.1729
PD						
	2	−0.0445	−0.3341	−0.0416	−0.0407	−0.0406
	3	−0.0682	−0.5361	−0.0668	−0.0654	−0.0680
	4	−0.1314	−1.0759	−0.1340	−0.1312	−0.1240
	5	−0.1912	−1.6889	−0.2103	−0.2059	−0.1930
AD						
	2	−0.0718	−0.3382	−0.0421	−0.0412	−0.0781
	3	−0.1105	−0.6626	−0.0825	−0.0808	−0.1111
	4	−0.1682	−1.1474	−0.1429	−0.1399	−0.1730
	5	−0.1960	−1.7417	−0.2169	−0.2123	−0.1968

Note. All the coefficients are significantly lower than 0 at a 5% significance level.
AD, anxiety/depression; DCE, discrete choice experiment; MO, mobility; PD, pain/discomfort; SC, self-care; TTO, time trade-off; UA, usual activities.

Bayesian approach [12,26]. Here, we used the Bayesian approach, which is more flexible for modeling random effects, applying a noninformative distribution as the prior distribution. The disutility of cTTO is assumed to be linearly related to the DCE disscore; that is, $\boldsymbol{\beta}_{TTO} = c \bullet \boldsymbol{\beta}_{DCE}$, where $\boldsymbol{\beta}_{TTO}$ is a vector of cTTO coefficients, $\boldsymbol{\beta}_{TTO}$ is a vector of latent DCE coefficients, and c is a constant. Our Bayesian approach was almost identical to that of Rowen et al. [12]. Nevertheless, although Rowen et al. used simple logistic regression, we analyzed the DCE part by using a conditional logit model. The numerical calculation was performed by WinBUGs, setting the number of iterations and the burn-in sample to 33,000 and 3,000, respectively, although samples only from every third iteration were used.

To compare the single cTTO-based and the three DCE-based scoring methods, we calculated the summary statistics of all 3125 health states and the root mean square error between the observed mean cTTO scores and the predicted scores. For each model, we counted the number of health states in which the predicted and observed cTTO scores differed by more than 0.05 and 0.1. In addition, we compared the kernel density functions of the 3125 health states on the basis of the four scoring methods and the 243 states of the EQ-5D-3L. Statistical analyses were performed by SAS 9.4 (SAS Institute Inc., Cary, NC), R 3.1.0, and WinBUGs.

Results

Characteristics of the Analysis Set

The 1098 respondents were interviewed by 35 investigators from March to June 2014 in five Japanese cities (Tokyo, Okayama, Nagoya, Osaka, and Niigata, surveyed in that order). The quality control process of the EuroQol Group, however, revealed that three of the investigators (one in Tokyo and two in Okayama) did not follow the survey manual procedures when interviewing 40% or more of their respondents (taking insufficient time to explain

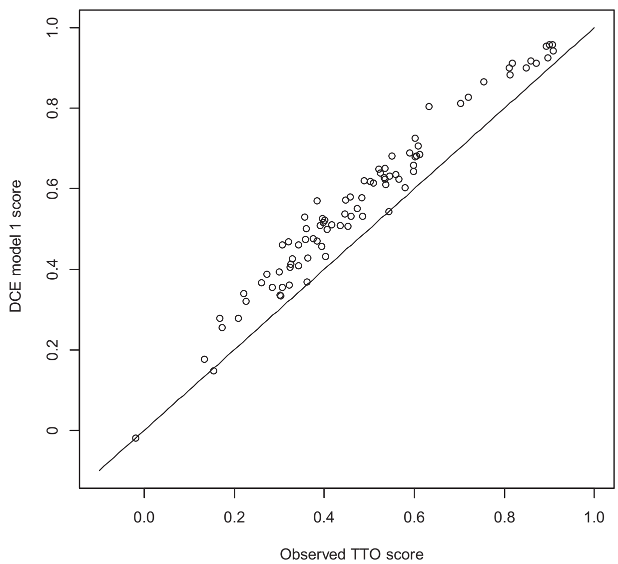

Fig. 2 – Relation between observed TTO and the score predicted by the DCE model 1. DCE, discrete choice experiment; TTO, time trade-off.

Table 3 – Characteristics of scores predicted by each model.

Score	TTO	DCE model 1 (anchored to worst state)	DCE model 2 (mapping)	DCE model 3 (hybrid)	EQ-5D-3L
Lowest score	−0.025	−0.019	−0.091	−0.025	−0.106
Median score	0.450	0.539	0.455	0.450	0.464
Second highest score	0.895	0.958	0.865	0.900	0.812
No. of negative scores	1	1	6	3	6
Inconsistency	0	0	0	0	
RMSE	0.0243	0.0852	0.0314	0.0237	
Maximum of predicted score (observed TTO)	0.0878	0.1860	0.1009	0.0864	
No. of predicted scores (observed TTO > 0.05)	9	70	16	7	
No. of predicted scores (observed TTO > 0.1)	0	30	1	0	

DCE, discrete choice experiment; EQ-5D-3L, three-level EuroQol five-dimensional questionnaire; RMSE, root mean square error; TTO, time trade-off.

the survey). All or some of the answers obtained by these investigators (involving 72 respondents) were excluded from the analysis.

The analysis set included data from 1026 respondents: 194 from Tokyo, 147 from Okayama, 210 from Nagoya, 235 from Osaka, and 240 from Niigata. Table 1 presents the basic characteristics of the respondents (sex, age, location, marital status, educational background, type of employment, personal income, and annual household income). In 2012, the actual Japanese median household income was ¥4.3 million (¥ = Japanese yen), whereas the average was ¥5.4 million. Married and unmarried people accounted for 61.1% and 22.8% of the population, respectively. Overall, 19.1% had graduated from university. Note that this statistic reflects the actual distribution of the Japanese population, but we sampled the same number of respondents from each age category. The background of the respondents, however, was comparable with that of the general population.

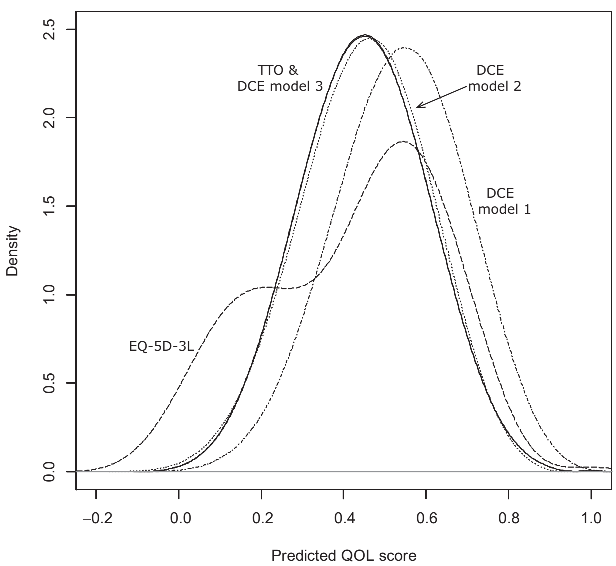

Fig. 3 – Density estimation of the QOL score obtained by each model. DCE, discrete choice experiment; EQ-5D-3L, three-level EuroQol five-dimensional questionnaire; QOL, quality of life; TTO, time trade-off.

Modeling Results from Four Different Analyses

The 1,026 respondents collectively yielded 14,364 cTTO data and 7,182 DCE data. Table 2 presents the coefficients estimated by the four models based on cTTO and DCE data. The coefficients of cTTO, model 2, and model 3 were quite similar. In the DCE analysis by model 1, the constant γ was estimated as 0.1245. The estimated mapping function f(x) in the DCE analysis by model 2 was $0.1219x + 0.09425$. Model 1 overestimated the cTTO scores of almost all health states. Figure 2 plots the relationship between the DCE scores obtained by model 1 and the observed cTTO scores.

"Inconsistency" was not observed in any of the analyses; that is, the coefficients of low health levels were not greater than those of high health levels. Table 3 presents the characteristics of the predicted scores. Less than 0.2% of the health states in the whole EQ-5D-5L model yielded negative scores, although 2.5% (6 of 243) of the health states were predicted as WTD in the EQ-5D-3L. The hybrid model exhibited a lower root mean square error and fewer health states with error exceeding 0.05 than did the other models, including the simple cTTO model, although the differences were quite small. Figure 3 shows the kernel density of the 3125 health states (243 states in the EQ-5D-3L). The probability density functions of the cTTO and the DCE model 3 (hybrid) overlap, and the density functions of the cTTO and the DCE model 2 (mapping) are almost the same. Because the mapping and hybrid analyses of the 3125 health states were highly correlated (correlation coefficient = 0.98), both methods yielded similar value sets. The density function of the EQ-5D-5L differs from that of the EQ-5D-3L by being monomodal and lower at poor health states. Figure 4 plots the mean error between the predicted and the observed cTTO scores.

Discussion

Both TTO and DCE are standard methods for measuring preference health status (described as the EQ-5D-5L system in this survey); nevertheless, TTO provides cardinal information, whereas DCE yields ordinal information. We compared various value sets for the EQ-5D-5L based on DCE and/or cTTO data. Data collected by computer-based surveying in Japan were processed with four types of statistical methods. Inconsistency was not observed in any of the analyses. Limitations of the conditional logit model to estimate DCE coefficients should however be considered. Gu et al. [27] pointed out that 1) the coefficients are

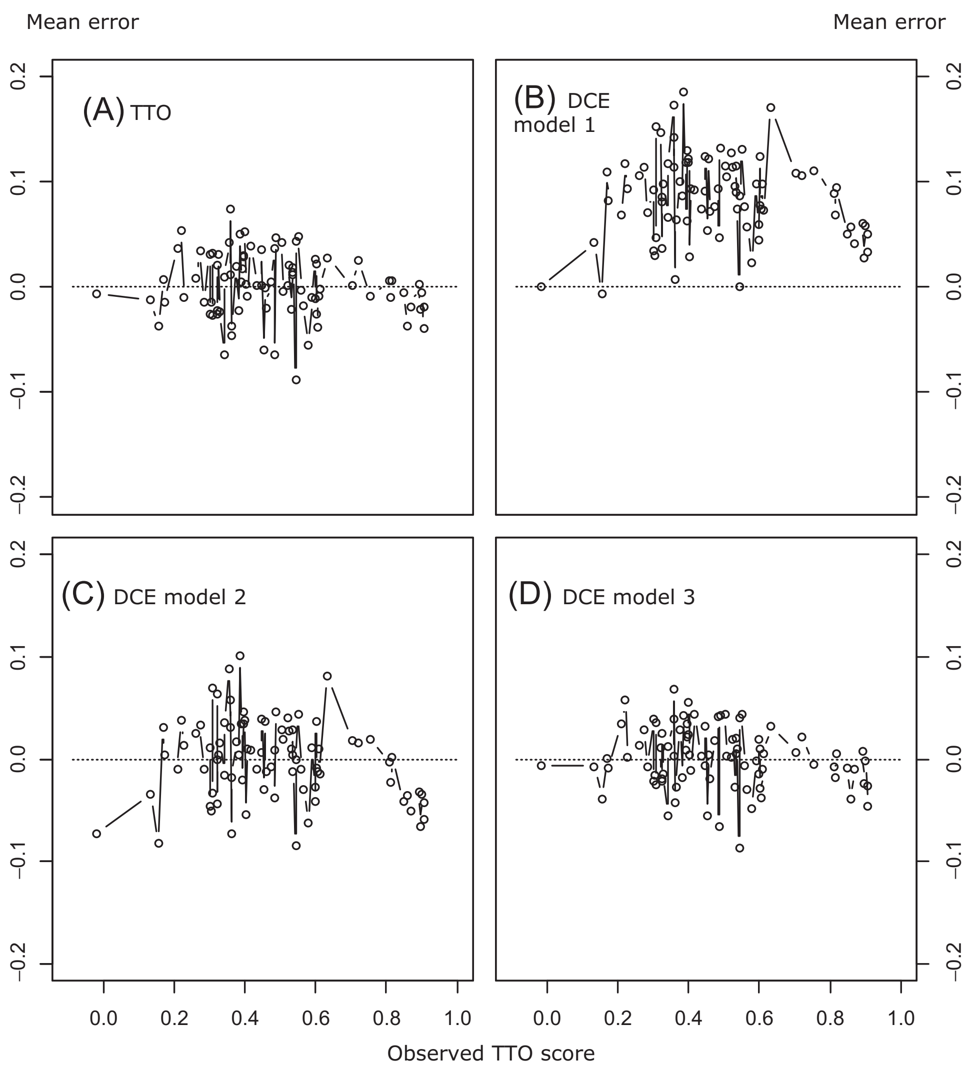

Fig. 4 – Mean error between observed TTO and predicted score. DCE, discrete choice experiment; TTO, time trade-off.

assumed to be the same for all people, 2) the model cannot consider panel structure, and 3) the independence of irrelevant alternatives is assumed.

In our survey, the value set based on cTTO was very similar in DCE model 2 (mapping) and DCE model 3 (hybrid). The study design, such as the ranking task before cTTO and/or the order of the task (DCE is followed by cTTO), may also influence these results. The EQ-5D-5L kernel obtains a lower probability density function than does the EQ-5D-3L in regions of smaller QOL scores, possibly because the linguistic expressions are changed at the lowest mobility level; "I am confined to bed" (in the EQ-5D-3L) was modified to "I am unable to walk about" (in the EQ-5D-5L). Respondents will evaluate former health states as worse than the latter.

The values predicted by DCE model 1 (anchored to the worst health state) are not well fitted to the observed cTTO data. Specifically, this model overestimates the empirical cTTO results. Figure 2 shows that the line y = x locates below almost all the health states. The observed mean cTTO score is more than 0 for all but the worst health state (55555). The worst health states were valued as WTD by 22.9% of the respondents, although only 7.5% of the cTTO scores (averaged among all health states) were less than 0. In the cTTO survey, the lead-time TTO measured the WTD states differently from the normal TTO. Therefore, the worst state may represent a special case, and may be unsuitable as an anchoring point for scaling the DCE data from 0 to 1.

At present, we lack a criterion standard for converting latent DCE coefficients to QOL scores. Participants might more easily respond to DCE questions than to cTTO tasks, which trade duration of a health state with QOL. In our survey, however, the cTTO information was essential for converting the DCE data to QOL scores. In some studies, the "duration of life" and "death," which were described in the Introduction section, were included as attributes or choices. If their respective coefficients could be used, cTTO data were not needed to change the DCE coefficient to the QOL scale. Our study, however, does not include them. This highlights the limitation of constructing a scoring algorithm from DCE data. In addition, the best conversion method cannot be theoretically determined exclusively from our data because the "true" DCE-based QOL score cannot be determined. We discussed the performances of the models by comparing their results with the observed cTTO data. Nevertheless, the model that best fits the observed cTTO data is not necessarily better than other models. If the best fit to cTTO data determines the best model, the DCE data are not needed. In addition, model selection based on information criteria (e.g., Akaike information criterion or Bayesian information criterion) seems meaningless for our survey because the different models process different data sets (cTTO, DCE, or both cTTO and DCE). This is another limitation of our study.

The hybrid model simultaneously estimates the coefficients of the cTTO and DCE models. The hybrid model, however, is sensitive to the number ratio of cTTO and DCE responses. Our data include more cTTO responses than DCE responses, and one continuous cTTO response contains more information than one discrete DCE response. Therefore, the cTTO responses exert a strong influence on the results of the hybrid model.

The coefficients of cTTO and DCE in the hybrid model are not assumed to be independent, but are constrained such that the two coefficients are linearly related. The same relation is assumed in the mapping method. Mapping methods also constitute a type of hybrid model because the mapping function is estimated from both cTTO and DCE data. Therefore, we consider that the mapping and hybrid methods differ not by their "hybrid" status, but on how they estimate the linear relation between the two coefficients. The mapping method is based on two-stage estimation; that is, the mapping function is determined after calculating the DCE coefficients, whereas the hybrid method determines the TTO and DCE coefficients and their linear relation in a single step. Although the estimation method is different, the value sets obtained by the mapping and hybrid methods are similar in this survey. The mapping method may have been adequate in terms of simplicity if the estimator was not significantly changed as it had been in our study.

In our Japanese valuation study, no inconsistency was identified in the cTTO and latent DCE coefficients. Nevertheless, if some inconsistencies are observed, the hybrid method may be able to improve its coefficients [26] by mixing cTTO and DCE information. In that case, however, the collected data might have some problems resulting from the characteristics of the questionnaire, study design, data collection process, and so on.

Although the DCE promises to overcome the limitation of TTO (or SG) when constructing the value sets of newly developed questionnaires, it faces challenges when converting latent DCE scores to the QOL scale. Here, we showed the DCE scoring algorithm for the EQ-5D-5L and compared coefficients by using different conversion methods. These DCE value sets are also useful for scoring the EQ-5D-5L. We also demonstrated the properties and feasibility of each model. A standard method is expected to be realized as more empirical data accumulate. In addition, we created a DCE scoring algorithm using TTO data; other designs mentioned in the Introduction section might, however, be able to function without using TTO or SG data. Which method is better remains to be determined.

Acknowledgment

We thank Makoto Kobayashi (CRECON Medical Assessment Inc.) for advice and support for this project.

Source of financial support: This study was funded by Health and Labour Science Research Grants, Ministry of Health, Labour and Welfare (H26-seisaku-shitei-012).

REFERENCES

[1] Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53–72.
[2] Szende A, Oppe M, Devlin N. EQ-5D Value Sets: Inventory, Comparative Review and User Guide. Dordrecht, The Netherlands: Springer, 2007.
[3] Tsuchiya A, Ikeda S, Ikegami N, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ 2002;11:341–53.
[4] Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the measurement of health utilities in published cost-utility analyses. Value Health 2006;9:213–8.
[5] van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012;15:708–15.
[6] Janssen BM, Oppe M, Versteegh MM, Stolk EA. Introducing the composite time trade-off: a test of feasibility and face validity. Eur J Health Econ 2013;14(Suppl. 1): S5–13.
[7] Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. J Health Econ 2012;31:306–18.
[8] Pullenayegum E, Xie F. Scoring the 5-level EQ-5D: can latent utilities derived from a discrete choice model be transformed to health utilities derived from time tradeoff tasks? Med Decis Making 2013; 33:567–78.
[9] Xie F, Pullenayegum E, Gaebel K, et al. Eliciting preferences to the EQ-5D-5L health states: discrete choice experiment or multiprofile case of best-worst scaling? Eur J Health Econ 2014;15:281–8.
[10] Ikeda S, Shiroiwa T, Igarashi I, et al. Developing a Japanese version of the EQ-5D-5L value set (in Japanese). J Natl Inst Public Health 2015;64:47–55.
[11] Viney R, Norman R, Brazier J, et al. An Australian discrete choice experiment to value EQ-5D health states. Health Econ 2014;23: 729–42.
[12] Rowen D, Brazier J, Van Hout B. A comparison of methods for converting DCE values onto the full health-dead QALY scale. Med Decis Making 2015;35:328–40.
[13] 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.
[14] 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.
[15] Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095–108.
[16] Robinson A, Spencer A. Exploring challenges to TTO utilities: valuing states worse than dead. Health Econ 2006;15:393–402.
[17] 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 (Suppl. 1):S25–31.
[18] Devlin NJ, Tsuchiya A, Buckingham K, Tilling C. A uniform time trade off method for states better and worse than dead: feasibility study of the "lead time" approach. Health Econ 2011;20:348–61.
[19] Attema AE, Versteegh MM, Oppe M, et al. Lead time TTO: leading to better health state valuations? Health Econ 2013;22:376–92.
[20] Devlin N, Buckingham K, Shah K, et al. A comparison of alternative variants of the lead and lag time TTO. Health Econ 2013;22:517–32.
[21] Bliemer MCJ, Rose JM, Hess S. Approximation of Bayesian efficiency in experimental choice designs. J Choice Model 2008;1:98–126.
[22] Luo N, Li M, Stolk EA, Devlin NJ. The effects of lead time and visual aids in TTO valuation: a study of the EQ-VT framework. Eur J Health Econ 2013;14(Suppl. 1):S15–24.
[23] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271–92.
[24] Ramos-Goni JM, 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(Suppl. 1): S33–42.
[25] Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ 2010;11:215–25.
[26] Ramos-Goni JM, Pinto-Prades JL, Oppe M, et al. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. Med Care [Epub ahead of print] December 17, 2014.
[27] Gu Y, Norman R, Viney R. Estimating health state utility values from discrete choice experiments—a QALY space model approach. Health Econ 2014;23:1098–114.