Valuation Survey of EQ-5D-Y Based on the International Common Protocol: Development of a Value Set in Japan

Takeru Shiroiwa, Shunya Ikeda, Shinichi Noto, Takashi Fukuda, and Elly Stolk

Background. EQ-5D-Y is a preference-based measure for children and adolescents (aged 8–15 y). This is the first study to develop an EQ-5D-Y value set for converting EQ-5D-Y responses to index values. **Methods**. We recruited 1047 respondents (aged 20–79 y) from the general population, stratified by gender and age group, in 5 Japanese cities. All data were collected through face-to-face surveys. Respondents were asked to value EQ-5D-Y states for a hypothetical 10-y-old child from a proxy perspective using composite time tradeoff (cTTO) and a discrete choice experiment (DCE). The discrete choice data were analyzed using a mixed logit model. Latent DCE values were then converted to a 0 (death)/1 (full health) scale by mapping them to the cTTO values. **Results**. The mean observed cTTO value of the worst health state [33333] was 0.20. Analysis of the DCE data showed that the coefficients of the domains related to mental functions ("Having pain or discomfort" and "Feeling worried, sad, or unhappy") were larger than those for the domains related to physical and social functions. By converting latent DCE values to a utility scale, we constructed a value set for EQ-5D-Y. No inconsistencies were observed. The minimum predicted score was 0.288 [33333], and the second-best score was 0.957 [12111]. **Conclusion**. A value set for EQ-5D-Y was successfully constructed. This is the first survey of an EQ-5D-Y value set. Interpreting the differences between EQ-5D-Y and EQ-5D-5L value sets is a future task with implications for health care policy.

Keywords
composite TTO, discrete choice experiment, EQ-5D-Y, preference-based measure, QALY

EQ-5D-Y instrument is a preference-based measure that has been constructed to measure the health states of children and adolescents (aged 8–15 y). According to a review by Rowen et al.,[1] approximately 10 preference-based measures for children and adolescents have been developed as of 2020, including the Child Health Utility 9D (CHU9D),[2] the Assessment of Quality of Life-6 Dimensions (AQoL-6D),[3] the Health Utilities Index Mark 2 (HUI2),[4] and HUI3,[5] among others. In 2009, EQ-5D-Y was introduced as a child-friendly version of EQ-5D by the EuroQol Group. EQ-5D-Y is conceptually equivalent to EQ-5D,[6] but the wording of the severity levels of the dimensions has been adapted to be more relevant to younger populations. The instrument allows children (or their proxies) to report functioning on 5

basic dimensions of health, and subsequently, a value from a societal perspective can be attached to every reported health state reflecting how preferable that state is.

In 2020, Ramos Goñi et al.[7] presented an international valuation protocol for EQ-5D-Y that enabled the Japanese valuation survey to commence. However, there are some controversial issues related to valuation with respect to pediatric instruments. When originally investigating a value set for EQ-5D-Y, Kreimeier et al.[8] showed

Corresponding Author
Takeru Shiroiwa, Center for Outcomes Research and Economic Evaluation for Health (C2H), National Institute of Public Health, 2-3-6 Minami, Wako, Saitama, 351-0197, Japan (t.shiroiwa@gmail.com).

that the values obtained for EQ-5D-Y states exceeded EQ-5D values, which indicated that the labels attached to the levels of the 2 instruments corresponded to different severity levels. Higher values were placed on health states by children when compared with values placed on health states by adults. This tendency was also reported by Shah et al.[9] These studies confirmed the necessity to develop EQ-5D-Y value sets and that, in this respect, the application of adult value sets was not appropriate.

The protocol for valuing EQ-5D-Y follows the well-documented protocol for the valuation of EQ-5D-5L, which uses composite time tradeoff (cTTO) and a discrete choice experiment (DCE) as valuation methods.[10,11] In the EQ-5D-5L valuation protocol, cTTO is the primary valuation method and, optionally, the DCE responses can be used to enrich the data set. In contrast, the EQ-5D-Y valuation protocol involves a 2-step approach. The first step involves modeling DCE responses to derive values on a latent scale. In the second step, cTTO values are used to anchor the DCE-derived values onto the full health-dead scale. Similar 2-step approaches have been used to generate utilities for CHU9D in Australia and China and for the Infant Health-related Quality of Life Instrument (IQI). The advantages of this 2-step approach are its feasibility and flexibility. Moreover, unbundling DCE and cTTO allows researchers to exercise choice over how and to whom these tasks are administered, which facilitates experimentation and enables quick dissemination of new results into the protocol, if required.[12,13]

Most other variations among approaches to generate utilities for different instruments appear to be entirely commonplace, with the exception of choices about perspective. Adult health states can usually be valued by adults themselves, but it may not be possible to elicit values from children. Possible approaches to the valuation of child health include eliciting values from adolescents or asking adults to value EQ-5D-Y states from a proxy perspective. In the Australian CHU9D valuation survey,[14,15] the AQoL[3] and the 16-Dimension (16D)[16] health states were valued by adolescents for themselves. However, values were elicited from the general population for a child (i.e., from proxies) in the valuation of IQI[13] and in pilot work concerning EQ-5D-Y.[7] The latter approach has now also been adopted in the EQ-5D-Y valuation protocol, to be consistent with the taxpayer perspective, and because it was clear that the tasks could not be administered with children.

In Japan, there are no preference-based measures for children or adolescents featuring a Japanese value set. This has led to some difficulties in evaluating the cost-effectiveness of technology for children. Hence, the aim of this study was to produce a Japanese value set for EQ-5D-Y, following the protocol suggested.[7]

Methods

Instruments

The EQ-5D-Y instrument is composed of 5 dimensions, each assessed at 3 levels. It has the following domains: "mobility," "looking after myself," "doing usual activities," "having pain or discomfort," and "feeling worried, sad, or unhappy." It covers the same basic dimensions of health as the adult version EQ-5D, but the words and phrases have been simplified to be more child friendly and are thus slightly different from those in EQ-5D-3L.[6,8] For example, level 3 for mobility was changed from "confined to bed" to "unable to walk around," and the fifth dimension was changed from "anxiety/depression" into "feeling worried, sad, or unhappy." Furthermore, in the Japanese version of EQ-5D-Y, the use of kanji characters was limited. EQ-5D-Y is designed for self-report by children and adolescents aged 8 to 15 y.[17] It might be possible for children and adolescents aged 12–15 y to use the adult EQ-5D version. However, for children aged 4–7 y, a proxy version of EQ-5D-Y is recommended.

The Japanese version of EQ-5D-Y was prepared by a Japanese research group, which included the authors of this article, based on a first draft provided by the EuroQol Group. The processes of translation, back translation, and harmonization for the first draft were completed by the EuroQol Group, independently of the Japanese group. The Japanese EQ-5D-Y has been confirmed for its psychometric properties.[18]

cTTO and DCE

Preferences with respect to EQ-5D-Y health states in the general population (not in the population of children and

adolescents) were measured using the cTTO and DCE methods using the EuroQol Group's recently published protocol for EQ-5D-Y valuation.[7] cTTO is a TTO variant that adopts conventional TTO for the valuation of better-than-dead states and lead-time TTO for the valuation of worse-than-dead states.[19]

According to this protocol,[7] the requirement is to include at least 10 states in the cTTO tasks and 150 pairs in the DCE tasks. The value set can be established based on a combination of cTTO and DCE data, which may be linked, for example, using a mapping approach. The sample size should be at least 200 for the cTTO task and 1000 for the DCE task. Inclusion of more states/pairs/individuals is allowed. However, estimation of a value set based on cTTO responses requires the expansion of the TTO task.

In the cTTO task, participants were asked to consider which option was better for a hypothetical 10-y-old child from a proxy perspective: living for 10 y in a health state described by EQ-5D-Y or living x years in full health. In lead-time TTO, a series of choices was offered between years of life in full health and a life with "10 y in full health followed by 10 y in the EQ-5D-Y state presented."

In the DCE survey, similar to the cTTO survey, the participants were required to imagine a hypothetical 10-y-old child's health state. Then, 2 health states (states A and B) from a combination of EQ-5D-Y descriptions were presented. The participants chose the one they preferred between the 2 options from a proxy perspective. Modeling of the DCE responses produced values on a latent scale, which could be converted to values on a quality-adjusted life-year (QALY) scale by anchoring on the values derived from the cTTO task.

Participants and Survey Process

Computer-assisted personal interviews were conducted in 5 cities in Japan (Tokyo, Niigata, Osaka, Okayama, and Fukuoka). These cities are representative of various regions in Japan and are geographically dispersed. The inclusion criteria were as follows: 1) aged 20 to 79 y, 2) current Japanese residency, 3) ability to visit the survey room in 1 of the 5 cities, 4) ability to provide informed consent, and 5) ability to complete the tasks in Japanese. The participants were recruited based on nonrandom quota sampling by a research company (ANTERIO Inc.), which sampled 1047 respondents throughout Japan (i.e., considering the size of the population, approximately 300 respondents from Tokyo and from Osaka, 200 from Fukuoka, and 100 from Niigata and from Okayama). The same number of respondents was collected by gender and age group in each city. The

interviews were conducted one-on-one at centrally located interview sites. The interviews were fully scripted.

The computer-assisted personal interview tool used in this study was the EuroQol portable valuation technology (EQ-PVT), developed by the EuroQol Group and translated into Japanese. The EQ-PVT was used to implement the TTO and DCE tasks. It stored all responses and the data needed to create quality control reports concerning interviewer performance. Apart from the valuation perspective, implementation of the cTTO task was consistent with that in version 2.1 of the EQ-VT protocol.[10]

All participants were asked to complete both the cTTO and DCE tasks, with the former always performed before the latter. After the cTTO and DCE tasks, the participants' demographic information was collected. In the cTTO phase, the first 3 tasks (wheelchair example) were presented as a practice exercise. First, they considered a hypothetical situation of living for 10 y in a wheelchair as a 10-y-old child. Next, they were asked to consider 2 unlabeled states from the same perspective: "much better than being in a wheelchair" and "much worse than being in a wheelchair, so bad that one would prefer to die immediately." After this introduction to the task, participants practiced valuing health states using cTTO for 3 states defined by the EQ-5D-Y descriptive system. Finally, they were asked to undertake the real valuation tasks.

One block, including 6 health states, was randomly allocated to each participant from 5 blocks. As the worst EQ-5D-Y state was included in all blocks, a total of 26 health states were used. The 26 health states included 18 states representing an orthogonal array, all 5 mild states (4 dimensions at level 1 and 1 dimension at level 2), and 3 other states. Health states were presented in a random order.

In the DCE task, 1 block consisting of 15 pairs of EQ-5D-Y health states was randomly allocated to each participant from 10 blocks (an experimental design created by the EuroQol Group). A distinctive feature of the DCE design was that all pairs included attribute-level overlap. An efficient design approach was used to create the experimental design. The order in which the questions were presented was randomized, and the presentation positions (left or right) of the 2 health states were also randomized in the DCE survey.

The survey was conducted on weekends (Friday, Saturday, and Sunday) from February to March 2019. Before administering the survey, all the investigators received training for approximately 1 d. To ensure quality and consistency among investigators, the number of investigators was limited to 11 (90–100 samples per

interviewer). To reduce the interviewer effect, all interviewers received strict quality control (QC) checks by the EuroQol Group after the survey each week, as described in by Ramos-Goñi et al.[20]; subsequently, feedback was provided to each interviewer. The valuation survey was conducted for the first 3 wk in Tokyo, and QC was conducted every week. After all interviewers passed the QC check 3 times in the Tokyo survey, they continued the survey in the other cities.

Statistical Analyses

A mixed logit model was used for the analysis of the DCE data (model 1). A mixed logit model can consider the heterogeneity of coefficients without an irrelevant alternative assumption, whereas a simple conditional logit model assumes that all responses are independent. When choices are analyzed based on random utility theory, U_{ij} (the disutility respondent j derives from choosing item i) can be divided into an explainable component (V_{ij}) and a random component (ε_{ij}),

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

$$V_{ij} = \beta_{12}X_{12} + \beta_{12}X_{12} + \ldots + \beta_{pq}X_{pq} + \ldots + \beta_{53}X_{53},$$

where β_{pq} represents the effects of the qth level (q = 2 or 3, where the first level is the reference term) of the pth (1 $\leq$ p $\leq$ 5) item. This model accounted for the panel structure in the data and for heteroscedasticity. At the same time, a simple conditional model was applied (model 2). To confirm the consistency with the DCE results, a linear mixed model was applied to cTTO values of 26 states to estimate each item's coefficient (model 3). Interaction with any level 3 responses was considered by adding the N3 term (N3 = 1, if any level 3 responses were included in the health states) to model 3 (model 4). As suggested by Stolk et al.,[10] the TTO value was censored at 1. Considering these distribution characteristics, the Tobit model was also used for the cTTO data (model 5).

Mixlogit of STATA15 was used to estimate each coefficient of the mixed logit model. SAS 9.4 was used for the linear mixed model and the other statistical analyses.

Constructing the Value Set from DCE Data

To convert the latent DCE values to a scale anchored at full health (1) and dead (0), the modeled DCE values were mapped onto the observed TTO values. The linear relationship function f(.) between the latent DCE values and the cTTO values of the 26 health states was estimated as $cTTO_i = f(DCE_i) + \varepsilon_i$, where $cTTO_i$ is the observed mean cTTO value and DCE_i is the latent DCE value for the ith health state (1 $\leq$ i $\leq$ 26).[21] The hybrid model[22] was also a candidate for the analysis of DCE and cTTO data in constructing a value set. The model can simultaneously treat both DCE and cTTO data, which is different from the above 2-step approach. However, this was not applied to the data set because the hybrid model is dependent on the relative amounts of TTO and DCE data collected, the ratio of which is arbitrarily set by researchers and not well balanced in the EQ-5D-Y protocol. Because it is uncertain how these features affect the performance of the hybrid model, a mapping approach was applied.

Results

A total of 1047 respondents from 5 cities (308 from Tokyo, 310 from Osaka, 210 from Fukuoka, 110 from Okayama, and 109 from Niigata) participated in the EQ-5D-Y valuation survey. The participants' characteristics are summarized in Table 1. The participants' age and gender distributions were well balanced. With respect to household income, 50.1% reported earnings of less than JPY 6 million (USD 55,000; USD 1 = JPY 110 as of April 2019), compared with the median household income of all Japanese families of JPY 5.4 million (USD 49,000) in 2016. All participants completed the cTTO and DCE surveys; however, 2 responses were not recorded in the DCE survey using EQ-PVT. A total of 6082 cTTO values and 10,458 DCE choices were collected. The mean and median response times in the 6 cTTO tasks were 9.1 min (*s* = 2.3 min) and 8.8 min (interquartile range 7.6–10.3 min), respectively. Few issues were encountered during the QC process. Only about 2.5% of the interviews were flagged as not meeting QC standards. With respect to the cTTO tasks, the mean number of moves before reaching a point of indifference was 4.8. No data were excluded on account of the QC process. The distribution of cTTO values and the cTTO values for each health state were similar between interviewers. Thus, high levels of protocol compliance and the presence of few interviewer effects in the data were demonstrated.

cTTO Results

Table 2 shows the average cTTO values for the 26 health states described by EQ-5D-Y. The predicted score (based on Table 3) and root mean square error are also shown.

Table 1 Background Characteristics of Respondents

	n	%
Location		
Tokyo	308	29.4
Osaka	310	29.6
Fukuoka	210	20.1
Okayama	110	10.5
Niigata	109	10.4
Gender		
Male	523	50.0
Female	524	50.1
Age, y		
20–29	174	16.6
30–39	174	16.6
40–49	175	16.7
50–59	175	16.7
60–69	174	16.6
70–79	175	16.7
With children		
Yes	703	67.1
No	344	32.9
Employment		
Employed or self-employed	677	64.7
Retired	76	7.3
Student	43	4.1
Homemaker	197	18.8
Others	54	5.2
Marital status		
Married	677	64.7
Unmarried	266	25.4
Divorced	62	5.9
Bereaved	41	3.9
Other	1	0.1
Education		
Junior high school	27	2.6
High school	366	35.0
College	234	22.3
University or graduate	419	40.0
Other	1	0.1
Household income		
<JPY 2 million	71	6.8
JPY 2 million–4 million	196	18.7
JPY 4 million–6 million	257	24.6
JPY 6 million–10 million	311	29.7
JPY 10 million–15 million	111	10.6
≥JPY 15 million	31	3.0
Unknown	70	6.7
EQ-5D-5L ($\bar{x}$ [s])	0.95 [0.08]	
EQ-VAS ($\bar{x}$ [s])	82.3 [12.5]	

The TTO values for health states [11112] and [21111] were 0.94 (the highest), and the score for health state [33333] was 0.20 (the lowest). A total of 910 respondents (86.9%) preferred the worst EQ-5D-Y state (33333) to death, and only 137 respondents (13.1%) evaluated it as worse than dead. Figure 1 shows the distribution of the cTTO values. The peak of the distribution was at cTTO value = 0.95, and the density of the distribution with a cTTO value <0 was very low (3.2%). As the misery score (the sum of level scores across dimensions) was higher, the mean cTTO value was lower, and the standard deviation increased with the misery score (Figure 2).

DCE Results and the EQ-5D-Y Value Set

Table 3 presents the parameter estimates obtained from the DCE data using a mixed logit model. The mixed logit model fit the DCE data better than the simple conditional logit model. Similar coefficients were obtained from models 3 to 5, which analyzed cTTO data. The N3 interaction term in model 4 was significantly negative. No inconsistencies were observed for any of the domains in any model; lower levels had lower negative scores. Note that the DCE results were estimated on a latent scale, and it was not possible simply to compare coefficients between the DCE and cTTO models.

Using the coefficients in Table 3 from the mixed logit model, latent DCE values were computed for EQ-5D-Y states. Next, the mapping function was estimated to predict cTTO values based on the modeled latent DCE values. The estimated equation from the regression of the cTTO value (disutility) to the latent DCE values was cTTO value = −0.058 × latent DCE value + 0.975. The relationship between the observed cTTO values and the derived DCE values is shown in Figure 3.

In Table 4, the DCE coefficients (from model 1) are rescaled. When the EQ-5D-Y index value is calculated, this table can be used. For example, the score for the EQ-5D-Y health state 23213 can be calculated from Table 3 by 1 + (−0.025 [intercept] − 0.040 [mobility] − 0.070 [looking after myself] − 0.038 [doing usual activities] − 0 [having pain or discomfort] − 0.156 [feeling worried, sad, or unhappy]) = 0.67. The coefficients for the domains "Having pain or discomfort" and "Feeling worried, sad, or unhappy" were smaller than those for the other domains. The coefficient for the "Looking after myself" domain had the largest value. Figure 4 shows the distribution of the predicted scores for all 243 EQ-5D-Y health states. The root mean square error and intraclass correlation coefficients between the 26 estimated and empirical cTTO health states were 0.026 and 0.986, respectively. Both of these results showed that the predictions obtained using this function were similar to the empirical cTTO values.

The minimum predicted score was 0.288 [33333], and the second-best score was 0.957 [12111]. Unlike for the adult Japanese EQ-5D-3L states (second-best score 0.812), no large gap was observed between the best and second-best health states.

Table 2 Mean Composite Time-Tradeoff Scores of 26 Health States

Health State	n	$\bar{x}.$	s	Predicted Score[a]	MSRE
11112	211	0.94	0.08	0.93	0.054
11121	206	0.92	0.11	0.90	0.068
11211	206	0.93	0.16	0.94	0.060
12111	211	0.93	0.09	0.96	0.047
21111	210	0.94	0.06	0.94	0.040
11122	211	0.88	0.14	0.85	0.098
21211	206	0.89	0.18	0.90	0.086
12212	210	0.86	0.12	0.87	0.087
22121	211	0.83	0.21	0.84	0.116
11313	209	0.68	0.25	0.72	0.181
13221	211	0.80	0.15	0.79	0.116
23112	206	0.80	0.23	0.82	0.131
31131	211	0.64	0.26	0.62	0.184
12331	211	0.60	0.29	0.59	0.204
32113	209	0.65	0.26	0.71	0.190
13133	209	0.47	0.33	0.48	0.217
21332	206	0.57	0.29	0.51	0.207
22223	211	0.63	0.28	0.64	0.196
22232	209	0.58	0.31	0.56	0.209
31223	210	0.63	0.25	0.61	0.170
33311	209	0.70	0.29	0.72	0.189
22233	210	0.49	0.29	0.45	0.215
32322	211	0.62	0.27	0.64	0.182
23323	210	0.57	0.25	0.53	0.195
33232	211	0.48	0.33	0.46	0.238
33333	1047	0.20	0.37	0.29	0.284

MSR, mean square root error.
[a]Predicted scores are calculated using Table 4.

Table 3 Results of Analysis of the DCE and cTTO Data[a]

	DCE				cTTO					
	Model 1 (Mixed Logit)		Model 2 (Conditional Logit)		Model 3 (Repeated-Measures ANOVA)		Model 4 (Model 3 + N3)		Model 5 (Tobit)	
Dimension/Level	Estimate	*P* Value	Estimate	*P* Value	Estimate	*P* Value	Estimate	*P* Value	Estimate	*P* Value
Intercept	—	—			0.973	<0.0001	1.025	<0.0001	1.024	<0.0001
Mobility										
2	−0.699	<0.0001	−0.545	<0.0001	−0.030	<0.0001	−0.040	<0.0001	−0.038	<0.0001
3	−1.546	<0.0001	−1.455	<0.0001	−0.117	<0.0001	−0.129	<0.0001	−0.128	<0.0001
Looking after myself										
2	−0.319	<0.0001	−0.272	<0.0001	−0.039	<0.0001	−0.045	<0.0001	−0.052	<0.0001
3	−1.204	<0.0001	−1.028	<0.0001	−0.099	<0.0001	−0.108	<0.0001	−0.106	<0.0001
Doing usual activities										
2	−0.658	<0.0001	−0.596	<0.0001	−0.032	<0.0001	−0.041	<0.0001	−0.033	0.003
3	−1.757	<0.0001	−1.515	<0.0001	−0.107	<0.0001	−0.121	<0.0001	−0.120	<0.0001
Having pain or discomfort										
2	−1.340	<0.0001	−1.163	<0.0001	−0.032	<0.0001	−0.044	<0.0001	−0.053	<0.0001
3	−4.681	<0.0001	−3.209	<0.0001	−0.239	<0.0001	−0.254	<0.0001	−0.260	<0.0001
Feeling worried, sad, or unhappy										
2	−0.850	<0.0001	−0.728	<0.0001	−0.038	<0.0001	−0.044	<0.0001	−0.047	<0.0001
3	−2.708	<0.0001	−2.039	<0.0001	−0.191	<0.0001	−0.204	<0.0001	−0.195	<0.0001
N3							−0.051	<0.0001		
AIC	10,980		11,933		−745		−766		2400	

DCE, discrete choice experiment; cTTO, composite time tradeoff; ANOVA, analysis of variance; AIC, akaike information criteria.
[a]DCE coefficients are latent utility and thus are not comparable with cTTO coefficients.
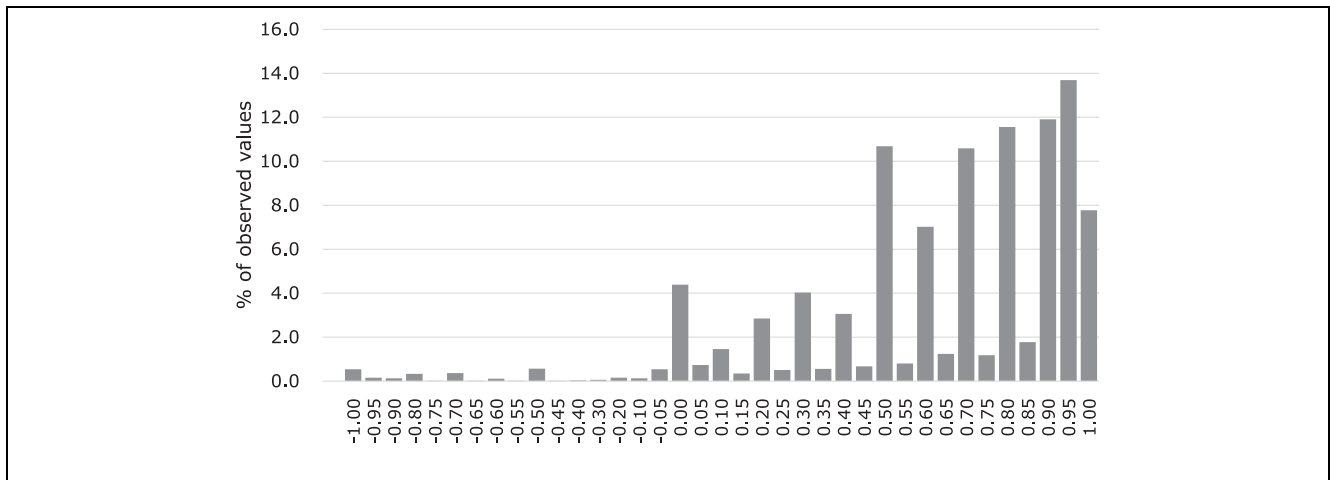
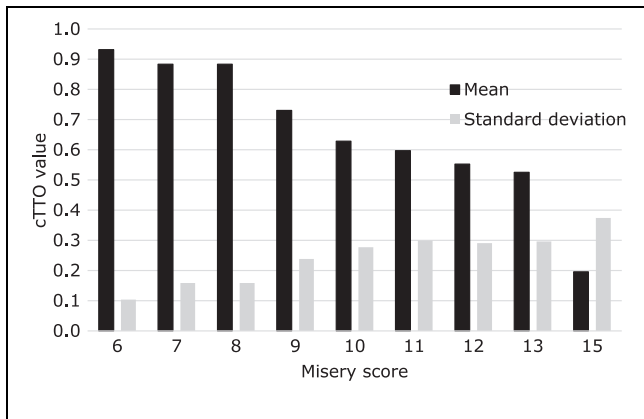**Figure 1** Composite time-tradeoff (cTTO) value distribution.

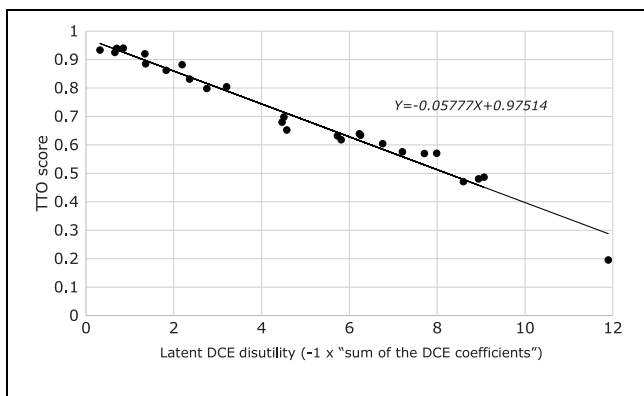**Figure 2** Mean and standard deviation of composite time-tradeoff (cTTO) value by misery score.

Figure 3 Relationship between latent discrete choice experiment scores and composite time-tradeoff values of 26 EQ-5D-Y states.

Table 4 Rescaled Discrete Choice Experiment Coefficient for Calculation of EQ-5D-Y Index Based on the Preferred Model

Domain	Level	Rescaled DCE Coefficient
Intercept		−0.025
Mobility	2	−0.040
	3	−0.089
Looking after myself	2	−0.018
	3	−0.070
Doing usual activities	2	−0.038
	3	−0.101
Having pain or discomfort	2	−0.077
	3	−0.270
Feeling worried, sad, or unhappy	2	−0.049
	3	−0.156

Discussion

The objective of this study was to establish a Japanese value set for EQ-5D-Y. Preferences were elicited from 1047 participants, using DCE and cTTO, for EQ-5D-Y health states, and regression techniques were used to predict values for each of the states. The feasibility of the study protocol was confirmed: respondents were able to perform the tasks, and adequate levels of task engagement were observed. For example, the mean number of moves before reaching the point of indifference was sufficiently large. The interviewers also performed well, with a small percentage of flagged interviews: high levels of protocol compliance were observed, and there were no significant interviewer effects, which was supported by the similar distribution of cTTO values and of cTTO values for each health state. A mixed logit model was used
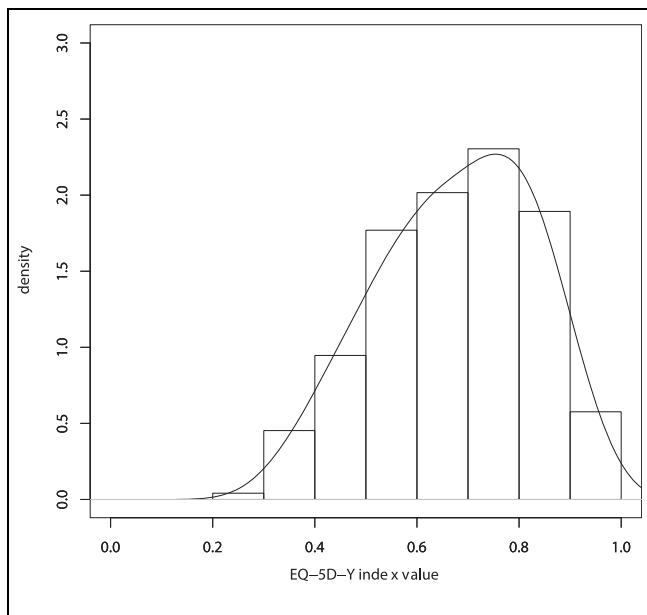
Figure 4 Distribution of EQ-5D-Y predicted value (all health states).

- TTO values for children are higher because people are more reluctant to trade time off on behalf of a child. Therefore, the values are calibrated differently. The same level of quality of life induces a smaller tradeoff where children are concerned.
- The descriptors used in EQ-5D-Y use wording relevant to children, and the choice of words makes the health states look less severe (22222 ≠ 22222). For example, the youth version has labels "some" or "a lot," whereas the corresponding labels in the adult version are "moderate" or "extreme/unable to."
- Children can actually have a better quality of life while experiencing the same health problems as an adult. Parents or other caregivers may dedicate more time to their children regardless of their health status. Moreover, children who are developing more independence may be less likely to rely on others for some things as compared with adults.

All of these explanations may hold some truth, which would imply that there is limited comparability of values derived from EQ-5D-Y and EQ-5D. The consequences for users are contemplated below.

A choice was made by the EuroQol Group to attach proxy-reported values from adults for a 10-y-old child to EQ-5D-Y health states. Another option could be to collect self-reported values from adolescents or young adults. For example, the CHU9D value set for Australia has been based on adolescents' (aged 11–17 y) best-worst scaling responses and young adults' (aged 18–29 y) TTO results.[14] Normally, arguments of inclusiveness support adolescents' views, but the inherent limitation is that only partial information can be obtained, because tasks involving comparisons to death are not appropriate for this age group. Hence, the CHU-9D team derived TTO values from people closest in age: young adults. An open question is whether self-reported values from young adults or proxy-reported values from adults for a child are more relevant for the valuation of health outcomes in children. Moreover, it is unknown how the self-reported values derived from adolescents and adults (which were found to differ) compared with proxy-reported values from adults. Disentangling age and proxy effects is an important area for future research, and in the absence of evidence, researchers need to remain open-minded toward alternative approaches and to future updates of the protocol.

This study has some limitations. Participants were not recruited through a rigorous random sampling procedure because with such an approach, the time required for the survey would be too long; rather, the subjects were recruited from a few preselected geographical areas. In

to estimate the values for each of the states based on the DCE observations. All parameters were logically consistent. The least preferred item was level 3 for the "Having pain or discomfort" domain, and the second least was level 3 for the "Feeling worried, sad, or unhappy" domain.

The Japanese EQ-5D-Y value set has a narrow value range as compared with previously reported international results. The observed value of the worst EQ-5D-Y state was 0.20 (and had a predicted value of 0.28). In European countries, the mean observed value for the worst state [33333] reported by Kreimeier et al.[8] was −0.14. However, this difference in value range is not unique to EQ-5D-Y: Japanese values for the adult EQ-5D-5L were also higher than the corresponding European values.[23] The finding that the values for EQ-5D-Y states exceeded the values for the EQ-5D states for adults is consistent with the results published by Kreimeier et al.[24] The results thus appear to have good face validity and are also supported by the high level of protocol compliance and the lack of interviewer effects. Hence, the value set obtained reliably reflects the preferences of Japanese participants.

It is important to consider why the values attached to EQ-5D-Y states exceed those for the corresponding EQ-5D-3L or 5L health states (e.g., state 22222), since the consequences with respect to the use of EQ-5D-Y values alongside EQ-5D-3L or 5L values will depend on this finding. These differences are multicausal.

addition, although the cTTO tasks are already complex, the use of cTTO for valuing child health adds an extra layer of complexity: it may have been difficult for adults to imagine what impact health problems would have had on a hypothetical 10-y-old.

The observed differences in the EQ-5D-Y value ranges for young people and the EQ-5D for adults have implications for users and policy makers. Using the EQ-5D-Y weights presented in this article, it will be possible to compare relative levels of quality of life across different groups of children but not to compare levels of quality of life observed in children and adults. Similarly, if the EQ-5D-Y is to be applied routinely in assessing the cost-effectiveness of new life-saving health care interventions for children, the results can be used to support reimbursement decisions in pediatric settings. However, if children and adults are treated for the same condition, they can have different values attached to comparable quality-of-life outcomes, leading to differences in estimates of QALYs gained. Because the QALY weights for children are higher than the QALY weights for adults, a similar degree of improvement in an adult's health state will generate fewer QALYs for a child. Thus, simple comparisons of cost-effectiveness ratios for adults and children should be avoided.

To conclude, we expect that this study will promote patient-centered research and economic evaluation in the area of health care technologies for children and adolescents. Important issues that need to be addressed in future research have also been identified.

Acknowledgments

ORCID iDs

Takeru Shiroiwa https://orcid.org/0000-0003-3055-9932
Elly Stolk https://orcid.org/0000-0001-5968-0416

References

1. Rowen D, Rivero-Arias O, Devlin N, Ratcliffe J. Review of valuation methods of preference-based measures of health for economic evaluation in child and adolescent populations: where are we now and where are we going? *Pharmacoeconomics*. 2020;38(4):325–40.
2. Stevens K. Valuation of the Child Health Utility 9D Index. *Pharmacoeconomics*. 2012;30(8):729–47.
3. Moodie M, Richardson J, Rankin B, Iezzi A, Sinha K. Predicting time trade-off health state valuations of adolescents in four Pacific countries using the Assessment of Quality-of-Life (AQoL-6D) instrument. *Value Health*. 2010;13(8):1014–27.
4. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Med Care*. 1996;34(7):702–22.
5. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care*. 2002;40(2):113–28.
6. 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(10):1727–36.
7. Ramos-Goñi JM, Oppe M, Stolk E, et al. International valuation protocol for the EQ-5D-Y-3L. *Pharmacoeconomics*. 2020;38(7):653–63.
8. Kreimeier S, Oppe M, Ramos-Goni JM, et al. Valuation of EuroQol Five-Dimensional Questionnaire, Youth Version (EQ-5D-Y) and EuroQol Five-Dimensional Questionnaire, Three-Level Version (EQ-5D-3L) health states: the impact of wording and perspective. *Value Health*. 2018;21(11):1291–8.
9. Shah KK, Ramos-Goñi JM, Kreimeier S, Devlin NJ. An exploration of methods for obtaining 0 = dead anchors for latent scale EQ-5D-Y values. *Eur J Health Econ*. 2020;21(7):1091–103.
10. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goni JM. Overview, update, and lessons learned from the international EQ-5D-5L valuation work: version 2 of the EQ-5D-5L valuation protocol. *Value Health*. 2019;22(1):23–30.
11. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health*. 2014;17(4):445–53.
12. Chen G, Xu F, Huynh E, Wang Z, Stevens K, Ratcliffe J. Scoring the Child Health Utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff. *Qual Life Res*. 2019;28(1):163–76.
13. Krabbe PFM, Jabrayilov R, Detzel P, Dainelli L, Vermeulen KM, van Asselt ADI. A two-step procedure to generate utilities for the Infant health-related Quality of life Instrument (IQI). *PLoS One*. 2020;15(4):e0230852.
14. Ratcliffe J, Huynh E, Chen G, et al. Valuing the Child Health Utility 9D: using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. *Soc Sci Med*. 2016;157:48–59.
15. Ratcliffe J, Flynn T, Terlich F, Stevens K, Brazier J, Sawyer M. Developing adolescent-specific health state values for economic evaluation: an application of profile case best-worst scaling to the Child Health Utility 9D. *Pharmacoeconomics*. 2012;30(8):713–27.
16. Apajasalo M, Sintonen H, Holmberg C, et al. Quality of life in early adolescence: a sixteen-dimensional health-related measure (16D). *Qual Life Res*. 1996;5(2):205–11.
17. Euroqol Group. *EQ-5D-Y User Guide*. 2020. Available from: https://euroqol.org/wp-content/uploads/2020/09/

EQ-5D-Y-User-Guide_version-2.0.pdf. Accessed December 18, 2020.

18. Shiroiwa T, Fukuda T, Shimozuma K. Psychometric properties of the Japanese version of the EQ-5D-Y by self-report and proxy-report: reliability and construct validity. *Qual Life Res*. 2019;28(11):3093–105.

19. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goni JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. 2016;34(10):993–1004.

20. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJ, Stolk E. Quality control process for EQ-5D-5L valuation studies. *Value Health*. 2017;20(3):466–73.

21. 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(3):328–40.

22. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med Care*. 2017;55(7):e51–8.

23. Shiroiwa T, Ikeda S, Noto S, et al. Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. *Value Health*. 2016;19(5):648–54.

24. Kreimeier S, Greiner W. EQ-5D-Y as a health-related quality of life instrument for children and adolescents: the instrument's characteristics, development, current use, and challenges of developing its value set. *Value Health*. 2019;22(1):31–7.