VALUE IN HEALTH 21 (2018) 1291–1298

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/jval

Methodology

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

Simone Kreimeier, MSc[1,*], Mark Oppe, PhD[2,**], Juan M. Ramos-Goñi, MSc[2], Amanda Cole, PhD[3], Nancy Devlin, PhD[3], Michael Herdman, MSc[3], Brendan Mulhern, MRes[4], Koonal K. Shah, PhD[3], Elly Stolk, PhD[2], Oliver Rivero-Arias, DPhil[5], Wolfgang Greiner, PhD[1]

[1]Bielefeld University, Faculty of Health Science, Department of Health Economics and Health Care Management; [2]Office of the EuroQol Research Foundation, Rotterdam, Netherlands; [3]Office of Health Economics, London, UK; [4]Centre for Health Economics Research and Evaluation, University of Technology, Sydney, Australia; [5]University of Oxford, UK

ABSTRACT

Background: Valuations of health states were affected by the wording of the two instruments (EQ-5D-3L and EQ-5D-Y) and by the perspective taken (child or adult). **Objectives:** There is a growing demand for value sets for the EQ-5D-Y (EQ-5D instrument for younger populations). Given the similarities between EQ-5D-Y and EQ-5D-3L, we investigated whether valuations of health states were affected by the differences in wording between the two instruments and by the perspective taken in the valuation exercise (child or adult). **Study Design:** Respondents were randomly assigned to EQ-5D-3L or EQ-5D-Y (instrument) and further into two groups that either valued health states for an adult or for a 10-year-old child (perspective). The valuation tasks were composite time trade-off (C-TTO) and discrete choice experiments (DCE), including comparisons with death (DCE + death). Members of the adult general population in four countries (Germany, Netherlands, Spain, England) participated in computer-assisted personal interviews. **Methods:** Two-way multivariate analysis of variance (MANOVA) and post hoc tests were used to compare C-TTO responses and chi-square tests were conducted to compare DCE + death valuations. **Results:** A significant interaction effect between instrument and perspective for C-TTO responses was found. Significant differences by perspective (adult and child) occurred only for the EQ-5D-3L. Significant differences in values between instruments (EQ-5D-3L and EQ-5D-Y) occurred only for the adult perspective. Both significant results were confirmed by the DCE + death results. When comparing EQ-5D-3L for adult perspective and EQ-5D-Y for child perspective, values were also significantly different. **Conclusions:** The results identified an interaction effect between wording of the instrument and perspective on elicited values, suggesting that current EQ-5D-3L value sets should not be employed to assign values to EQ-5D-Y health states.

Keywords: adolescents, children, EQ-5D-Y, health-related quality of life (HRQoL), health state values, perspective.

Introduction

In the context of economic evaluation of health care interventions, the number of cost-utility analyses in pediatric populations has increased [1,2]. However, only a few generic preference-based instruments developed for children and adolescents have value sets available to calculate utility [3,4]. There are conceptual and methodological challenges surrounding the design of valuation studies for child-specific, preference-based health related quality of life (HRQoL) measures, compared to adult preference-based measures. One challenge relates to different opinions regarding who should value pediatric health states: should it be adults from the general population or children and adolescents themselves? The arguments that the taxpayer perspective should be adopted and that completing a valuation task requires abstract thinking both speak in favour of a general population sample [2,5–7]. However, evidence also shows that preferences of children and adolescents themselves can be assessed, and that they place

* Address correspondence to: Simone Kreimeier, Bielefeld University, Faculty of Health Science, Department of Health Economics and Health Care Management, Universitätsstraße 25, 33615 Bielefeld, Germany.

** Address correspondence to: Mark Oppe, Office of the EuroQol Research Foundation, Marten Meesweg 107, 3068 AV Rotterdam, Netherlands.

E-mail: simone.kreimeier@uni-bielefeld.de

https://doi.org/10.1016/j.jval.2018.05.002

Table 1 – Wording of the EQ-5D-3L and EQ-5D-Y instruments.

EQ-5D-3L	EQ-5D-Y
Mobility	**Mobility (walking about)**
• I have no problems walking about	• I have no problems walking about
• I have some problems walking about	• I have some problems walking about
• I am confined to bed	• I have a lot of problems walking about
Self-care	**Looking after myself**
• I have no problems with self-care	• I have no problems washing or dressing myself
• I have some problems washing or dressing myself	• I have some problems washing or dressing myself
• I am unable to wash or dress myself	• I have a lot of problems washing or dressing myself
Usual activities	**Doing usual activities**
(e.g. work, study, housework, family or leisure activities)	*(for example, going to school, hobbies, sports, playing, doing things with family or friends)*
• I have no problems with performing my usual activities	• I have no problems doing my usual activities
• I have some problems with performing my usual activities	• I have some problems doing my usual activities
• I am unable to perform my usual activities	• I have a lot of problems doing my usual activities
Pain/discomfort	**Having pain or discomfort**
• I have no pain or discomfort	• I have no pain or discomfort
• I have moderate pain or discomfort	• I have some pain or discomfort
• I have extreme pain or discomfort	• I have a lot of pain or discomfort
Anxiety/depression	**Feeling worried, sad, or unhappy**
• I am not anxious or depressed	• I am not worried, sad, or unhappy
• I am moderately anxious or depressed	• I am a bit worried, sad, or unhappy
• I am extremely anxious or depressed	• I am very worried, sad, or unhappy

different weights on impairments than adults [8–14]. Apart from the question of who should provide the values there is the question of which valuation task is most appropriate to use. Choosing a valuation technique is complicated, and the framing and wording of the valuation task are important [7,15–19].

In 2010 the EQ-5D-Y was developed as a version of the EQ-5D-3L for measuring HRQoL of children and adolescents aged 8 to 15 years using the standard five-dimensional, three-level format (Table 1) [16,20]. While the instrument has been included in an increasing number of studies, it cannot be used in cost-utility analyses, because there are currently no value sets available [21,22]. The similarity between the EQ-5D-3L and the EQ-5D-Y, and the fact that more than 20 value sets are currently available for the EQ-5D-3L [23], raises the question of whether EQ-5D-3L value sets could potentially be used for the EQ-5D-Y or whether separate value sets are required.

Evidence suggests that values given to health states of children and adolescents are lower than those given to adult health states (i.e., a given health problem is considered less desirable) [24–27]. Thus, new value sets for calculating a preference-based index for the EQ-5D-Y may be necessary [15]. However, for the case of the EQ-5D-Y, this question about a separate value set has yet to be answered empirically. A prior study examined this by asking respondents to value eight EQ-5D-Y states using the visual analogue scale (VAS) from three different perspectives: 1) yourself; 2) a hypothetical adult; and 3) a hypothetical child. In all three countries (England, Germany, Spain) adults gave lower values for children than for themselves or for other adults. However, no systematic relationship was found [21]. In addition, the generalizability of the results was limited, because only a few health states were included. Further, the VAS method used did not include the valuation of "death," which is needed to allow quality-adjusted life-year (QALY) calculations. The study only included the EQ-5D-Y, and not the EQ-5D-3L, so the impact of the difference in wording between the instruments (Table 1) was not considered.

In order to establish whether EQ-5D-3L value sets can be used for the EQ-5D-Y or whether separate value sets are required, differences in both the wording of the two instruments and in the perspective adopted in the valuation task (i.e., valuing health states for an adult [for themselves] or for a child [for another young person]) are important. Differences in both instrument and perspective could lead to differences in the relative importance of dimensions and to differences in the scales of the value sets produced, necessitating separate EQ-5D-Y valuation studies. Therefore, the aim of the study was to investigate whether values elicited for health states varied when: 1) those health states were defined using the EQ-5D-3L or EQ-5D-Y descriptive systems and 2) the values were elicited for an adult or a child.

Methods

Design

Because we wanted to investigate the impact of the wording of the instrument and the impact of the perspective adopted in the valuation task, our study comprised four arms: EQ-5D-3L (adult perspective), EQ-5D-3L (child perspective), EQ-5D-Y (adult perspective), and EQ-5D-Y (child perspective). Respondents were randomly allocated to one of the four arms. We opted to elicit values from the adult general population not only to keep the traditional view of using the values of the taxpayer in this kind of valuation task, but also because including comparisons with death (required to produce a value set on the QALY scale) was deemed unsuitable for children.

Valuation Exercise

We used an adapted version of the EuroQol Valuation Technology (EQ-VT) protocol, which was developed for the valuation of the EQ-5D-5L [28] and includes a C-TTO and a DCE task. In the C-TTO, respondents are asked to compare a life in an impaired health state for 10 years to a life in full health but with a shorter duration [29]. In the DCE, respondents are asked to consider two different health states (no duration attached) and indicate which of the two they think is better [30]. We included two modifications to the EQ-VT protocol. The first was the addition of a ranking task; the second, a modification of the DCE task to also include forced choice paired comparisons of EQ-5D health states with immediate death (DCE + death) [30]. The valuation protocol included: 1) introduction; 2) background information, including respondent's own health using EQ-5D, age, sex, and experience with serious illness; 3) a ranking task, where the respondent was asked to rank the 10 dimension-level descriptions indicating some or extreme problems on each of the five dimensions (Table 1) of the EQ-5D from best (or "least bad") to worst; 4) C-TTO [29] tasks, including instructions focusing on an example health state describing requiring the use of a wheelchair, three practice states, nine tasks involving the valuation of EQ-5D health states, a feedback module to review respondents' responses, and feedback questions about task complexity; 5) DCE + death tasks, including instructions, nine paired comparisons between EQ-5D health states (A vs. B) plus each health state compared to immediate death (A vs. death followed by B vs. death), so overall 27 comparisons, and feedback questions about task complexity; 6) extended background questions, including education, employment status, experience working directly with children, experience of serious illness in children, whether the participant lives with children, and feedback questions to get insight into how the framing (age/age range of the described child, own vs. hypothetical child, taking different perspectives) might affect responses; and 7) general thank you and goodbye. Depending on the study arm, EQ-5D-3L or EQ-5D-Y was used throughout the interview.

Perspectives Taken within the Valuation Exercise

Given the two perspectives, the framing of the elicitation task differed slightly. For the adult perspective, respondents were asked to imagine themselves being in the health states when completing the tasks—the conventional approach of EuroQol valuation studies—while for the child perspective, respondents were asked: "Imagine a 10-year-old child being in the health states when completing the tasks." We used this approach, because we considered that specifying the age might help respondents imagine the child and the impact of the health state. In the DCE, where two different health profiles are presented on the screen, we asked: "Please imagine yourself living in these health profiles and choose the health profile you think is the better of the two," in the adult arm and "Please imagine a 10-year-old child living in these health profiles and choose the health profile you think is the better of the two," in the child arm.

Selection of Included Health States, Pairs, and Sample Size

We used the set of 17 health states that had been used for the valuation of the EQ-5D-3L in the Netherlands and Japan [31,32]. Those 17 states were divided into two blocks of nine health states, where both blocks included the worst state "33333," using the blocking algorithm in the R-package AlgDesign. Ten different DCE pairs, containing 19 different EQ-5D health states, were included in the study (two pairs included the worst EQ-5D-3L/EQ-5D-Y state: 33333). The DCE pairs were created in such a way that the nine states included in the first block of the C-TTO tasks were also included in the DCE + death tasks. This was done to allow direct comparisons between C-TTO and DCE with respect to whether or not a state is considered to be worse than death. The 10 states that were included in the DCE pairs, but were not taken from the TTO task, were selected using a Bayesian efficient design algorithm which minimized D-error. The priors that were used to calculate the D-errors in the design optimization algorithm to select the (pairs of) health states came from an earlier EQ-5D-3L DCE study [30]. All 19 health states included in the DCE design were also compared with immediate death. We initially created a DCE design using two blocks of nine pairs each where one of the EQ-5D states in each pair of the second DCE block matched an EQ-5D state in the second TTO block. However, due to an error, the final implementation included only the 10 pairs that contained matches with EQ-5D states from the first TTO block as described above.

Given the heteroscedasticity present in C-TTO observations (i.e., the worse the health state the larger the SD), a threshold of standard error < 0.05 for the health state with the largest SD was used to estimate sample size; using this criterion, approximately 100 observations per health state were needed to achieve this minimum level of precision for mean observed C-TTO values. This, combined with the fact that there were two blocks of health states in each of the four study arms, led to an estimated sample size of $2 \times 4 \times 100 = 800$ respondents in total. Because four countries participated and we did not expect major country-specific variation in the differences of the values between the adult and the child perspective, this implied a target sample size in each country of 200.

Data Collection

The study was approved by an ethics committee in each country. Data collection took place between May and July 2015 in four countries: Germany, Netherlands, Spain, and England). Organization of data collection differed somewhat between countries. In Germany, Spain, and the Netherlands, interviews were conducted at a central site. In England, interviews were carried out in respondents' homes. Respondents were recruited using a convenience strategy in each country. All countries monitored the sample during data collection to ensure representativeness for the country with respect to age and sex. The respondents received a shopping voucher or cash as an incentive. Computer-assisted personal interviews were used in all countries, and interviewers received intensive training and conducted pilot interviews using EQ-VT before beginning fieldwork. Interviewers were also required to follow a set protocol for the different study arms. A quality control process using the EQ-VT QC tool was implemented to ensure protocol compliance [33].

Statistical Analysis

Descriptive statistics were used to summarize the collected data. We used proportions and sample sizes to present background characteristics of the sample and DCE + death responses, although we used means, standard deviation, and box plots to present mean and confidence intervals of the C-TTO responses. All results are presented by study arms.

In order to test whether values differed across study arms, we used a 2-way MANOVA analysis for C-TTO responses. We first tested for possible interaction effects between wording of the instrument and perspective (if an interaction exists, no further exploration of wording and/or perspective should be conducted due to confounding) [34]. For post hoc comparisons, we used the Hotelling T^2 test and T-test and adjusted P values using a Bonferroni correction for one-to-one comparison of study arms and health states, respectively [34]. Given that respondents were randomized to one of two blocks for each study arm that contained different health states, C-TTO valuations were compared within each block. In total, six post hoc pair comparisons between study arms were performed (EQ-5D-3L-Adult vs. EQ-5D-3L-Child; EQ-5D-3L-Adult vs. EQ-5D-Y-Adult; EQ-5D-3L-Adult vs. EQ-5D-Y-Child; EQ-5D-3L-Child vs. EQ-5D-Y-Adult; EQ-5D-3L-Child vs. EQ-5D-Y-Child; EQ-5D-Y-Adult vs. EQ-5D-Y-Child). That yielded 12 total comparisons when accounting for the two blocks of the C-TTO design. In order to test whether observed choice probabilities in DCE + death paired comparisons differed between study arms, we used the chi-square test. For post hoc comparisons, we used the same chi-square test, but adjusted *P* values using a Bonferroni correction. The experimental designs were created using R version 3.1.0. Analyses were carried out using Stata/MP 14.

Results

The sample size was 805 respondents, 200 in the Netherlands, Germany, and England and 205 in Spain. No significant differences in sample sizes by study arm in each country were found. Age groups were evenly represented across study arms, except for age 65 and older. Overall, there were more female than male respondents in all of the study arms. Samples sizes across study arms were not significantly different for education, experience with illness, or experience working with children, whether the participant had children or whether the participant was at any point the primary caregiver of a child (Table 2).

The observed mean values for the 17 health states included in C-TTO design showed differences between study arms (Table 3). The descriptive results show that valuations based on the child perspective were, on the whole, higher than those based on the adult perspective (comparing EQ-5D-3L-Adult vs. EQ-5D-3L-Child and EQ-5D-Y-Adult vs. EQ-5D-Y-Child) (Table 3). Values for EQ-5D-3L health states were generally lower than those for corresponding EQ-5D-Y health states. Regression models of the TTO data for the four study arms can be found in supplemental online material 1.

The proportion selecting option A in the DCE or DCE + death also showed differences between study arms (Table 4). Overall, it could be observed that "immediate death" is chosen less frequently in the child perspective arm than in the adult perspective arm. When comparing the choice probabilities of EQ-5D-3L-Adult vs. EQ-5D-3L-Child and EQ-5D-Y-Adult vs. EQ-5D-Y-Child, the probability for choosing the health state instead of "immediate death" is in general higher in the child arm than in the adult arm.

MANOVA showed statistically significant interactions between wording and perspective in C-TTO responses in both blocks (P values < 0.0001). Therefore, neither wording nor perspective can be analyzed as independent factors. Post hoc paired comparisons between study arms confirmed evidence of differences in three comparisons: 1) EQ-5D-3L-Adult vs. EQ-5D-3L-Child (P value $= 0.0252$); 2) EQ-5D-3L-Adult vs. EQ-5D-Y-Adult (P value < 0.0001); and 3) EQ-5D-3L-Adult vs. EQ-5D-Y-Child (P value < 0.0001). Further post hoc comparisons to identify which health states were causing the differences indicated that

Table 2 – Background characteristics by study arm.

	Study arm 1: EQ-5D-3L Adult perspective n (%)	Study arm 2: EQ-5D-3L Child perspective n (%)	Study arm 3: EQ-5D-Y Adult perspective n (%)	Study arm 4: EQ-5D-Y Child perspective n (%)	Total n (%)
Country					
England	52 (6.5)	50 (6.2)	48 (6.0)	50 (6.2)	200 (24.8%)
Germany	49 (6.1)	45 (5.6)	49 (6.1)	57 (7.1)	200 (24.8%)
Spain	50 (6.2)	53 (6.6)	45 (5.6)	52 (6.5)	200 (24.8%)
Netherlands	54 (6.7)	47 (5.8)	52 (6.5)	52 (6.5)	205 (25.5%)
Total	205 (25.5)	195 (24.2)	194 (24.1)	211 (26.2)	805 (100.0%)
Age (groups)					
<25	28 (3.5)	33 (4.1)	31 (3.9)	34 (4.2)	126 (15.7)
25–34	43 (5.3)	34 (4.2)	44 (5.5)	34 (4.2)	155 (19.3)
35–44	40 (5.0)	39 (4.8)	32 (4.0)	37 (4.6)	148 (18.4)
45–54	45 (5.6)	40 (5.0)	38 (4.7)	31 (3.9)	154 (19.1)
55–64	34 (4.2)	30 (3.7)	27 (3.4)	43 (5.3)	134 (16.6)
65–74	13 (1.6)	14 (1.7)	13 (1.6)	23 (2.9)	63 (7.8)
>75	2 (0.2)	5 (0.6)	9 (1.1)	9 (1.1)	25 (3.1)
Sex					
Female	120 (14.9)	115 (14.3)	118 (14.7)	119 (14.8)	472 (58.6)
Male	85 (10.6)	80 (9.9)	76 (9.4)	92 (11.4)	333 (41.4)
Education*					
Low	55 (6.8)	35 (4.3)	41 (5.1)	52 (6.5)	183 (22.7)
Medium	71 (8.8)	76 (9.4)	76 (9.4)	80 (9.9)	303 (37.6)
High	78 (9.7)	84 (10.4)	77 (9.6)	79 (9.8)	318 (39.5)
Experience of serious illness					
Yourself (% Yes)	135 (16.8)	121 (15.0)	130 (16.1)	144 (17.9)	530 (65.8)
Family (% Yes)	135 (16.8)	121 (15.0)	130 (16.1)	144 (17.9)	530 (65.8)
Caring for other (% Yes)	61 (7.6)	51 (6.3)	61 (7.6)	64 (8.0)	237 (29.4)
Employed in role working directly with children (% Yes)	64 (8)	74 (9.2)	65 (8.1)	72 (8.9)	275 (34.2)
Participant has own children (% Yes)	111 (13.8)	110 (13.7)	104 (12.9)	128 (15.9)	453 (56.3)
Participant is primary caregiver of a child (% Yes)	114 (14.2)	111 (13.8)	103 (12.8)	123 (15.3)	451 (56.0)

* One response was missing in the EQ-5D-3L adult arm.

Table 3 – Values given to the 17 health states included in the composite time trade-off design by block and study arm.

Health State	Block	Study arm 1:EQ-5D-3L Adult perspective			Study arm 2:EQ-5D-3L Child perspective			Study arm 3:EQ-5D-Y Adult perspective			Study arm 4: EQ-5D-Y Child perspective		
		n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
11112	1	102	0.89	0.25	96	0.94	0.13	108	0.95	0.10	101	0.96	0.11
11133			0.23	0.68		0.24	0.71		0.30	0.65		0.23	0.67
11312			0.69	0.40		0.84	0.25		0.80	0.29		0.86	0.25
13311			0.58	0.50		0.80	0.30		0.74	0.35		0.86	0.19
21111			0.93	0.13		0.96	0.07		0.94	0.16		0.95	0.12
23232			0.18	0.65		0.25	0.64		0.39	0.56		0.31	0.66
32211			0.45	0.51		0.71	0.40		0.79	0.30		0.87	0.25
32223			0.16	0.64		0.36	0.60		0.42	0.52		0.52	0.49
11113	2	103	0.55	0.51	99	0.66	0.49	86	0.46	0.56	110	0.70	0.42
11121			0.91	0.14		0.90	0.18		0.87	0.25		0.92	0.21
11131			0.52	0.49		0.33	0.62		0.42	0.61		0.49	0.61
11211			0.94	0.09		0.94	0.16		0.90	0.23		0.95	0.21
12111			0.92	0.15		0.93	0.16		0.91	0.14		0.94	0.22
22222			0.71	0.29		0.73	0.31		0.70	0.36		0.83	0.23
32313			0.18	0.55		0.25	0.61		0.28	0.63		0.54	0.47
33323			-0.03	0.60		0.07	0.59		-0.02	0.68		0.16	0.63
33333	1 & 2	205	-0.32	0.62	195	-0.20	0.63	194	-0.17	0.65	211	-0.14	0.67

Table 4 – Choice probabilities of choosing health state A over health state B or death in the DCE by study arm.

DCE pair (A vs. B)	EQ-5D-3L-Adult %	EQ-5D-3L-Child %	EQ-5D-Y-Adult %	EQ-5D-Y-Child %
11332 vs. 22222	13.7	11.8	6.7	10.0
13213 vs. 32331	86.3	88.2	69.6	75.4
11113 vs. 11121	11.2	20.5	21.6	23.7
31212 vs. 12111	1.0	1.0	4.1	4.3
32121 vs. 11211	0.5	0.5	2.6	5.2
31231 vs. 32313	50.2	43.1	52.1	46.0
33323 vs. 21133	20.0	36.4	34.0	53.1
11131 vs. 13222	56.1	28.7	45.9	37.4
33333 vs. 23333	1.5	0.0	0.0	0.5
DCE (A vs. Death)				
11332 vs. Death	71.7	65.1	72.2	76.3
22222 vs. Death	97.6	97.9	95.4	99.1
13213 vs. Death	86.8	91.3	82.0	92.4
32331 vs. Death	57.6	52.8	34.5	32.2
11113 vs. Death	91.2	92.8	87.1	93.8
11121 vs. Death	99.5	99.5	99.5	100.0
31212 vs. Death	82.4	95.9	97.4	98.6
12111 vs. Death	99.5	100.0	100.0	100.0
32121 vs. Death	83.4	95.9	95.4	98.1
11211 vs. Death	100.0	100.0	99.5	99.5
31231 vs. Death	56.1	55.4	76.8	76.8
32313 vs. Death	53.7	74.9	64.4	80.1
33323 vs. Death	31.2	46.7	44.8	56.9
21133 vs. Death	63.9	55.4	59.3	55.9
11131 vs. Death	85.9	75.9	76.3	78.2
13222 vs. Death	94.1	95.9	95.4	98.6
33233 vs. Death	25.5	30.2	32.4	34.7
23333 vs. Death	32.0	32.3	24.4	37.3
33333 vs. Death	17.1	17.9	25.3	28.0

DCE = discrete choice experiment

Table 5 – Overview of significant differences between study arms for both mean values derived from the C-TTO and choice probabilities within the discrete choice experiments*.

C-TTO		
Compared study arms	**Health state**	**P value**
EQ-5D-3L-Adult vs. EQ-5D-3L-Child	13311	0.0081
	32211	0.0027
EQ-5D-3L-Adult vs. EQ-5D-Y-Adult	32211	< 0.0001
	32223	0.0351
EQ-5D-3L-Adult vs. EQ-5D-Y-Child	11312	0.0135
	13311	< 0.0001
	32211	< 0.0001
	32223	< 0.0001
	22222	0.0216
	32313	< 0.0001
DCE and DCE + death		
Compared study arms	**Pair**	**P value**
EQ-5D-3L-Adult vs. EQ-5D-3L-Child	11131 vs. 13222	0.00841
EQ-5D-3L-Adult vs. EQ-5D-Y-Adult	32121 vs. 11211	0.0003
	33323 vs. 21133	< 0.0001
	32313 vs. Death	0.0004
	33323 vs. Death	0.0346
EQ-5D-3L-Child vs. EQ-5D-Y-Child	32121 vs. 11211	0.0007

C-TTO = composite time trade-off; DCE + death = discrete choice experiment, including death.

* This table contains only those states (mean of C-TTO results) and state pairs (choice probabilities of DCE + death results) that showed significant differences by study arm. The states and state pairs that are not mentioned in the table showed no significant differences between study arms.

neither health states at the top of the scale (mild states: 11112, 11121, 11211, 12111, 21111) nor at the bottom (severest state 33333) were significantly different between study arms. Significant differences between C-TTO valuations over health states by study arm are presented in Table 5, together with significant differences in DCE choice probabilities. There were no statistically significant differences between the four arms with respect to the value for health state "33333" either for the study population as a whole (Fig. 1) or when analyzed by country. Further analyses by country showed that, in the Netherlands, values for health states tended to be higher when valued from the adult perspective rather than the child perspective, which was the opposite of the result observed in the three other countries. In addition to the presented findings, results regarding the complexity of TTO and DCE tasks and the sort of child respondents imagined during the interview are shown in supplemental online material 2.

Discussion

Our results showed that participants from the adult general population placed different values on some health states (see Table 5) depending on the wording of the instrument (EQ-5D-3L vs. EQ-5D-Y) and the perspective (adult vs. child) used. However, we did not find significant differences for either wording or perspective regarding state 33333 in either C-TTO or DCE + death in any of the four countries. We therefore do not have evidence to suggest that the scale (i.e., the distance between 11111 and 33333) of a hypothetical country-specific EQ-5D-Y value set should be different from the scale of the EQ-5D-3L value set of the same country (when the same elicitation procedure is used). However, both our study and the study by Kind et al. found evidence that differences do exist in (at least some of) the other health states [21]. This leads us to postulate that the relative importance of the domains and levels differs for both wording of the instrument and perspective. However, the impact on scale needs further research. We can therefore conclude that separate value sets for the EQ-5D-Y instrument are needed.

Overall, values for EQ-5D-3L states were lower than those for corresponding EQ-5D-Y states. It seems as if the differences based on the wording of the instrument are larger in health states describing level 3 problems (i.e., the worst response category; see Table 1) in mobility, self-care or usual activities compared to health states with a lower severity level in those dimensions. This can be observed in the C-TTO values as well as in the DCE + death results. The explanation for that might be the

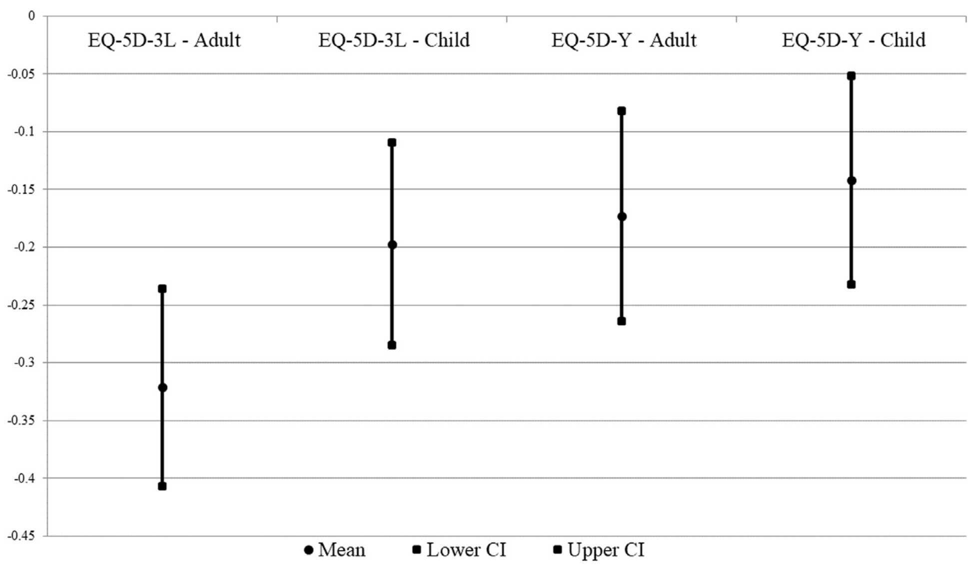

Fig. 1 – Box plot of composite-time trade-off (C-TTO) responses for health state 33333 by study arm in the total sample.

way in which level 3 problems in the dimensions are worded in EQ-5D-3L compared to EQ-5D-Y. Wording that describes a more severe limitation of health is used in EQ-5D-3L; for example, the EQ-5D-3L wording for the level 3 in mobility is "confined to bed," although the wording in the EQ-5D-Y is "a lot of problems walking about" (see Table 1). Our results contrast with those reported by Kind et al. While Kind et al. found that in general respondents gave lower VAS ratings for the child perspective than for the adult perspective [21], we found that in general they gave higher C-TTO values for the child perspective than for the adult perspective. Compared to the Kind et al. study, our study included more health states, used different valuation techniques, and considered the differences in wording between the EQ-5D-Y and the EQ-5D-3L. It is clear that the valuation technique has an impact on the results. While both the C-TTO and DCE + death techniques specifically included direct comparison with (immediate) death, the VAS method used by Kind et al. did not. This fact may explain the different directions of the results between studies. We conjecture that the inclusion of shortened lives in the C-TTO valuation task and death in the DCE led to a stronger preference for longer lives when done for children (even if they are impaired), resulting in higher values for child health states relative to adult health states in our study. Given that the QALY model requires the utility scale to be anchored on full health and death, caution is warranted when EQ-5D-Y values and EQ-5D-3L values, anchored on death, are being compared for adults and for children, because we are currently not able to account for differences in the perception of this anchor between the two groups. More research on valuation methods using the perspective of a (10-year-old) child that address the anchoring of values to "death" is necessary.

The C-TTO results suggest that the trade-off between time and HRQoL made by participants for children seems to be different from the equivalent trade-off for adults. There seems to be a shift in preferences of HRQoL and length of life when there is a change in the delineated target group. It might therefore be impossible, or at least difficult, to use the same QALY league table and the same cost-per-QALY decision-making threshold to value interventions for children and adults and to use them in information decisions about the distribution of resources between the two groups. Different league tables and cost-per-QALY thresholds might be necessary.

The results of our study were limited by the fact that only one of the two blocks of the DCE + death tasks was included in the study due to a technical error. Because of this we could only investigate the impact of the wording of the instruments and perspective for 19 health states in the DCE + death and compare differences between C-TTO and DCE + death for nine health states. Additionally, we could not estimate DCE + death models to compare those to the C-TTO models. A second limitation was that differences in our data across countries, due to unobservable characteristics and different recruitment strategies in the countries, may affect the results from the pooled data. However, the power was not sufficient to explore country-specific analysis. Finally, the fact that in the adult perspective respondents valued their own health and in the child perspective they valued another person's health could also have influenced our results.

Conclusion

The wording of the instrument and the perspective adopted in the valuation task affect health state values elicited in a general population sample. Therefore, EQ-5D-3L value sets should not be applied to EQ-5D-Y health states, and specific value sets for the EQ-5D-Y are necessary. EQ-5D-Y values can be derived on a QALY scale, but differences between health state values for adults and children should be interpreted with care since the anchors of the QALY scale appear to be interpreted differently across those groups.

Acknowledgments

The authors want to express their gratitude to the respondents who completed the interviews. The authors want to express their special gratitude to Arnd Jan Prause for his constant and useful technical support.

Research was funded by a research grant (reference number: 2014160) awarded by the EuroQol Research Foundation.

The study reported in the paper was already presented at the 1) 32nd EuroQol Plenary Meeting 2015 in Krakow, Poland; and at the 2) EuHEA Conference 2016 in Hamburg, Germany.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2018.05.002.

REFERENCES

[1] Kromm SK, Bethell J, Kraglund F, et al. Characteristics and quality of pediatric cost-utility analyses. Qual Life Res 2012;21:1315–25.
[2] Thorrington D, Eames K. Measuring Health Utilities in Children and Adolescents: A Systematic Review of the Literature. PLoS One 2015;10: e0135672.
[3] Chen G, Ratcliffe J. A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. Pharmacoeconomics 2015;33:1013–28.
[4] Solans M, Pane S, Estrada MD, et al. Health-Related Quality of Life Measurement in Children and Adolescents: A Systematic Review of Generic and Disease-Specific Instruments. Value Health 2008;11:742–64.
[5] Gold MR, Siegel JE, Russell LB, et al. Cost-Effectiveness in Health and Medicine. New York, NY: Oxford University Press, 1996.
[6] National Institute for Health and Clinical Excellence (NICE). Guide to the Methods of Technology Appraisal. 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword. [Accessed May 18, 2018].
[7] Noyes J, Edwards RT. EQ-5D for the Assessment of Health-Related Quality of Life and Resource Allocation in Children: A Systematic Methodological Review. Value Health 2011;14:1117–29.
[8] Ratcliffe J, Huynh E, Stevens K, et al. Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best–Worst Scaling. Health Econ 2015;25:486–96.
[9] Wassermann J, Aday LA, Begley CE, et al. Measuring health state preferences for hemophilia: development of a disease-specific utility instrument. Haemophilia 2005;11:49–57.
[10] Saigal S, Stoskopf B, Feeny D, et al. Differences in Preferences for Neonatal Outcomes Among Health Care Professionals, Parents, and Adolescents. Journal of the American Medical Association 1999;281:1991–7.
[11] Ratcliffe J, Flynn T, Terlich F, et al. 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;30:713–27.
[12] Ratcliffe J, Couzner L, Flynn T, et al. Valuing Child Health Utility 9D health states with a young adolescent sample: A feasibility study to compare best-worst scaling discrete-choice experiment, standard gamble and time trade-off methods. Appl Health Econ Health Policy 2011;9:15–27.
[13] Crump RT, Beverung LM, Lau R, et al. Reliability, Validity, and Feasibility of Direct Elicitation of Children's Preferences for Health States. Med Decis Making 2017;37:314–26.
[14] Yi MS, Britto MT, Wilmott RW, et al. Health values of adolescents with cystic fibrosis. J Pediatr 2003;142:133–40.
[15] Jelsma J, McKenzie J, Rama L. Differences in subjective construction of health-related quality of life (as described by the components of the EQ-5D) for adults and children. Proceedings of the 28th Scientific Plenary Meeting of the EuroQol Group. Rotterdam: EuroQol Research Foundation, 2011.

[16] Ravens-Sieberer U, Wille N, Badia X, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. Qual Life Res 2010;19:87–897.
[17] Ratcliffe J, Couzner L, Flynn T, et al. Valuing Child Health Utility 9D health states in a young adolescent sample, Working Paper No. 2009/10. 2009;10:1–22. Available from: https://www.flinders.edu.au/centres-files/fccchcr/documents/Working%20Paper%202009_10.pdf. [Accessed May 18, 2018].
[18] Griebsch I, Coast J, Brown J. Quality-adjusted life-years lack quality in pediatric care: a critical review of published cost-utility studies in child health. Pediatrics 2005;115:e600–8.
[19] Dolan P, Olsen JA, Menzel P, Richardson J. An inquiry into the different perspectives that can be used when eliciting preferences in health. Health Econ 2003;12:545–51.
[20] Wille N, Badia X, Bonsel G, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res 2010;19(6):875–86.
[21] Kind P, Klose K, Gusi N, et al. Can adult weights be used to value child health states? Testing the influence of perspective in valuing EQ-5D-Y. Qual Life Res 2015;24:2519–39.
[22] EuroQol. EQ-5D-Y User Guide. Basic information on how to use the EQ-5D-Y instrument. Version 1.0. Rotterdam: EuroQol Research Foundation, 2014. Available from: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-Y_User_Guide_v1.0_2014.pdf. [Accessed May 18, 2018].
[23] EuroQol. EQ-5D-3L Valuation. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/valuation/. [Accessed May 18, 2018].
[24] Tsuchiya A. The Value of Health at Different Ages. Discussion Paper 184. 2001;184:1–15. Available from: http://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20184.pdf. [Accessed May 18, 2018].
[25] Tsuchiya A. Age-related preferences and age weighting health benefits. Soc Sci Med 1999;48:267–76.
[26] Busschbach JJ, Hessing DJ, de Charro FT. The utility of health at different stages in life: a quantitative approach. Soc Sci Med 1993;37:153–8.
[27] Lewis PA, Charny M. Which of two individuals do you treat when only their ages are different and you can't treat both? J Med Ethics 1989;15:28–34.
[28] Shah K, Rand-Hendriksen K, Ramos JM, et al. Improving the quality of data collected in EQ-5D-5L valuation studies: a summary of the EQ-VT research methodology programme. Proceedings of the 31st Scientific Plenary Meeting of the EuroQol Group. Rotterdam: EuroQol Research Foundation, 2014.
[29] Oppe M, Devlin N, 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.
[30] Stolk EA, Oppe M, Scalone L, Krabbe PF. Discrete Choice Modeling for the Quantification of Health States: The Case of the EQ-5D. Value Health 2010;13:1005–13.
[31] 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.
[32] 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.
[33] Ramos-Goñi JM, Oppe M, Slaap B, et al. Quality Control Process for EQ-5D-5L Valuation Studies. Value Health 2017;20:466–73.
[34] Mardia KV, Kent JT, Bibby J. Factor Analysis. In: Mardia KV, Kent JT, Bibby J,eds. Multivariate analysis. London, UK: Academic Press, 1979.