



Estimating an EQ-5D-Y-3L Value Set for China

Zhihao Yang^{1,2} · Jie Jiang² · Pei Wang³ · Xuejing Jin⁴ · Jing Wu⁵ · Yu Fang⁶ · Da Feng⁷ · Xiaoyu Xi⁸ · Shunping Li⁹ · Mingxia Jing¹⁰ · Bin Zheng¹¹ · Weidong Huang¹² · Nan Luo¹³

Accepted: 1 November 2022 / Published online: 18 November 2022
© The Author(s) 2022

Abstract

Introduction The standard EQ-5D-Y-3L valuation protocol applies DCE data as the primary preference source to model the relative importance of dimensions while cTTO data served to anchor the DCE coefficients onto the QALY scale. This study aims to estimate an EQ-5D-Y-3L value set for China following this protocol, but with a larger cTTO design to better understand the role of cTTO data in estimating EQ-5D-Y-3L value sets.

Methods In total, 150 choice sets and 28 EQ-5D-Y-3L health states were valued using DCE and cTTO methods with two independent samples, respectively. General public from 14 different regions were recruited using quota sampling method to achieve representativeness. We compared two modelling strategies: (1) fit the DCE data with mixed logit model with correlated coefficients and a subsequent mapping procedure for anchoring; (2) fit the DCE and TTO data jointly in a hybrid model. Two evaluation criteria (1) coefficient significance and monotonicity; (2) prediction accuracy of the observed cTTO values were used to select the value set.

Results In total, 1476 individuals participated in the study, with 1058 participated the DCE interview and 418 participated the cTTO interview. The highest mean TTO value was 0.924 for state 11112 and the lowest mean TTO value was – 0.088 for state 33333. The hybrid model with an A3 term performed the best and was selected as the value set.

Discussion Following the international protocol and using a larger cTTO design, this study established the EQ-5D-Y-3L value set using a hybrid model for China. Future EQ-5D-Y-3L valuation study could consider using a larger cTTO design for estimating the value set.

Key Points for Decision Makers

This study estimated an EQ-5D-Y-3L value set for China using a representative sample from 4 different regions.

Two types of preference data (composite time trade-off and discrete choice experiment) were collected, and different modelling approaches were explored.

A hybrid model with 10 main-effects parameters and an additional parameter accounting for the effect of having severe problems on all dimensions was chosen as the value set.

1 Introduction

As the most-used health-related quality of life (HRQoL) instrument worldwide including China, EQ-5D has been employed for varying purposes including monitoring population health [1, 2], quantifying disease burden [3, 4] and, most prominently, estimating health utility values for economic evaluation of health interventions and technologies [5]. For the purpose of estimating health utility values, EQ-5D value sets have been established in China [6, 7]. In 2009, the child-friendly version of EQ-5D was developed by the EuroQol Group to cope with the growing need to measure the HRQoL of children and adolescents [8]. This newly developed instrument has been psychometrically validated for different health conditions in China [9–12]. To date, there are no value sets available for the EQ-5D-Y-3L in China, limiting the use of this instrument for calculating health utility values.

Although the EQ-5D-Y-3L retained the five core health dimensions and three response levels from its adult

✉ Zhihao Yang
Zhihao_yang_cn@126.com

Extended author information available on the last page of the article

counterpart, the EQ-5D-3L, the descriptive system of EQ-5D-Y-3L, including both the health dimensions and the response levels, was altered [8]. Thus, the EQ-5D-3L value set could not be applied to calculate EQ-5D-Y-3L health state values [13, 14]. This prompted a methodological studies research programme in search of the most appropriate method to establish EQ-5D-Y-3L value sets [13, 15, 16]. In 2020, the EuroQol Group published the international EQ-5D-Y-3L valuation protocol [17].

In the protocol, discrete choice experiment (DCE) data are used as the primary health preference data to model the relative importance of the five health dimensions, and the composite time trade-off (cTTO) data are used to anchor the DCE modelling results onto the quality-adjusted life year (QALY) scale [17]. Use of DCE data has the advantages of better data collection feasibility and avoidance of the comparison with dead in the task [17]. In comparison, a concern with respect to the use of the cTTO method for valuing children's health is that the task can be upsetting and abhorrent to some respondents, posing potential ethical issues. In addition, cTTO values were found to be high for EQ-5D-Y-3L states, especially for mild/moderate health states, which could lead to insensitivity and hamper the modelling process [17]. Adopting a taxpayer's perspective [17] and considering the possible challenges of involving adolescents in valuation studies [18], the protocol aims to elicit only adult respondents' preferences using both the DCE and cTTO methods. This indicates that in EQ-5D-Y-3L valuation studies, participants will value health for someone else, but not for themselves. Specifically, participants are asked to provide valuations for a hypothetical 10-year-old child, without specifying the relationship between this hypothetical child and the participant. This perspective change, and a preference for longer life over quality of life when valuing children compared with valuing adults, are found to be the two major reasons for high values in the cTTO task [19, 20].

To date, six countries have published their EQ-5D-Y-3L value sets following the international EQ-5D-Y-3L valuation protocol [21–26]. All six studies used DCE methods for determining relative preferences and cTTO values for anchoring, though the choices of models and anchoring methods differed. An important finding which emerged from these studies was that the worst possible state (33333) had a much lower value than the other states, creating a gap at the lower end of the valuation space. This questions the appropriateness of using 33333 to anchor DCE values, given that it is the outlier state. Alternative options available are the mapping method (i.e. regressing DCE latent values onto observed cTTO values, and using the regression coefficient to rescale the DCE coefficients) and hybrid modelling (i.e. modelling cTTO and DCE data simultaneously), which have been widely used in published EQ-5D-5L value sets [27, 28].

Given the increasing need for a local EQ-5D-Y-3L value set for China's health technology assessment (HTA) projects, this study aims to estimate an EQ-5D-Y-3L value set based on the Chinese general public's health preferences. To identify the best approach to establish the value set, we compared mapping (the most-used approach) with hybrid modelling. We expanded the cTTO design since both approaches could benefit from collecting cTTO values for more health states [29]. In addition, using a larger cTTO design allowed us to evaluate the gap effect observed in previous studies that only collected cTTO values for 10 health states.

2 Methods

Our study followed the published international EQ-5D-Y-3L valuation protocol but expanded the original cTTO design of 10 health states with an orthogonal design of 18 health states, resulting in 28 health states in total. The study was approved by Jinan University ethics committee (Approval letter number: JNUKY-2021-002). In this section, we describe the methodological details used to collect two types of valuation data and the model evaluation process.

2.1 Experimental Design

In the protocol of the EQ-5D-Y-3L valuation study, different valuation methods served different purposes. The purpose of the cTTO is to provide anchors for the latent DCE values, and the DCE is the primary method to estimate the relative importance of dimensions/levels. Hence, the sample size for the DCE component was larger than that of the cTTO [17]. Following this consideration, the study used two independent samples to collect cTTO data ($n = 400$) and DCE data ($n = 1000$). Note that the sample size of the cTTO component was larger than the recommended sample size in the protocol, as we included more health states for the cTTO design.

Overall, there were 150 choice sets and 28 EQ-5D-Y-3L health states to be valued using DCE and cTTO methods, respectively. The DCE design was a Bayesian efficient design featuring a two-dimension overlap, a good level spread (i.e. each response level of each dimension appearing an approximately similar number of times in the study design), and utility balance [15]. Dimension overlap has been demonstrated to be an efficient method to reduce attribute non-attendance [30]. The design was constructed using a two-step procedure by first collecting data using a design with a sample of 127 participants, and then incorporating the collected information as priors to derive the final design. In total, there were 150 choice sets divided into 10 blocks, with each block containing 15 choice sets. This design did

not include a dominant task or test–retest task to evaluate data quality.

The cTTO design included the 10 health states from the original protocol and another 18 states from an orthogonal design, which exhibited good predictive performance in estimating an EQ-5D-3L value set [29]. All 28 states were divided into three blocks of 10 states, with state 33333 included in every block. The total number of health states and observations per health state permitted the estimation of a value set based only on TTO values.

2.2 Sampling Strategy and Participant Recruitment

In addition to the experimental design aspects, another key consideration was to recruit a representative sample for China. We used the quota sampling method to ensure the sample represented the Chinese general population in terms of gender, age, education attainment and registered residence area (rural/urban, also known as *hukou*). Quotas for these four criteria were set based on the 6th National Population Census of China [31]. In total, 14 provinces/cities were selected to cover five different geographical parts of China, including North China (Beijing, Tianjin, Heilongjiang, Shandong), East China (Shanghai, Jiangsu), South China (Guangdong, Fujian), Central China (Hubei), and West China (Guizhou, Chongqing, Yunnan, Xinjiang, Shanxi).

Given the target sample size of 1400 participants ($n = 1000$ for DCE and $n = 400$ for cTTO), 100 participants were recruited from each province/city. In Tianjin, Shanghai, Jiangsu, Guangdong, Hubei, Guizhou, Chongqing and Shanxi, 50 participants completed the cTTO interviews and another 50 participants completed the DCE interviews. In other regions, 100 participants completed the DCE interviews. We used non-probability sampling methods to recruit participants, including both snowball sampling and purposive sampling methods. The interviewers first recruited respondents from their acquaintances and then rolled out to others via word of mouth. In each recruitment province/city, a local researcher monitored the quota fulfilment and aided interviewers in recruiting hard-to-reach participants. Individuals who (i) were Chinese citizens and had lived in China over the past 5 years, (ii) gave informed consent, (iii) met the four quota criteria, and (iv) did not participate in any earlier EQ-5D valuation studies, were recruited.

2.3 Data Collection Procedures

Both cTTO and DCE data were collected using the EQ-VT platform. Consenting participants were invited to a face-to-face, one-on-one computer-assisted personal interview. There were four sections in each interview: (1) the interviewer explained the study and obtained informed consent;

(2) consenting participants provided their demographic information and reported their health using EQ-5D-Y-3L; (3) participants completed either a block of cTTO tasks or a block of DCE tasks. For the cTTO interview, each participant completed five practice states (two wheelchair examples and three practice EQ-5D-Y-3L states), before valuing formal EQ-5D-Y-3L states; (4) participants responded to debriefing questions, rating the difficulty of the valuation task and providing information related to their children, if any.

Two training workshops were arranged prior to the data collection. For the interviewers who conducted the cTTO data collection ($n = 8$, one interviewer per region), we invited all interviewers to participate in a 2-day training and practice workshop. The workshop covered the study background and design, providing the interviewers with opportunities to perform practice interviews with one another. Post-training, each interviewer conducted two rounds of five practice interviews. Quality control (QC) reports were prepared, and feedback was provided to the interviewers after each round (5–10 interviews per interviewer) following the EuroQol QC protocol [32]. Overall, the QC process evaluated the protocol compliance of the interviewers using four criteria (sufficient time to explain the example task, completing the form task, whether lead-time TTO was used, and whether inconsistent responses occurred), and further evaluated the presence of interviewer effects using time, number of moves, and value distributions of the cTTO interview data. For the interviewers who conducted the DCE data collection, a 2-hour online training workshop was arranged. QC control was not implemented for the DCE data collection process.

2.4 Data Analysis and Model Evaluation

There is no established consensus on how to estimate the EQ-5D-Y-3L value set [18]. Our study design allowed us two options: (1) use the DCE modelling results and a subsequent anchoring procedure to estimate the value set, or (2) use a hybrid model to model DCE and cTTO data jointly. All published studies that followed the standard protocol [17] applied the former approach due to the lack of health states in the cTTO design for modelling a main-effects model [22, 27]. It should be noted that both approaches have their limitations, as the former approach only uses the cTTO values for anchoring purposes, and the latter approach only supports the use of a conditional logit model for modelling DCE data.

For the first approach, we decided to use the mixed logit model (with random correlated coefficients) that accounted for preference and scale heterogeneity [15]. For anchoring, we opted for the mapping method, given the discontinuity of cTTO values for the worst state 33333 that has been reported

in published studies. Thus, for the first approach, we initially fitted the DCE data with a mixed logit model and calculated the latent values for all health states. Next, we regressed the DCE latent values onto the observed cTTO values using an ordinary least squares (OLS) model and obtained the rescale coefficient. Finally, the DCE mixed logit modelling coefficients were rescaled [22]. For the second approach, the hybrid modelling, the ‘hyreg’ function of Stata [33] was used to model the DCE and cTTO data (accounting for heteroscedasticity) jointly [27]. For details on installing and implementing the ‘hyreg’ function, see Ramos-Goñi et al. [27].

For both approaches, a main-effects only model was first used (see Appendix 1 in the electronic supplementary material for the model equations). In an EQ-5D-Y-3L main-effects model, health state utility values were explained by 11 variables, one intercept, and two dummy variables to represent the move from level 1 in any dimension to level 2 and level 3. Having observed a discontinuity of cTTO values between the worst state 33333 and other states in previous studies [20, 21, 23], we explored the

use of an additional term ‘All 3’ (A3) to capture this gap between state 33333 and other states. Note that the A3 term was only tested in the hybrid model, since 33333 was not included in the DCE design and could not be modelled. We compared the performance of these two approaches/three models using two criteria: (1) coefficient significance and monotonicity; and (2) prediction accuracy of the observed cTTO values. Prediction accuracy was measured by computing the mean absolute error (MAE) between the predicted values and the observed mean values for the 28 health states.

3 Results

Data collection was conducted between January 2020 and October 2021. In total, 1476 individuals participated in the study, of whom 1058 participated in the DCE interviews, and 418 in the cTTO interviews. Table 1 reports demographic information for the sample. Based on the

Table 1 Sample demographic characteristics

	6th National Population Census	Whole sample, <i>N</i> = 1476	DCE, <i>n</i> = 1058	cTTO, <i>n</i> = 418
Gender	Female (48.81%)	721, 51.15%	519, 49.05%	202, 48.33%
	Male (51.19%)	755, 48.85%	539, 50.95%	216, 51.67%
Age group	18–29 (25.70%)	385, 26.08%	277, 26.18%	108, 25.84%
	30–39 (20.42%)	287, 19.44%	205, 19.38%	82, 19.62%
	40–49 (21.85%)	322, 21.82%	229, 21.64%	93, 22.25%
	50–59 (15.19%)	223, 15.11%	162, 15.31%	61, 14.59%
	> 60 (16.84%)	259, 17.55%	185, 17.49%	74, 17.70%
Residency type ^a	Urban (66.67%)	950, 64.36%	681, 64.37%	269, 64.35%
	Rural (33.33%)	524, 35.50%	375, 35.44%	149, 35.65%
Education level	Primary (33.75%)	473, 32.05%	335, 31.66%	138, 33.01%
	Junior high (41.70%)	605, 40.99%	439, 41.49%	166, 39.71%
	Senior high (15.02%)	212, 14.36%	153, 14.46%	59, 14.11%
	University and above (9.53%)	186, 12.60%	131, 12.38%	55, 13.16%
Health insurance	Urban employee	522, 35.37%	361, 34.12%	161, 38.52%
	Residents	819, 55.49%	604, 57.09%	215, 51.44%
	Commercial	42, 2.85%	30, 2.84%	12, 2.87%
	Others	48, 3.25%	36, 3.40%	12, 2.87%
	No insurance	45, 3.05%	27, 2.55%	18, 4.31%
No. of children	0	384, 26.02%	275, 25.99%	109, 26.08%
	1	626, 42.41%	441, 41.68%	185, 44.26%
	2	358, 24.25%	262, 24.76%	96, 22.97%
	3	83, 5.62%	62, 5.86%	21, 5.02%
	> 4	25, 1.69%	18, 1.70%	7, 1.67%
Self-reported EQ-VAS	Mean, SD	86.04, 10.39	86.51, 10.43	84.87, 10.20

Gender, age group, residency type and education level were used for quota criteria for sampling cTTO composite time trade-off, DCE discrete choice experiment

^aTwo participants from the DCE arm reported as ‘unknown’

6th National Population Census (31), our sample was representative.

3.1 Data Descriptions

In total, 418 respondents participated in the cTTO interviews. On average, respondents took 35.70 minutes (SD 12.42) and 13.21 moves (SD 9.00) to complete an interview.

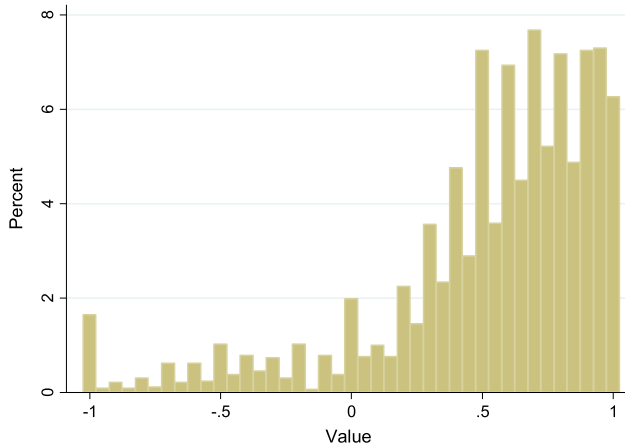


Fig. 1 Composite time trade-off (cTTO) value distribution

Around 21.89% of observations ($n = 915$) were negative values. Figure 1 shows the distribution of the cTTO values. Overall, cTTO values were distributed primarily at the higher end of the scale (≥ 0.5), with a long tail towards negative values and a ‘spike’ at -1.0 (1.65%). The highest mean TTO value was 0.924 (SE 0.011) for state 11112, and the lowest mean TTO value was -0.088 (SE 0.025) for state 33333. State 33333 was the only state with a negative observed mean value. In total, 1058 participants completed the DCE interviews, with an average duration of 9.61 minutes (SD 7.12).

3.2 Value Set Modelling

Table 2 presents the modelling coefficients and model characteristics of the three models, with an additional OLS model as a reference. In both the mixed effects model with random correlated coefficients and the hybrid main-effects only model, the sc2 term (corresponding to the effect of level 2 problems in self-care) was not significant at the 0.05 level. The rest of the coefficients were consistent and significant in all three models. In the hybrid model with the A3 term, the term was statistically significant, and the coefficient magnitude was larger than the sc3 term (corresponding to the effect of level 3 problems in self-care). In addition,

Table 2 Modelling coefficients and performances

	OLS model with A3 term (for reference)	Hybrid model with A3 term	Hybrid main effects model	Mixed effect, on latent scale	Mixed effects, anchored (coefficients only)
Model coefficients, β (SE)					
mo2	0.055 (0.014)	0.050 (0.005)	0.049 (0.005)	0.282 (0.062)	0.040
mo3	0.145 (0.014)	0.183 (0.005)	0.186 (0.005)	1.357 (0.071)	0.190
sc2	0.030 (0.014)	0.011 (0.005)	0.010* (0.005)	0.017* (0.054)	0.002
sc3	0.124 (0.014)	0.127 (0.005)	0.130 (0.005)	1.109 (0.067)	0.155
ua2	0.025* (0.014)	0.046 (0.005)	0.045 (0.005)	0.303 (0.050)	0.042
ua3	0.146 (0.015)	0.170 (0.006)	0.175 (0.005)	1.183 (0.058)	0.166
pd2	0.070 (0.014)	0.078 (0.004)	0.079 (0.004)	0.682 (0.054)	0.096
pd3	0.263 (0.014)	0.267 (0.008)	0.275 (0.008)	1.811 (0.071)	0.254
ad2	0.033 (0.014)	0.061 (0.004)	0.061 (0.005)	0.593 (0.053)	0.083
ad3	0.173 (0.014)	0.172 (0.006)	0.177 (0.006)	0.992 (0.059)	0.139
A3	0.192 (0.023)	0.158 (0.028)			
constant	0.046 (0.015)	0.012 (0.003)	0.011 (0.003)		
Model performances					
MAE	0.040	0.030	0.035		0.051
Inconsistency	0	0	0	0	
Non-significant coefficients	1	0	1	1	
Value of 12111	0.924	0.977	0.989		0.998
Value of 33333	-0.089	-0.089	0.046		0.097

MAE mean absolute error, OLS ordinary least squares

*Coefficient not significant at 0.05 level

after including the A3 term, the hybrid model resulted in the lowest MAE results and predicted a negative value for state 33333. Applying the two model evaluation criteria, the hybrid model with the A3 term outperformed both the mixed effects logit model with correlated coefficients and the main effects hybrid model and was therefore chosen to be the value set.

4 Discussion

Using two different preference elicitation methods, our study collected health preference data from two general Chinese population samples encompassing 14 different geographical regions. A strength of our study is that we used an expanded cTTO design and explored the use of the hybrid model, which has not been explored in any published EQ-5D-Y-3L valuation study. The expanded cTTO design allowed us to investigate the gap effect in EQ-5D-Y-3L cTTO values. After fitting the preference data into different models, we recommend using the hybrid model with an additional A3 term to generate the EQ-5D-Y-3L value set for China. This model made the most use of the collected data, reported the lowest prediction errors, and had no inconsistent or insignificant coefficients. In addition, the model predicted a negative value for state 33333, which was observed in the cTTO data but was not achieved with the other two models. The availability of this value set can facilitate economic evaluations of paediatric interventions in China.

Using the coefficients of level 3 problems as an indicator, the mixed effects logit model and the hybrid model resulted in a different rank order for the five dimensions. In both models, the top and second most important dimensions were pain/discomfort and mobility, respectively; however, the worried/sad/unhappy dimension was the least important in the mixed effects logit model and ranked third in the hybrid model with an A3 term. This may suggest that respondents focused on different aspects of the health states in the DCE and cTTO tasks, and that the time component may have also played a role in driving people's preferences for the health states [34]. The DCE method does not require participants to trade-off between life years and quality of life, while the cTTO method elicits values by requiring participants to shorten life expectancy in exchange for better health. It is possible that during valuation, respondents focused on the overall severity in the cTTO tasks but shifted focus to the dimension differences in the DCE tasks. It should be noted that the current cTTO method may be suboptimal for valuing EQ-5D-Y-3L states that are considered worse than dead (WTD). In other words, when a 10-year lead time is used in the context of valuing health states for a 10-year-old child, a 10-year duration of full health lead time and a 10-year duration of being in the valued states means that the state being

valued occurs at ages 20–30 years. Theoretically, a lag-time TTO method [16, 35, 36], in which the hypothetical life B in the WTD task starts with an impaired health state for 10 years and is followed by a full health state for another 10 years, is more appropriate. Thus far, only Shah et al. have investigated the use of lag-time TTO in valuing EQ-5D-Y-3L state 33333 [16]. Future studies should investigate the potential of using lag-time TTO in valuing a set of EQ-5D-Y-3L health states.

In line with the consensus of not solely using the value of 33333 for anchoring purposes [18], a value gap between state 33333 (−0.088) and the second worst health state 33323 (0.219) was observed in this study, as in other published EQ-5D-Y-3L valuation studies [21–26]. For example, in the Japanese EQ-5D-Y-3L value set, although all observed values were positive, a value gap of 0.28 existed between 33333 and the second worst state 33232 [22]. This value gap implies that a simple additive main-effects model may not perform well when state 33333 has been included in an EQ-5D-Y-3L valuation study. Hence, anchoring the latent DCE values using the cTTO value of 33333 could ensure the accuracy of the value range, but at the cost of reduced predictive accuracy for the other states. In this study, the use of the A3 term in the hybrid model successfully reduced the overall prediction error and predicted a negative value for state 33333. However, the current DCE design does not allow the exploration of this gap effect, as the DCE design does not include 33333. This could be a disadvantage of any model that only uses the DCE data (e.g. the mapping approach tested in this study). To address this, one could expand the DCE design with a pair including state 33333. This design allows the exploration of the gap effect in DCE data and facilitates data quality assessment, given that 33333 is the worst state in the EQ-5D-Y-3L descriptive system and any pair including state 33333 is a dominant pair [37]. Future research is required to understand the reasons for this value gap.

Unlike the cTTO value distributions observed in most EQ-5D studies [6, 7], there was not a strong sign of censoring at −1 in our study. This may indicate the unwillingness to trade off all the life years of a child for better health [19, 20] and subsequently led to high cTTO values. The high values raise concerns about the sensitivity of the resultant EQ-5D-Y-3L values [17]. However, the higher EQ-5D-Y-3L values mean life-saving interventions for children will gain more QALYs than life-saving interventions for adults. In fact, no value set can favour both quality-improving and life-saving interventions over another value set. Hence, children and adolescents will not always be disadvantaged if EQ-5D-Y-3L values are used together with EQ-5D-3L or -5L values to inform resource allocation decisions. Before the implications of the systematically different EQ-5D values for

adult and children health states are fully understood, efforts to make EQ-5D-Y-3L values lower may not be tenable.

It appears that the values for EQ-5D-Y-3L state 33333 in Asian countries (China: -0.088 ; Japan: 0.200 [22]; Indonesia: -0.086) have been significantly higher than those elicited in European countries (Slovenia: -0.691 [23]; Germany: -0.260 [21]; Spain: -0.389 [24]; Hungary: -0.517 [26]; the Netherlands: -0.372 [25]), implying cultural differences in the valuation of child health states. However, this pattern was not evident for EQ-5D-5L valuation studies, that is, the predicted value of 55555 was -0.341 for China [7], -0.019 for Japan [38], -0.719 for Indonesia [28], -0.166 for Spain [27], -0.462 for Germany [39], -0.642 for Hungary [40], and -0.446 for the Netherlands [41]. These results suggest that when compared with the European populations, Asian populations are less willing to trade-off life years for a child. Future studies should examine the differences in values between instruments and the implications for applications that entail switching between instruments.

Our study also found that the social preferences with respect to EQ-5D-Y-3L and EQ-5D-3L/EQ-5D-5L health states were different [42] in China. Mobility was considered the most important dimension, followed by pain/discomfort, while usual activities was the least important dimension in both the EQ-5D-3L and EQ-5D-5L Chinese value sets [6, 7]. In contrast, pain/discomfort was the most important dimension, and mobility and self-care were the least important dimensions in this study. Similar differences between EQ-5D-3L/EQ-5D-5L and EQ-5D-Y-3L were observed in other countries [20–23], suggesting that EQ-5D-3L value sets should not be applied to EQ-5D-Y-3L data. The differences could be attributed to the differing disutility perspectives associated with different EQ-5D health dimensions for children and adults. For example, mobility and self-care are important to adults since they are supposed to be independent; these dimensions are not of foremost importance for children because they are most likely to be dependent on their parents or caregivers. Nevertheless, the differences may also have been partially due to the change in valuation perspectives (self for EQ-5D vs others for EQ-5D-Y-3L), and in descriptive systems (EQ-5D-3L vs EQ-5D-Y-3L) [13, 14].

This study is not without limitations. First, we collected cTTO data from only 8 regions. This may explain the discrepancy between the mixed logit model and the hybrid model, given that the DCE data covered more regions and there may be preference heterogeneity across regions. Second, following the EQ-5D-Y-3L protocol, the relationship between the imagined child and the respondent was not specified, which could have increased data variability [43]. Anecdotally, interviewers reported that various relationships were hypothesised by respondents; for example, the child being their own child, their friend's or relative's child, a random child, or themselves when they were 10 years old. In future

EQ-5D-Y-3L valuation studies, it may be more appropriate to standardise the relationship (e.g. a 10-year-old child you do not know) for all valuation tasks.

The published EQ-5D-Y-3L valuation protocol represented best practice at the time of development and our knowledge with respect to the valuation of children's health has been deepening and expanding over time. For example, it has been reported that adolescents could complete the DCE tasks and exhibit different health preferences compared with adults [44]. Additionally, following the current EQ-5D-Y-3L protocol, we fixed the age of the imagined child at 10 years old in the valuation task. The effect of the hypothetical child's age appears to be minimal on the latent utility values derived from the DCE tasks [45]. Hence, the resultant value set may be applied to the health states for the full age range of children and adolescents, although future studies are needed concerning the effect of age on cTTO values. As our knowledge about children's health valuation grows, an updated EQ-5D-Y-3L valuation protocol and a new EQ-5D-Y-3L value set can be expected to be developed.

5 Conclusion

Using an expanded cTTO design, this study collected both cTTO and DCE preference data for EQ-5D-Y-3L health states in a large representative sample in China. By comparing two analytical approaches, we found that the hybrid model outperformed the DCE mixed logit model (anchoring through mapping) in terms of coefficient significance and predictive performance. Furthermore, by adding an A3 term, the hybrid model addressed the value gap (between 33333 and other states) issue that has been reported in some other published EQ-5D-Y-3L studies. We recommend the hybrid model with A3 term as the EQ-5D-Y-3L value set for China.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40273-022-01216-9>.

Acknowledgments The authors thank the EuroQol Research Foundation for funding this study (Grant number: 20191020). The authors thank Elly Stolk, Bram Roudijk from the EuroQol Research Foundation, Juan M. Ramos-Goni from the Maths in Health for suggestions on data analysis. The authors also thank the interviewers: Guangjie Zhang, Xinru Jiang, Sha Li, Yue Sun, Ruijie Liu, Siyu He, Beibei Yan Jiaxuan Chen, Yuling Huang, Renfeng Su, Honglin Li, Xuelin Yao, Lingyi Zhang, Zehao Zheng, Dingyao Wang, Weihua Guo, Zhiyu Chen, Hongjie Chen, Qianyun Chai, Yuting Feng, Changhao Liang, Bo Liu, Wanji Zheng, Fan Du, Yuanyuan Tian for their dedicated work.

Declarations

Conflict of interest Dr. Yang reports grants from The EuroQol Research Foundation, during the conduct of the study; grants from The EuroQol Research Foundation, outside the submitted work; and Dr.

Yang is an EuroQol member. Dr. Jin grants from The EuroQol Research Foundation, outside the submitted work; and Dr. Jin is a member of the EuroQol Group. Dr. Luo reports grants from The EuroQol Research Foundation, during the conduct of the study; grants from The EuroQol Research Foundation, outside the submitted work; and Dr. Luo is a member of the EuroQol Group.

Author contributions Concept and design: ZY, NL. Acquisition of data: all co-authors. Analysis and interpretation of data: ZY, NL. Drafting of the manuscript: all co-authors. Critical revision of the paper for important intellectual content: all co-authors. Statistical analysis: ZY, NL. Obtaining funding: ZY, NL.

Data availability statement The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Disclosure statement This article is published in a special edition journal supplement wholly funded by the EuroQol Foundation.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. Yang Z, Busschbach J, Liu G, et al. EQ-5D-5L norms for the urban Chinese population in China. *Health Qual Life Outcomes*. 2018;16:210.
2. Yao Q, Liu C, Zhang Y, et al. Changes in health-related quality of life of Chinese populations measured by the EQ-5D-3L: a comparison of the 2008 and 2013 National Health Services Surveys. *Health Qual Life Outcomes*. 2019;17:43.
3. Weng G, Hong Y, Luo N, et al. Comparing EQ-5D-3L and EQ-5D-5L in measuring the HRQoL burden of 4 health conditions in China. *Eur J Health Econ*. 2022. <https://doi.org/10.1007/s10198-022-01465-7>.
4. Jiang J, Hong Y, Zhang T, et al. Comparing the measurement properties of the EQ-5D-5L and the EQ-5D-3L in hypertensive patients living in rural China. *Qual Life Res*. 2021;30:2045–60.
5. Wang A, Rand K, Yang Z, et al. The remarkably frequent use of EQ-5D in non-economic research. *Eur J Health Econ*. 2022;23:1007–14.
6. Liu GG, Wu H, Li M, et al. Chinese time trade-off values for EQ-5D health states. *Value Health*. 2014;17:597–604.
7. Luo N, Liu G, Li M, et al. Estimating an EQ-5D-5L Value Set for China. *Value Health*. 2017;20:662–9.
8. 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;19:875–86.
9. Zhou W, Shen A, Yang Z, et al. Patient-caregiver agreement and test-retest reliability of the EQ-5D-Y-3L and EQ-5D-Y-5L in paediatric patients with haematological malignancies. *Eur J Health Econ*. 2021;22:1103–13.
10. Wong CKH, Cheung PWH, Luo N, et al. A head-to-head comparison of five-level (EQ-5D-5L-Y) and three-level EQ-5D-Y questionnaires in paediatric patients. *Eur J Health Econ*. 2019;20:647–56.
11. Wong CKH, Cheung PWH, Luo N, et al. Responsiveness of EQ-5D Youth version 5-level (EQ-5D-5L-Y) and 3-level (EQ-5D-3L-Y) in patients with idiopathic scoliosis. *Spine (Phila Pa 1976)*. 2019;44:1507–14.
12. Wang P, Sun Y, Yang Z-H, et al. Testing measurement properties of two EQ-5D youth versions and KIDSCREEN-10 in China. *Eur J Health Econ*. 2021. <https://doi.org/10.1007/s10198-021-01307-y>.
13. 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;21:1291–8.
14. Lipman SA, Reckers-Droog VT, Karimi M, et al. Self vs. other, child vs. adult. An experimental comparison of valuation perspectives for valuation of EQ-5D-Y-3L health states. *Eur J Health Econ*. 2021;22:1507–18.
15. Mott DJ, Shah KK, Ramos-Goni JM, et al. Valuing EQ-5D-Y-3L health states using a discrete choice experiment: do adult and adolescent preferences differ? *Med Decis Mak*. 2021;41:584–96.
16. Shah KK, Ramos-Goni JM, Kreimeier S, et al. An exploration of methods for obtaining 0 = dead anchors for latent scale EQ-5D-Y values. *Eur J Health Econ*. 2020;21:1091–103.
17. Ramos-Goni JM, Oppe M, Stolk E, et al. International Valuation Protocol for the EQ-5D-Y-3L. *Pharmacoeconomics*. 2020;38:653–63.
18. Devlin N, Pan T, Kreimeier S, et al. Valuing EQ-5D-Y: the current state of play. *Health Qual Life Outcomes*. 2022;20:105.
19. Reckers-Droog V, Karimi M, Lipman S, et al. Why do adults value EQ-5D-Y-3L health states differently for themselves than for children and adolescents: a think-aloud study. *Value Health*. 2022;25:1174–84.
20. Astrom M, Conte H, Berg J, et al. “Like holding the axe on who should live or not”: adolescents’ and adults’ perceptions of valuing children’s health states using a standardised valuation protocol for the EQ-5D-Y-3L. *Qual Life Res*. 2022;31:2133–42.
21. Kreimeier S, Mott D, Ludwig K, et al. EQ-5D-Y value set for Germany. *Pharmacoeconomics*. 2022. <https://doi.org/10.1007/s40273-022-01143-9>.
22. Shiroiwa T, Ikeda S, Noto S, et al. Valuation survey of EQ-5D-Y based on the International Common Protocol: development of a value set in Japan. *Med Decis Mak*. 2021;41:597–606.
23. Prevolnik Rupel V, Ogorevc M, Group IHH. EQ-5D-Y value set for Slovenia. *Pharmacoeconomics*. 2021;39:463–71.
24. Ramos-Goni JM, Oppe M, Estevez-Carrillo A, et al. Accounting for unobservable preference heterogeneity and evaluating alternative anchoring approaches to estimate country-specific EQ-5D-Y value sets: a case study using Spanish preference data. *Value Health*. 2022;25:835–43.
25. Roudijk B, Sajjad A, Essers B, et al. A value set for the EQ-5D-Y-3L in the Netherlands. *Pharmacoeconomics*. 2022. <https://doi.org/10.1007/s40273-022-01192-0>.
26. Rencz F, Ruzsa G, Bató A, et al. Value set for the EQ-5D-Y-3L in Hungary. *Pharmacoeconomics*. 2022. <https://doi.org/10.1007/s40273-022-01190-2>.
27. 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*. 2017;55:e51–8.
28. Purba FD, Hunfeld JAM, Iskandarsyah A, et al. The Indonesian EQ-5D-5L value set. *Pharmacoeconomics*. 2017;35:1153–65.

29. Yang Z, Luo N, Bonsel G, et al. Selecting health states for EQ-5D-3L valuation studies: statistical considerations matter. *Value Health*. 2018;21:456–61.
30. Jonker MF, Donkers B, de Bekker-Grob E, et al. Attribute level overlap (and color coding) can reduce task complexity, improve choice consistency, and decrease the dropout rate in discrete choice experiments. *Health Econ*. 2019;28:350–63.
31. Tabulation on the 2010 population census of the People's Republic of China. Population Census Office under the State Council 2010.
32. Ramos-Goni JM, Oppe M, Slaap B, et al. Quality control process for EQ-5D-5L valuation studies. *Value Health*. 2017;20:466–73.
33. StataCorp. *Stata Statistical Software: Release 14*. College Station: StataCorp LLC; 2015.
34. Lipman SA, Zhang L, Shah KK, et al. Time and lexicographic preferences in the valuation of EQ-5D-Y with time trade-off methodology. *Eur J Health Econ*. 2022. <https://doi.org/10.1007/s10198-022-01466-6:1-13>.
35. 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.
36. 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.
37. Johnson FR, Yang JC, Reed SD. The internal validity of discrete choice experiment data: a testing tool for quantitative assessments. *Value Health*. 2019;22:157–60.
38. 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;19:648–54.
39. Ludwig K, Graf von der Schulenburg JM, Greiner W. German value set for the EQ-5D-5L. *Pharmacoeconomics*. 2018;36:663–74.
40. Rencz F, Brodszky V, Gulácsi L, et al. Parallel valuation of the EQ-5D-3L and EQ-5D-5L by time trade-off in Hungary. *Value Health*. 2020;23:1235–45.
41. Versteegh M, Vermeulen K, Evers S, et al. Dutch tariff for the Five-Level Version of EQ-5D. *Value Health*. 2016;19:343–52.
42. 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.
43. Lipman SA, Essers BAB, Finch AP, et al. In a child's shoes: composite time trade-off valuations for EQ-5D-Y-3L with different proxy perspectives. *Pharmacoeconomics*. 2022. <https://doi.org/10.1007/s40273-022-01202-1>.
44. Prevolnik Rupel V, Ramos-Goni JM, Ogorevc M, et al. Comparison of adult and adolescent preferences toward EQ-5D-Y-3L health states. *Value Health*. 2021;24:1350–9.
45. Ramos-Goni JM, Estevez-Carrillo A, Rivero-Arias O, et al. Does changing the age of a child to be considered in 3-level version of EQ-5D-Y discrete choice experiment-based valuation studies affect health preferences? *Value Health*. 2022;25:1196–204.

Authors and Affiliations

Zhihao Yang^{1,2}  · Jie Jiang² · Pei Wang³ · Xuejing Jin⁴ · Jing Wu⁵ · Yu Fang⁶ · Da Feng⁷ · Xiaoyu Xi⁸ · Shunping Li⁹ · Mingxia Jing¹⁰ · Bin Zheng¹¹ · Weidong Huang¹² · Nan Luo¹³

¹ Health Services Management Department, Guizhou Medical University, Gui'an, China

² College of Pharmacy, Jinan University, Guangzhou, China

³ School of Public Health, Fudan University, Shanghai, China

⁴ Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

⁵ School of Pharmaceutical Science and Technology, Tianjin University, Tianjin, China

⁶ Department of Pharmacy Administration, School of Pharmacy, Health Science Centre, Xi'an Jiaotong University, Xi'an, China

⁷ School of Pharmacy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

⁸ The Research Centre of National Drug Policy and Ecosystem, China Pharmaceutical University, Nanjing, China

⁹ Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

¹⁰ Department of Preventive Medicine, Medical College of Shihezi University, Shihezi, China

¹¹ Department of Pharmacy, Fujian Medical University Union Hospital, Fuzhou, China

¹² School of Health Management, Harbin Medical University, Harbin, China

¹³ Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore