Development of Chinese VAS Value Set for EQ-5D-3L and Difference analysis with the TTO Value Set Based on Same Nationally Representative Sample

Lang Zhuo (✉ zhuolang@xzhmu.edu.cn)

Xuzhou Medical University <https://orcid.org/0000-0003-3926-9894>

Lin Zhuo

Tianjin Medical University

Siting Chen

Xuzhou Medical University

Ting Cai

Xuzhou Medical University

Jianqiang Pan

Xuzhou Medical University

Yunxin Kong

Xuzhou Medical University

Haiting Jiang

Nanjing Medical University

Jiaying Chen

Nanjing Medical University

Research Article

Keywords: China, difference analysis, EQ-5D-3L, health-related quality of life, visual analog scale

Posted Date: April 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-367677/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective: To develop a Chinese VAS value set for EQ-5D-3L and compare it with the TTO value set based on nationally representative sample.

Methods: An adapted Measurement and Valuation of Health (MVH) protocol was applied with VAS method. EQ-5D-3L was used in face-to-face interviews conducted by trained interviewers with participants selected via multi-stage stratified clustered random sample. Fifteen hypothetical health statuses (11 random states in MVH protocol, plus 11111, 33333, unconscious and death) were assigned for assessment individually. Ordinary least square, general least square and weighted least square models were constructed with or without N3. Four categories of indices, including quality of original data, distribution of rescaled values, goodness of fit of models, and distribution of predicted values were adopted to compare between Chinese VAS and TTO value sets.

Results: All 5,939 participants aged 15 and over were interviewed; 5,884 eligible participants were included in constructing models. An OLS model that included ten dummy variables and a constant and without N3 was chosen as the best-fit VAS model (Adjusted $R^2=0.670$). The mean absolute error was 0.0319, and the correlation coefficient between predicted and mean values was 0.9837. The curve of predicted value of VAS model was uniformly lower than that of Chinese TTO value set. Comparing with the TTO value set, VAS method had higher responsive rate, less inconsistency, lower skewed values and better goodness of fit values.

Conclusions: VAS value set performed rationally in Chinese population for its simplicity, responsiveness and robustness. We recommend the VAS value set, Model 2, as the benchmark of scoring algorithm of HRQoL for Chinese population.

Introduction

With the extension of human lifespan, the measurement of health-related quality of life (HRQoL) plays an increasingly important role worldwide [1–3]. HRQoL has been measured using specialized instruments, including the SF-36[4], WHOQOL100[5] and EQ-5D[6], among others. As the most concise preference-based instrument, EQ-5D has been applied generally since its publication in 1990 [7].

The EuroQol five-dimensional questionnaire (EQ-5D) consists of a descriptive system and a visual analogue scale (VAS). The EQ-5D-3L describes HRQoL in terms of five dimensions (mobility, self-care, usual activities, pain/discomfort, and depression/anxiety); each assessed at three severity levels (no problem, moderate problem, and severe problem). The VAS, a 20cm vertical scale, solicits a self-rating value of HRQoL. Another commonly used elicitation technique is time trade-off (TTO), which forces trade-offs between 10-year time of life and quality of life under conditions of certainty[8]. After eliciting the preferences from the general population based on a selected subset of the EQ-5D states, a scoring algorithm could then be developed for predicting the utility weights for all 243 states[9].

To date, there are 34 value sets for EQ-5D-3L available on EuroQol website[10]. More than half countries constructed value sets using the VAS method, others using the TTO method, and few of them using both methods[10]. We published the Chinese TTO value set for EQ-5D-3L in 2018[11].

Although TTO is adopted as a mainstream approach to elicit health utility weights for EQ-5D[8], many debates are still ongoing about it, e.g. TTO is burdensome[12], has duration effect[13], interviewer effect[14], and more inconsistencies[15]. In 2020, We proved that VAS and TTO could be equivalent under certain conditions, but in general population, VAS method appears a more logical approach for valuation due to its greater reliability, simplicity and feasibility[16]. Hence, the Chinese VAS value set is critical for not only deriving a safe value of HRQoL for Chinese population but also exploring the potential bias between the two methods.

Though studies have shown that EQ-5D-5L was associated with reduced ceiling effects and improved discriminatory power[17], EQ-5D-3L has responsiveness advantage[18] and has been widely applied, e.g. EQ-5D-3L has been adopted in the National Health Services Survey (NHSS) of China three times in a row from 2008 to 2018. Therefore, developing a Chinese VAS value set for the EQ-5D-3L has irreplaceable key function to deepen the use of the instrument. The main objectives of this study are 1) to develop a hypothesis-based VAS value set for EQ-5D-3L; 2) to compare the Chinese VAS value set with the TTO value set.

Methods

Health States Description and Selection

An EQ-5D-3L health state, 12333 for example, represents no problems with mobility (level 1), moderate problems with self-care(level 2), and severe problems with usual activities, pain/discomfort and anxiety/depression (level 3)[6; 19]. According to the Measurement and Valuation of Health (MVH) protocol, 43 health states were selected for valuation. Each respondent was assigned to value 11 randomly selected health states (two very mild states, three mild states, three moderate states, and three severe states) and 4 particular states (1111, 33333, "unconscious" and "dead"[6].

Sample Size

For this study, Chinese aged 15 years and older in 2014 represent the target population. Based on Dolan's estimates, a sample size of 3,235 should be adequate for modeling [6]. Considering the regional variation and potential deviation of the Chinese population, a sample size of 6,000 respondents was selected.

Sampling Method

A multi-stage, stratified, clustered random sample was drawn from the target population from five areas involving the provinces of Jiangsu, Shaanxi, Guangdong, Hebei, and Chongqing. One county (rural area) and one district (urban area) from each province were selected based on economic development status. In each selected area, five rural and five urban streets were sampled. In each selected street, 60

households were systematically drawn according to the existing household code, and all family members 15 years and older were interviewed [20].

Interviewer Training

A total of 10 supervisors trained at Nanjing Medical University (NMU) were sent to the sample counties/districts for interviewer training and quality control. A total of 108 interviewers who had taken part in NHSS 2013 were chosen from the local Health Service Center. To decrease the interviewer effects [21], the procedure was described in a guidebook. A three-day training workshop was conducted, including lectures, demonstrations, mock interviews, and pilot interviews.

The Valuation Tasks

Face-to-face interviews were conducted in the participant's homes based on the MVH protocol [6]. The questionnaire included 17 demographic questions, the EQ-5D-3L descriptive system, the VAS rating scale and the 15 pre-selected health states. The interviewer asked the respondent to describe and rate their health. Then respondents ranked the 15 pre-selected health states and rated them on the VAS regarding their severity level. After the task, the first three states were re-valued. To help the participants valuing the state, it was marked on a board of visual aid. (Figure S1 in Supplemental Materials).

Data Entry and Cleaning

The supervisors checked the questionnaires by the end of each day. Any missing data and/or serious inconsistencies including 1) All states valued the same except for non-traders; 2) All states valued worse than dead were returned to the interviewers for updating. Five percent of the respondents were randomly re-interviewed by the supervisor to confirm the accuracy of data collection. Data were input into the database by supervisors for the first time and latter by NMU's staff with the aid of real-time checking function in Epidata3.1.

Exclusion criteria included: 1) participants who could not complete the questionnaire; 2) less than 15 years old; 3) more than four inconsistencies i, 4) involving extreme outliers. Extreme outliers are values that met all the following criteria: 1) points away from continuous groups in a boxplot; 2) the distance between the outlier and the nearer quartile is more than 3 times of interquartile range; 3) lower than the 5th percentile or higher than the 95th percentile [22]. Inconsistency means a worse state valued 0.1 or higher than any logically better state[23].

Modeling

The definition of dummy variables and specifications of the models constructed in this study are presented in Table 1. The main effect was represented by a set of ten dummy variables. MO2, SC2, UA2, PD2 and AD2 representing the movements from no problems (level 1) to moderate problems (level 2), and MO3, SC3, UA3, PD3 and AD3 representing the movements from no problems (level 1) to severe problems (level 3). If level 2 or 3 was found in any of the dimensions, it was represented by the constant and variable N3, respectively [9].

Table 1
Definition of the dummy variables and description of the models

Dummy variables	Definition	
MO2	1 if mobility is level 2; 0 otherwise	
MO3	1 if mobility is level 3; 0 otherwise	
SC2	1 if self-care is level 2; 0 otherwise	
SC3	1 if self-care is level 3; 0 otherwise	
UA2	1 if usual activities is level 2; 0 otherwise	
UA3	1 if usual activities is level 3; 0 otherwise	
PD2	1 if pain/discomfort is level 2; 0 otherwise	
PD3	1 if pain/discomfort is level 3; 0 otherwise	
AD2	1 if anxiety/depression is level 2; 0 otherwise	
AD3	1 if anxiety/depression is level 3; 0 otherwise	
N3	1 if any dimension is level 3; 0 otherwise	
Model Specification	f (x)	Methods
Model 1	f (MO2 MO3 SC2 SC3 UA2 UA3 PD2 PD3 AD2 AD3 N3)	OLS with constant and with N3
Model 2	f (MO2 MO3 SC2 SC3 UA2 UA3 PD2 PD3 AD2 AD3)	OLS with constant and without N3
Model 3	f (MO2 MO3 SC2 SC3 UA2 UA3 PD2 PD3 AD2 AD3 N3)	GLS multilevel effect with constant and with N3
Model 4	f (MO2 MO3 SC2 SC3 UA2 UA3 PD2 PD3 AD2 AD3)	GLS multilevel effect with constant and without N3
Model 5	f (MO2 MO3 SC2 SC3 UA2 UA3 PD2 PD3 AD2 AD3 N3)	WLS with constant and with N3
Model 6	f (MO2 MO3 SC2 SC3 UA2 UA3 PD2 PD3 AD2 AD3)	WLS with constant and without N3
Model A	Model 2 + sex	OLS with constant and without N3, robust
Model B	Model 2 + sex + age group	Ditto
Model C	Model 2 + sex + age group + district	Ditto

Dummy variables	Definition	
Model D	Model 2 + sex + age group + district + marital status	Ditto
Model E	Model 2 + sex + age group + district + education	Ditto
Model F	Model 2 + sex + age group + district + employment	Ditto
Model G	Model 2 + sex + age group + district + economic status	Ditto
Model H	Model 2 + sex + age group + district + marital status + education + employment + economic status	Ditto

The VAS values (dependent variable) were calculated as $(X_{\text{raw}} - \text{Deadraw}) / (11111_{\text{raw}} - \text{Deadraw})$. X_{raw} means the original value on the scale of state X. The rescaled values are numerical, 1 refers to full health, and 0 to death[6].

In the regression models, disutility (1-rescaled VAS value) was set as the dependent variable (Y) [24–26]. Six models were constructed with three types of regression methods. Ordinary least squares (OLS) models were constructed for the basic characteristics of the model. General least squares (GLS) regression models with a random effect were constructed to rectify the significance potentially influenced by inner correlations as each respondent valued 13 health states. Finally, weighted least squares (WLS) models were also explored to rectify the possible heteroscedasticity. The inverse of the variance of residuals was employed as the weight. Multiple linear regression models were built using STATA/SE 12.0 (StataCorp, College Station, TX) with an α of 0.05[27].

Criteria for Choice of Model

We chose the best-fit model based on the following criteria [9]: 1) logical consistency in each dimension; 2) the sign of coefficients should be positive; 3) mean absolute error (MAE), adjusted R^2 , Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used for valuing goodness of fit [9]; 4) simplicity plays a more critical role when goodness of fit makes less sense; 5) easy for non-experts to understand [11].

Sensitivity and Subgroup Analyses and Reliability Tests

The database was randomly split into two groups; one group was used to construct model, and the other group for estimation of goodness of fit. To identify the effect on the model coefficients, demographic and socioeconomic variables such as sex, age group, marital status, district, education level, employment, and economic status were added to the final model.

Based on the re-valuated states, the following reliability tests were performed: the rates of consistency, Pearson correlation coefficient, Kappa, intra-class correlation coefficient, and Cronbach α .

Results

Data Exclusion and reliability tests

The survey was conducted from 10th July to 25th August 2014. A total of 6,041 individuals were interviewed with a mean time of 48.6 ± 16.7 minutes. Due to mental problems, 98 participants were unable to complete the interview, 4 were under 15 years, and no participants met the exclusion criteria. At this stage, 5,939 participants were kept for checking of outliers and inconsistencies.

A total of 104 extreme outliers were identified, involving 55 participants who were entirely excluded. A total of 571 participants were identified as having inconsistent values; 496 of whom had only one inconsistency, and no participants had more than 4 inconsistencies. Inconsistencies were all kept in the final dataset, as suggested by Lamers et al. [23]. The final sample consisted of 5,884 individuals, representing 99.1% of the whole sample.

Based on the three re-valued tasks, the reliability tests indicated a high level of reliability[28; 29], e.g., the rates of consistency were 78.2%, 79.0% and 80.6%; Pearson correlation coefficients were 0.90, 0.92 and 0.93; Kappa values were 0.72, 0.75 and 0.77; Intra-class correlation coefficients were 0.95, 0.96 and 0.97; Cronbach's α were 0.95, 0.96 and 0.97, respectively.

Sample Characteristics

Based on the probabilistic sampling method, the proportion of the sample demographic characteristics was similar to the general population in 2014[30]. The demographic and socioeconomic characteristics are shown in Table 2. The composing of the final sample compared with the Chinese population is presented in Appendix Table S1 in Supplementary Materials.

Table 2
Demographic, socioeconomic and health-related characteristics of the sample (n = 5,884)

Variables	Category	%	n
Sex	Male	48.1	2,830
	Female	51.9	3,054
Age (y)	15–24	11.0	644
	25–34	14.9	876
	35–44	17.2	1,011
	45–54	20.4	1,202
	55–64	20.1	1,180
	65–74	11.2	656
	75–97	5.3	315
Region	Urban	50.3	2,958
	Rural	49.7	2,926
District	North (Hebei)	21.1	1,241
	East (Jiangsu)	20.5	1,205
	South (Guangdong)	17.4	1,021
	Middle (Chongqing)	19.2	1,130
	West (Shaanxi)	21.9	1,287
Marital status	Single	13.9	815
	Married	79.9	4,703
	Divorced	1.6	93
	Widowed	4.6	270
	Other	0.1	3
Educational level	Below primary school	6.8	400
	Primary school	18.6	1,092
	Junior middle school	31.2	1,835
	Senior middle school	17.5	1,030
*Subjective perception compares with neighborhood			
* Reported by participants who had been diagnosed before by doctor in county hospital or higher			

Variables	Category	%	n
	Technology school	6.7	392
	Junior college	10.8	636
	University and above	8.5	499
Economic status*	Good	16.5	970
	Common	73.6	4,332
	Poor	9.9	582
Occupational status	Employed	60.4	3,551
	Retired	18.8	1,104
	Students	5.6	328
	Unemployed	1.7	98
	No occupation	13.7	803
Ethnicity	Han	88.5	5,207
	Other	11.5	677
Chronic disease**	Yes	25.1	1,474
Disability**	Yes	3.8	225
Self-rated health	Very good	24.9	1,464
	Good	46.8	2,751
	Fair	22.8	1,341
	Poor	5.0	291
	Very poor	0.6	37
Mobility dimension	No problem	94.5	5,562
	Moderate problem	5.2	305
	Severe problem	0.3	17
Self-care dimension	No problem	98.2	5,777
	Moderate problem	1.5	88
	Severe problem	0.3	19

*Subjective perception compares with neighborhood

* Reported by participants who had been diagnosed before by doctor in county hospital or higher

Variables	Category	%	n
Usual activities dimension	No problem	95.6	5,623
	Moderate problem	3.7	215
	Severe problem	0.8	46
Pain/discomfort dimension	No problem	85.6	5,034
	Moderate problem	14.12	831
	Severe problem	0.32	19
Anxiety/depression dimension	No problem	92.85	5,463
	Moderate problem	6.83	402
	Severe problem	0.32	19
11111	Yes	81.1	4,773
At least one at level 3	Yes	1.2	71
*Subjective perception compares with neighborhood			
* Reported by participants who had been diagnosed before by doctor in county hospital or higher			

Table 2 also shows health-related characteristics, including problems reported on each EQ-5D-3L.

Chronic disease (25.1%) and disability (3.8%) rates which had been diagnosed in a county hospital or higher were reported. In EQ-5D, the most reported problems were pain/discomfort (14.4%), followed by anxiety/depression (7.2%).

With the whole sample, the mean VAS value for the 43 evaluated states was 0.577 ± 0.323 . The minimum was -1.375 ; and the 1st, 5th and 10th percentiles at -0.333 , -0.0526 , 0.1333 respectively, indicating that no transformation is needed to rescaling negative values.

Analysis of Models

In Table 3, 6 regression models, coefficient estimates, and statistics are presented. All models and coefficients were significant ($P < 0.001$) but failed to pass the heteroscedasticity and inner correlation test. Differences for each dummy variable coefficients were all less than 0.03 in absolute terms, among which 90% were less than 0.01. Six predicted-value curves for Model 1 through Model 6 are shown in Appendix Figure S2. The six lines almost overlapped with each other. We chose the best-fit model based on the following process.

Table 3
Coefficient estimates and statistics of OLS, GLS and WLS models on VAS values

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	OLS	OLS	GLS	GLS	WLS	WLS
11111	1	1	1	1	1	1
Constant	-0.0242	-0.0274	-0.0245	-0.0272	-0.0142	-0.0185
M02	-0.0721	-0.0727	-0.0754	-0.0760	-0.0832	-0.0805
M03	-0.1769	-0.1788	-0.1787	-0.1803	-0.1770	-0.1881
SC2	-0.1117	-0.1105	-0.1120	-0.1110	-0.1126	-0.1097
SC3	-0.2332	-0.2349	-0.2343	-0.2358	-0.2207	-0.2276
UA2	-0.0802	-0.0880	-0.0803	-0.0874	-0.0852	-0.0900
UA3	-0.1397	-0.1518	-0.1404	-0.1514	-0.1290	-0.1550
PD2	-0.0920	-0.0898	-0.0884	-0.0865	-0.1009	-0.0980
PD3	-0.1601	-0.1667	-0.1596	-0.1656	-0.1565	-0.1847
AD2	-0.0839	-0.0851	-0.0827	-0.0838	-0.0833	-0.0888
AD3	-0.1623	-0.1702	-0.1645	-0.1716	-0.1423	-0.1691
N3	-0.0236	-	-0.0213	-	-0.0539	-
Adj. R-square	0.6704	0.6700	-	-	0.7390	0.7227
AIC	40706.11	40613.89	81954.97	81800.25	111284.00	101688.00
BIC	40595.17	40512.20	81825.54	81680.07	111173.10	101586.30
MAE	0.0317	0.0319	0.0317	0.0319	0.0304	0.0325
#>0.05	7	7	8	10	9	11
#>0.1	0	0	0	0	1	2
r	0.9847	0.9837	0.9846	0.9836	0.9846	0.9828
# Logical error	0	0	0	0	0	0
All models were significant and $P=0.0000$; $P<0.001$ for all regression coefficients unless otherwise stated						
* Significance for all coefficients $P<0.001$						
M2,mobility dimension level 2, etc.						
Adj. R-square, adjusted R-square						

Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
OLS	OLS	GLS	GLS	WLS	WLS
AIC, Akaike information criterion					
BIC, Bayesian information criterion					
MAE, mean absolute error between actual mean and predicted value					
# >0.05, number of MAE > 0.05 out of 42 states					
#>0.05, number of MAE > 0.1 out of 42 states					
r, correlation coefficient of mean and predict value					
#Logical error, number of inconsistency occurred in 7,533 compairable states					

Between the two OLS models, Model 1 (OLS with constant and with N3) was abandoned as N3 term increased the complexity without improving the fitness. Model 2 (OLS with constant and without N3) was preferred at this stage. Compare to Model 2, all coefficients in Model 3 (GLS with constant and with N3) and Model 4 (GLS with constant and without N3) were significant, same as Model 2. This meant potential inner correlation did not lead to an invalid significance of the coefficients in OLS methods. However, AIC, BIC, and the number of MAE > 0.05 notably increased. So, Model 2 still was preferred in the second stage. In the last stage, Model 5 (WLS with constant and with N3) and 6 (WLS with constant and without N3) were constructed based on the WLS method to address heteroscedasticity issues [31]. Although WLS models had little higher adjusted R-square, AIC, BIC and the number of MAE > 0.05 increased more. Additionally, different weights might lead to different models, producing new potential bias. So WLS models were also dropped at the final stage. Due to its consistency, parsimony, and transparency as well as its high level of goodness of fit, model 2 appeared preferable among all the models. Thus, we chose Model 2 as the best-fit model. The predicted value, "22321" as an example, could be calculated using the following algorithm:

$$1 - 0.0274 - 0.0727 * 1 - 0.1105 * 1 - 0.1518 * 1 - 0.0898 * 1 = 0.5478$$

All 243 predicted values with Model 2 are shown in Appendix Table S2 in Supplemental Materials. Appendix Figure S3 shows the observed mean values compared with the predicted values based on Model 2 for 43 valued states. The mean errors between observed and predicted values were minimal especially on the left of the curve. Appendix Table S4 shows goodness-of-fit indices for the whole sample and the split sample models indicating a high level of robustness.

Influencing Factors

Based on Model 2, the socio-demographic influence on health state valuation is presented in Appendix Table S3. The coefficient for females was significant but very small (-0.006). Only two age groups of

65–74 and 75–97 had significant coefficients of 0.011 and 0.014, respectively. For educational level, a negative gradient of coefficients was identified significantly. Among demographic variables, the college group and above exhibits the least coefficient of -0.050. Comparing Model 2 with those from Model A to Model H, all coefficients of ten dummy variables performed steadily. All these characteristics meant that VAS method is generally robust.

Comparisons with Other Models

Figure 1A shows the predicted values of Model 2 compared with the value sets of the United Kingdom, New Zealand, Belgium, Spain, Finland, Slovenia and Denmark[22]. Except for full health, Model 2 gave higher values than all the models from these countries.

Figure 1B shows the predicted values of Model 2 compared with the Chinese TTO value set derived from the same sample[11]. Except for full health, Model 2 gave uniformly lower values than Chinese TTO value set.

Table 4 shows four categories of indices for comparing Model 2 and Chinese TTO value set, including quality of original data, distribution of rescaled values, goodness of fit of models, distribution of predicted values of 5,939 cases of the sample[11]. The indices showed that the VAS method has less inconsistency of original data, lower skewness of rescaled values, higher goodness of fit, and more responsive predictions of the sample. These characteristics meant VAS was a more feasible approach than TTO in the Chinese population due to its simplicity, robustness and responsiveness. All referred value sets compared with Model 2 are presented in Appendix Table S5[11; 22].

Table 4
Four categories of indices for comparing the Model 2 with Chinese TTO value set

Categories	Indices	Model 2 (VAS)	Value set of Chinese TTO
Quality of original data	Excluded cases	55(0.93%)	436(7.34%)
	Inconsistencies	571(9.70%)	648(11.78%)
Distribution of rescaled values of 43 health states	Mean	0.577	0.713
	Standard deviation	0.323	0.446
	P ₂₅	0.382	0.700
	P ₅₀	0.625	0.900
	P ₇₅	0.833	0.983
	Skewness	-0.901	-2.838
	Kurtosis	3.925	13.321
Goodness of fit of models	Adj. R-square	0.6700	0.3543*
	MAE	0.0319	0.0838
	MAE > 0.05	9	24
	MAE > 0.1	0	12
	correlation coefficient	0.9837	0.8989
	AIC	40613.89	56444.80
	BIC	40512.20	56536.58
	VPC	0.68%	1.21%
Distribution of predicted values of 5,939 cases of the sample	Mean	0.964	0.985
* Substitutive value of the OLS model with constant			
Adj. R-square, adjusted R-square			
MAE, mean absolute error between observed mean and predicted value			
AIC, Akaike information criterion			
BIC, Bayesian information criterion			
VPC, variance partition coefficient with a multi-level Zero-model for interviewer effect			

Categories	Indices	Model 2 (VAS)	Value set of Chinese TTO
	Standard deviation	0.092	0.052
	P ₅₀	1	1
	P ₂₅	1	1
	P ₁₀	0.883	0.964
	P ₅	0.527	0.779
	Skewness	3.633	7.102
	Kurtosis	20.536	74.605
* Substitutive value of the OLS model with constant			
Adj. R-square, adjusted R-square			
MAE, mean absolute error between observed mean and predicted value			
AIC, Akaike information criterion			
BIC, Bayesian information criterion			
VPC, variance partition coefficient with a multi-level Zero-model for interviewer effect			

Discussion

Modeling

This study reported on the development of a hypothesis-based VAS value set for EQ-5D-3L with a nationally representative Chinese sample. The best fit model was based on OLS regression with main effects and a constant.

Although Dolan P. had long recommended GLS models[6], many authors still preferred the OLS models[32–34]. In our study, we constructed GLS and WLS models for precise estimation of the significance of coefficients. The final choice was based on full comparison according to our prior criteria.

We dropped N3 term for parsimony since it could not significantly improve the goodness of fit in Model 1: the MAE only increased by 0.0002 from 0.0317 to 0.0319 with N3 model, the number of MAE > 0.05 did not change and Pearson correlation coefficients only decreased by 0.001. Furthermore, almost half of the value set published worldwide have no N3 term [35].

The constant, an average loss of utility deviated from full health according to Dolan P.[6], was included in the final model as most value sets performed. Constant in Model 2 was only 0.0274, less than one-fifth of that of the U.K. [6]. Through the algorithm of Model 2, the measured value for full health (0.9726) was very close to the value 1 determined by the original hypothesis. The loss of HRQoL is due to the fact that the "fully healthy" reporter is not fully healthy, but does not reach a moderate level. These reporters were identified by a head-to-head study [36].

Comparison with Value Sets of other countries

Figure 1A shows that the line of Model 2 was consistently higher than that of other countries. This meant that under the same health condition, the utility expectation of Chinese people is higher than that of other countries. There are several possible explanations for this phenomenon. Firstly, different cultures lead to different understandings of quality of life [37–40]. China had experienced a period of sufferings in early twentieth century; therefore, Chinese people's tolerance of low quality of life is much higher than that of other countries. Secondly, as a developing country with many relatively undeveloped areas, Chinese people living in these areas have low expectations of quality life. Thirdly, in recent decades, China's health service system has been significantly improved, and the satisfaction with the quality of life has been greatly increased.

As to the most influential dimension, self-care had the most significant impact on HRQoL for severe problems in our study. In countries such as New Zealand and Belgium, anxiety/depression had the greatest impact, followed by pain/discomfort[22]. The results emphasize the importance of different dimensions to the overall HRQoL.

Comparison with Chinese TTO Value Set

Figure 1B shows that the line of Model 2 was consistently lower than the line of Chinese TTO value set. The same trend was also identified in Sweden [16]. These phenomena indicated that participants were reluctant to trade despite the low VAS score. Indices in Table 4 show a notable difference between VAS and TTO methods. They also showed obvious psychological burden under the TTO method[12]. For Chinese, there are several possible reasons for their unwillingness to trade. Firstly, Filial Piety. An ancient Chinese book, named *Classic of Filial Piety*, said that Filial Piety begins with taking good care of our body, skin and hair for alleviating parents' worries about our physical injury. The ultimate goal of filial piety is to do good for the people, to set an example for future generations for leaving parents a great sense of honor[41]. These deep-rooted moral values regulate everyone to stay away from any behavior that endangers his/her health[42]. Secondly, complex emotional combinations, such as emotional attachment to family members[42], affection to wealth [43], fear for death[44], etc. reduce the willingness to trade. Thirdly, some generally accepted beliefs, such as "better a living dog than a dead lion" lead to people's unwillingness to trade[45].

Compare with the Chinese TTO value set, the coefficient ranking of Model 2 was basically the same, since 7 of the 10 coefficients were in the same position[11]. Self-care (-0.2349) had the most significant impact on HRQoL for severe problems in Model 2, followed by mobility (-0.1788), anxiety/depression (-0.1702),

pain/discomfort (-0.1667) and usual activities (-0.1518). Moderate problems in all dimensions except self-care (-0.1105) were less than 0.1 in absolute terms. This phenomenon demonstrated the internal correlation between the two methods.

Socio-demographic Factors

Socio-demographic factors exerted minor influence on Model 2. Firstly, all coefficients of ten dummy variables changed little from Model A to H. Secondly, Adjusted R^2 increased from 0.670 to 0.677 from Model A to H. Thirdly, the highest coefficient of college or above is 0.05 in absolute term, which might be related to their higher expectation on HRQoL[46; 47]. Compared with Model 2, coefficients of Chinese TTO value set varied from Model A to H[11]. Furthermore, studies have shown that factors other than demographic variables may affect on the TTO values, such as attitudes toward death [48]. Therefore, the performance of VAS method in Chinese population is more stable than that of TTO method.

Strengths and Weaknesses

There were several strengths to our study. 1) A large nationally representative sample. Sensitivity analysis showed that the large sample size of 5,884 participants had ensured a robust model. 2) High-quality survey. The study was authorized by the National Health Commission (NHC) of China, supervised by NMU and implemented by experienced health service personnel from local health service center. 3) High level of reliability was proved through reliability test, sensitivity analysis, robustness of coefficients, small constant, and responsiveness of predicted values.

Interviewer effect is one of the weakness of this study. Even though we carried out extensive training, 108 interviewers still had different background and interview skills. Another weakness was the qualitative explanation of the difference between Model 2 and Chinese TTO value set. Further quantitative analysis of the difference will be informative and should be encouraged.

To our knowledge, this is the first try to develop a Chinese VAS value set for EQ-5D-3L, and provide abundant evidence to show that VAS method still has value and advantages..

Conclusions

Vas method has value and advantages in a wider range of application scenarios for its simplicity, responsiveness, and robustness. We recommend VAS value set, model 2, as the benchmark of scoring algorithm of HRQOL in Chinese population.

Abbreviations

CREATE: Checklist for reporting valuation studies

EQ-5D-3L: three-level version of EQ-5D

GLS: Generalized Least Square

HRQoL: Health-Related Quality of Life

MAE: Mean Absolute Error

MVH: Measurement and Valuation of Health

NHFPC: National Health and Family Planning Commission, formerly known as Chinese Ministry of Health

NHSS: National Health Service Survey

NMU: Nanjing Medical University

OLS: Ordinary Least Square

TTO: Time Trade-Off

VAS: Visual Analogue Scale

WLS: Weighted Least Square

Declarations

Competing interests

The authors declare that they have no Competing of interest.

Compliance with Ethical Standards

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Nanjing Medical University Ethics Committee approved this study (EA#20140706004).

Informed consent: Written informed consents were obtained from all individual participants included in the study. Participants were informed that it was voluntary to participate and that they could terminate at any time during the interview.

Funding

The project was funded National Natural Science Foundation of China (grant no. 71373183).

Acknowledgements

We wish to express thanks for the project funding from the National Natural Science Foundation of China (grant no. 71373183). We also thank the National Health Commission of the People's Republic of China and local health administrative institutes in five provinces for their support of the implementation of the survey. Additionally, we appreciate all of the efforts of the dedicated supervisors and interviewers.

References

1. Benedini, V., Caporaso, N., Corazza, G. R., Rossi, Z., Fornaciari, G., Cottone, M., Frosini, G., Caruggi, M., Ottolini, C., & Colombo, G. L. (2012). Burden of Crohn's disease: economics and quality of life aspects in Italy. *Clinicoecon Outcomes Res*, 4, 209-218.
2. Trippoli, S., Vaiani, M., Lucioni, C., & Messori, A. (2001). Quality of life and utility in patients with non-small cell lung cancer. Quality-of-life Study Group of the Master 2 Project in Pharmacoeconomics. *Pharmacoeconomics*, 19(8), 855-863.
3. Global Burden of Disease Cancer, C., Fitzmaurice, C., Allen, C., Barber, R. M., Barregard, L., Bhutta, Z. A., Brenner, H., Dicker, D. J., Chimed-Orchir, O., Dandona, R., Dandona, L., Fleming, T., Forouzanfar, M. H., Hancock, J., Hay, R. J., Hunter-Merrill, R., Huynh, C., Hosgood, H. D., Johnson, C. O., Jonas, J. B., Khubchandani, J., Kumar, G. A., Kutz, M., Lan, Q., Larson, H. J., Liang, X., Lim, S. S., Lopez, A. D., MacIntyre, M. F., Marczak, L., Marquez, N., Mokdad, A. H., Pinho, C., Pourmalek, F., Salomon, J. A., Sanabria, J. R., Sandar, L., Sartorius, B., Schwartz, S. M., Shackelford, K. A., Shibuya, K., Stanaway, J., Steiner, C., Sun, J., Takahashi, K., Vollset, S. E., Vos, T., Wagner, J. A., Wang, H., Westerman, R., Zeeb, H., Zoeckler, L., Abd-Allah, F., Ahmed, M. B., Alabed, S., Alam, N. K., Aldhahri, S. F., Alem, G., Alemayohu, M. A., Ali, R., Al-Raddadi, R., Amare, A., Amoako, Y., Artaman, A., Asayesh, H., Atnafu, N., Awasthi, A., Saleem, H. B., Barac, A., Bedi, N., Bensenor, I., Berhane, A., Bernabe, E., Betsu, B., Binagwaho, A., Boneya, D., Campos-Nonato, I., Castaneda-Orjuela, C., Catala-Lopez, F., Chiang, P., Chibueze, C., Chittheer, A., Choi, J. Y., Cowie, B., Damtew, S., das Neves, J., Dey, S., Dharmaratne, S., Dhillon, P., Ding, E., Driscoll, T., Ekwueme, D., Endries, A. Y., Farvid, M., Farzadfar, F., Fernandes, J., Fischer, F., TT, G. H., Gebru, A., Gopalani, S., Hailu, A., Horino, M., Horita, N., Hussein, A., Huybrechts, I., Inoue, M., Islami, F., Jakovljevic, M., James, S., Javanbakht, M., Jee, S. H., Kasaeian, A., Kedir, M. S., Khader, Y. S., Khang, Y. H., Kim, D., Leigh, J., Linn, S., Lunevicius, R., El Razek, H. M. A., Malekzadeh, R., Malta, D. C., Marcenes, W., Markos, D., Melaku, Y. A., Meles, K. G., Mendoza, W., Mengiste, D. T., Meretoja, T. J., Miller, T. R., Mohammad, K. A., Mohammadi, A., Mohammed, S., Moradi-Lakeh, M., Nagel, G., Nand, D., Le Nguyen, Q., Nolte, S., Ogbo, F. A., Oladimeji, K. E., Oren, E., Pa, M., Park, E. K., Pereira, D. M., Plass, D., Qorbani, M., Radfar, A., Rafay, A., Rahman, M., Rana, S. M., Soreide, K., Satpathy, M., Sawhney, M., Sepanlou, S. G., Shaikh, M. A., She, J., Shiue, I., Shore, H. R., Shrimme, M. G., So, S., Soneji, S., Stathopoulou, V., Stroumpoulis, K., Sufiyan, M. B., Sykes, B. L., Tabares-Seisdedos, R., Tadese, F., Tedla, B. A., Tessema, G. A., Thakur, J. S., Tran, B. X., Ukwaja, K. N., Uzochukwu, B. S. C., Vlassov, V. V., Weiderpass, E., Wubshet Terefe, M., Yebo, H. G., Yimam, H. H., Yonemoto, N., Younis, M. Z., Yu, C., Zaidi, Z., Zaki, M. E. S., Zenebe, Z. M., Murray, C. J. L., & Naghavi, M. (2017). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability,

and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol*, 3(4), 524-548.

4. Coteur, G., Feagan, B., Keininger, D. L., & Kosinski, M. (2009). Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. *Aliment Pharmacol Ther*, 29(9), 1032-1041.
5. van de Willige, G., Wiersma, D., Nienhuis, F. J., & Jenner, J. A. (2005). Changes in quality of life in chronic psychiatric patients: a comparison between EuroQol (EQ-5D) and WHOQoL. *Qual Life Res*, 14(2), 441-451.
6. Dolan, P. (1997). Modeling valuations for EuroQol health states. *Med Care*, 35(11), 1095-1108.
7. Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J. A., Abdalla, S., Aboyans, V., Abraham, J., Ackerman, I., Aggarwal, R., Ahn, S. Y., Ali, M. K., Alvarado, M., Anderson, H. R., Anderson, L. M., Andrews, K. G., Atkinson, C., Baddour, L. M., Bahalim, A. N., Barker-Collo, S., Barrero, L. H., Bartels, D. H., Basanez, M. G., Baxter, A., Bell, M. L., Benjamin, E. J., Bennett, D., Bernabe, E., Bhalla, K., Bhandari, B., Bikbov, B., Bin Abdulhak, A., Birbeck, G., Black, J. A., Blencowe, H., Blore, J. D., Blyth, F., Bolliger, I., Bonaventure, A., Boufous, S., Bourne, R., Boussinesq, M., Braithwaite, T., Brayne, C., Bridgett, L., Brooker, S., Brooks, P., Brugha, T. S., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Buckle, G., Budke, C. M., Burch, M., Burney, P., Burstein, R., Calabria, B., Campbell, B., Canter, C. E., Carabin, H., Carapetis, J., Carmona, L., Cella, C., Charlson, F., Chen, H., Cheng, A. T., Chou, D., Chugh, S. S., Coffeng, L. E., Colan, S. D., Colquhoun, S., Colson, K. E., Condon, J., Connor, M. D., Cooper, L. T., Corriere, M., Cortinovis, M., de Vaccaro, K. C., Couser, W., Cowie, B. C., Criqui, M. H., Cross, M., Dabhadkar, K. C., Dahiya, M., Dahodwala, N., Damsere-Derry, J., Danaei, G., Davis, A., De Leo, D., Degenhardt, L., Dellavalle, R., Delossantos, A., Denenberg, J., Derrett, S., Des Jarlais, D. C., Dharmaratne, S. D., Dherani, M., Diaz-Torne, C., Dolk, H., Dorsey, E. R., Driscoll, T., Duber, H., Ebel, B., Edmond, K., Elbaz, A., Ali, S. E., Erskine, H., Erwin, P. J., Espindola, P., Ewoigbokhan, S. E., Farzadfar, F., Feigin, V., Felson, D. T., Ferrari, A., Ferri, C. P., Fevre, E. M., Finucane, M. M., Flaxman, S., Flood, L., Foreman, K., Forouzanfar, M. H., Fowkes, F. G., Franklin, R., Fransen, M., Freeman, M. K., Gabbe, B. J., Gabriel, S. E., Gakidou, E., Ganatra, H. A., Garcia, B., Gaspari, F., Gillum, R. F., Gmel, G., Gosselin, R., Grainger, R., Groeger, J., Guillemin, F., Gunnell, D., Gupta, R., Haagsma, J., Hagan, H., Halasa, Y. A., Hall, W., Haring, D., Haro, J. M., Harrison, J. E., Havmoeller, R., Hay, R. J., Higashi, H., Hill, C., Hoen, B., Hoffman, H., Hotez, P. J., Hoy, D., Huang, J. J., Ibeanusi, S. E., Jacobsen, K. H., James, S. L., Jarvis, D., Jasrasaria, R., Jayaraman, S., Johns, N., Jonas, J. B., Karthikeyan, G., Kassebaum, N., Kawakami, N., Keren, A., Khoo, J. P., King, C. H., Knowlton, L. M., Kobusingye, O., Koranteng, A., Krishnamurthi, R., Laloo, R., Laslett, L. L., Lathlean, T., Leasher, J. L., Lee, Y. Y., Leigh, J., Lim, S. S., Limb, E., Lin, J. K., Lipnick, M., Lipshultz, S. E., Liu, W., Loane, M., Ohno, S. L., Lyons, R., Ma, J., Mabweijano, J., MacIntyre, M. F., Malekzadeh, R., Mallinger, L., Manivannan, S., Marcenes, W., March, L., Margolis, D. J., Marks, G. B., Marks, R., Matsumori, A., Matzopoulos, R., Mayosi, B. M., McAnulty, J. H., McDermott, M. M., McGill, N., McGrath, J., Medina-Mora, M. E., Meltzer, M., Mensah, G. A., Merriman, T. R., Meyer, A. C., Miglioli, V., Miller, M., Miller, T. R., Mitchell, P. B., Mocumbi, A. O., Moffitt, T. E., Mokdad, A. A., Monasta, L., Montico, M., Moradi-Lakeh, M., Moran, A., Morawska, L., Mori, R., Murdoch, M. E.,

- Mwaniki, M. K., Naidoo, K., Nair, M. N., Naldi, L., Narayan, K. M., Nelson, P. K., Nelson, R. G., Nevitt, M. C., Newton, C. R., Nolte, S., Norman, P., Norman, R., O'Donnell, M., O'Hanlon, S., Olives, C., Omer, S. B., Ortblad, K., Osborne, R., Ozgediz, D., Page, A., Pahari, B., Pandian, J. D., Rivero, A. P., Patten, S. B., Pearce, N., Padilla, R. P., Perez-Ruiz, F., Perico, N., Pesudovs, K., Phillips, D., Phillips, M. R., Pierce, K., Pion, S., Polanczyk, G. V., Polinder, S., Pope, C. A., 3rd, Popova, S., Porrini, E., Pourmalek, F., Prince, M., Pullan, R. L., Ramaiah, K. D., Ranganathan, D., Razavi, H., Regan, M., Rehm, J. T., Rein, D. B., Remuzzi, G., Richardson, K., Rivara, F. P., Roberts, T., Robinson, C., De Leon, F. R., Ronfani, L., Room, R., Rosenfeld, L. C., Rushton, L., Sacco, R. L., Saha, S., Sampson, U., Sanchez-Riera, L., Sanman, E., Schwebel, D. C., Scott, J. G., Segui-Gomez, M., Shahraz, S., Shepard, D. S., Shin, H., Shivakoti, R., Singh, D., Singh, G. M., Singh, J. A., Singleton, J., Sleet, D. A., Sliwa, K., Smith, E., Smith, J. L., Stapelberg, N. J., Steer, A., Steiner, T., Stolk, W. A., Stovner, L. J., Sudfeld, C., Syed, S., Tamburlini, G., Tavakkoli, M., Taylor, H. R., Taylor, J. A., Taylor, W. J., Thomas, B., Thomson, W. M., Thurston, G. D., Tleyjeh, I. M., Tonelli, M., Towbin, J. A., Truelsen, T., Tsilimbaris, M. K., Ubeda, C., Undurraga, E. A., van der Werf, M. J., van Os, J., Vavilala, M. S., Venketasubramanian, N., Wang, M., Wang, W., Watt, K., Weatherall, D. J., Weinstock, M. A., Weintraub, R., Weisskopf, M. G., Weissman, M. M., White, R. A., Whiteford, H., Wiersma, S. T., Wilkinson, J. D., Williams, H. C., Williams, S. R., Witt, E., Wolfe, F., Woolf, A. D., Wulf, S., Yeh, P. H., Zaidi, A. K., Zheng, Z. J., Zonies, D., Lopez, A. D., Murray, C. J., AlMazroa, M. A., & Memish, Z. A. (2012). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380(9859), 2163-2196.
8. Weernink, M. G. M., Groothuis-Oudshoorn, C. G. M., IJzerman, M. J., & van Til, J. A. (2016). Valuing Treatments for Parkinson Disease Incorporating Process Utility: Performance of Best-Worst Scaling, Time Trade-Off, and Visual Analogue Scales. *Value in Health*, 19(2), 226-232.
 9. Burstrom, K., Sun, S., Gerdtham, U. G., Henriksson, M., Johannesson, M., Levin, L. A., & Zethraeus, N. (2014). Swedish experience-based value sets for EQ-5D health states. *Qual Life Res*, 23(2), 431-442.
 10. EuroQol, G. (Checked at 16th Feb 2019). Overview of available value sets for EQ-5D-3L. from <https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/valuation/>
 11. Zhuo, L., Xu, L., Ye, J., Sun, S., Zhang, Y., Burstrom, K., & Chen, J. (2018). Time Trade-Off Value Set for EQ-5D-3L Based on a Nationally Representative Chinese Population Survey. *Value Health*, 21(11), 1330-1337.
 12. Green, C., Brazier, J., & Deverill, M. (2000). Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics*, 17(2), 151-165.
 13. Craig, B. M. (2009). The duration effect: a link between TTO and VAS values. *Health Econ*, 18(2), 217-225.
 14. Stiggelbout, A. M., Eijkemans, M. J. C., Kiebert, G. M., Kievit, J., Leer, J. W. H., & DeHaes, H. J. C. J. M. (1996). The 'utility' of the visual analog scale in medical decision making and technology assessment - Is it an alternative to the time trade-off? *International Journal of Technology Assessment in Health Care*, 12(2), 291-298.

15. Lee, J. Y., & Lee, E. (2016). Assessment of Utility for Heart Failure Using Visual Analogue Scale(Vas), Time-Trade Off(Tto) and Euroqol-5 Dimension(Eq-5d) in the Korean General Population. *Value in Health*, 19(7), A868-A869.
16. Wang, X., Zhuo, L., Ma, Y., Cai, T., Must, A., Xu, L., & Zhuo, L. (2020). Similar responses to EQ-5D-3L by two elicitation methods: visual analogue scale and time trade-off. *BMC Med Res Methodol*, 20(1), 118.
17. Devlin, N. J., Shah, K. K., Feng, Y., Mulhern, B., & van Hout, B. (2017). Valuing health-related quality of life: An EQ-5D-5L value set for England. *Health Econ*.
18. Golicki, D., Niewada, M., Karlinska, A., Buczek, J., Kobayashi, A., Janssen, M. F., & Pickard, A. S. (2014). Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQ VAS in stroke patients. *Qual Life Res*.
19. Oppe, M., Rand-Hendriksen, K., Shah, K., Ramos-Goni, J. M., & Luo, N. (2016). EuroQol Protocols for Time Trade-Off Valuation of Health Outcomes. *Pharmacoeconomics*.
20. Sun, S., Chen, J., Kind, P., Xu, L., Zhang, Y., & Burstrom, K. (2015). Experience-based VAS values for EQ-5D-3L health states in a national general population health survey in China. *Qual Life Res*, 24(3), 693-703.
21. Yang, Z., van Busschbach, J., Timman, R., Janssen, M. F., & Luo, N. (2017). Logical inconsistencies in time trade-off valuation of EQ-5D-5L health states: Whose fault is it? *PLoS One*, 12(9), e0184883.
22. Agota Szende, M. O., Nancy Devlin. (2007). EQ-5D Value Sets: Inventory, Comparative Review and User Guide (pp. 36-38). The Netherlands Springer.
23. Lamers, L. M., Stalmeier, P. F., Krabbe, P. F., & Busschbach, J. J. (2006). Inconsistencies in TTO and VAS values for EQ-5D health states. *Med Decis Making*, 26(2), 173-181.
24. Cleemput, I. (2010). A social preference valuations set for EQ-5D health states in Flanders, Belgium. *Eur J Health Econ*, 11(2), 205-213.
25. Wittrup-Jensen, K. U., Lauridsen, J., Gudex, C., & Pedersen, K. M. (2009). Generation of a Danish TTO value set for EQ-5D health states. *Scand J Public Health*, 37(5), 459-466.
26. Devlin, N., Hansen, P., & Herbison, P. (2000). Variations in self-reported health status: results from a New Zealand survey. *N Z Med J*, 113(1123), 517-520.
27. Boston, R. C., & Sumner, A. E. (2003). STATA: a statistical analysis system for examining biomedical data. *Adv Exp Med Biol*, 537, 353-369.
28. Leske, D. A., Hatt, S. R., & Holmes, J. M. (2010). Test-retest reliability of health-related quality-of-life questionnaires in adults with strabismus. *Am J Ophthalmol*, 149(4), 672-676.
29. Hussain, Z. I., Lim, M., & Stojkovic, S. (2014). The test-retest reliability of fecal incontinence severity and quality-of-life assessment tools. *Dis Colon Rectum*, 57(5), 638-644.
30. China, N. B. o. S. o. (2014). One Percent of Chinese Population of Census Data. from <http://data.stats.gov.cn/easyquery.htm?cn=C01>

31. Wooldridge, J. M. (2006). *Introductory econometrics: a modern approach*. Mason, OH: Thomson/South-Western.
32. Tsuchiya, A., Ikeda, S., Ikegami, N., Nishimura, S., Sakai, I., Fukuda, T., Hamashima, C., Hisashige, A., & Tamura, M. (2002). Estimating an EQ-5D population value set: the case of Japan. *Health Econ*, 11(4), 341-353.
33. Lee, Y. K., Nam, H. S., Chuang, L. H., Kim, K. Y., Yang, H. K., Kwon, I. S., Kind, P., Kweon, S. S., & Kim, Y. T. (2009). South Korean time trade-off values for EQ-5D health states: modeling with observed values for 101 health states. *Value Health*, 12(8), 1187-1193.
34. Augustovski, F. A., Irazola, V. E., Velazquez, A. P., Gibbons, L., & Craig, B. M. (2009). Argentine valuation of the EQ-5D health states. *Value Health*, 12(4), 587-596.
35. Xie, F., Gaebel, K., Perampaladas, K., Doble, B., & Pullenayegum, E. (2014). Comparing EQ-5D valuation studies: a systematic review and methodological reporting checklist. *Med Decis Making*, 34(1), 8-20.
36. Buchholz, I., Thielker, K., Feng, Y. S., Kupatz, P., & Kohlmann, T. (2014). Measuring changes in health over time using the EQ-5D 3L and 5L: a head-to-head comparison of measurement properties and sensitivity to change in a German inpatient rehabilitation sample. *Qual Life Res*.
37. Bailey, H., & Kind, P. (2010). Preliminary findings of an investigation into the relationship between national culture and EQ-5D value sets. *Qual Life Res*, 19(8), 1145-1154.
38. Jo, M. W., Yun, S. C., & Lee, S. I. (2008). Estimating quality weights for EQ-5D health states with the time trade-off method in South Korea. *Value Health*, 11(7), 1186-1189.
39. Peltola, H. (2000). Worldwide *Haemophilus influenzae* type b disease at the beginning of the 21st century: global analysis of the disease burden 25 years after the use of the polysaccharide vaccine and a decade after the advent of conjugates. *Clin Microbiol Rev*, 13(2), 302-317.
40. Shaw, J. W., Johnson, J. A., Chen, S., Levin, J. R., & Coons, S. J. (2007). Racial/ethnic differences in preferences for the EQ-5D health states: results from the U.S. valuation study. *J Clin Epidemiol*, 60(5), 479-490.
41. Wikipedia-Xiaojing. (2019). Classic-Filial-Piety. from https://en.wikipedia.org/wiki/Classic_of_Filial_Piety
42. Cai, J., Zhao, H., & Coyte, P. C. (2017). Socioeconomic Differences and Trends in the Place of Death among Elderly People in China. *Int J Environ Res Public Health*, 14(10).
43. Wang, Z., Du, W., Pang, L., Zhang, L., Chen, G., & Zheng, X. (2015). Wealth Inequality and Mental Disability Among the Chinese Population: A Population Based Study. *Int J Environ Res Public Health*, 12(10), 13104-13117.
44. Yin, Z., Li, J., Ma, K., Ning, X., Chen, H., Fu, H., Zhang, H., Wang, C., Bruera, E., & Hui, D. (2017). Development of Palliative Care in China: A Tale of Three Cities. *Oncologist*, 22(11), 1362-1367.
45. Jin, X., Liu, G. G., Luo, N., Li, H., Guan, H., & Xie, F. (2016). Is bad living better than good death? Impact of demographic and cultural factors on health state preference. *Qual Life Res*, 25(4), 979-986.

46. Sen, A. (2002). Health: perception versus observation. *BMJ*, 324(7342), 860-861.
47. Symon, Z., Daignault, S., Symon, R., Dunn, R. L., Sanda, M. G., & Sandler, H. M. (2006). Measuring patients' expectations regarding health-related quality-of-life outcomes associated with prostate cancer surgery or radiotherapy. *Urology*, 68(6), 1224-1229.
48. Augestad, L. A., Rand-Hendriksen, K., Stavem, K., & Kristiansen, I. S. (2013). Time trade-off and attitudes toward euthanasia: implications of using 'death' as an anchor in health state valuation. *Qual Life Res*, 22(4), 705-714.

Figures

Figure 1

Predicted value of Model2 compared with other related value sets

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [07SupplementaryTables.doc](#)

- [08SupplementaryFigureS1VisualAid.tif](#)
- [09SupplementaryFigureS2PredictedvaluesfromM1toM6.tif](#)
- [10SupplementaryFigureS3meanscomparedwithpredicted.tif](#)
- [11Titlesofsupplementaryfigures.doc](#)