Health and Quality of Life Outcomes

RESEARCH

Open Access

Mapping the ADDQoL to the EQ-5D-5L and SF-6Dv2 among Chinese patients with type 2 diabetes mellitus

Haoran Fang[1†], Tianqi Hong[2†], Xinran Liu[1,3], Chang Luo[1,3], Yuanyuan Hou[1,3*] and Shitong Xie[1,3*]

Abstract

Objective The Audit of Diabetes-Dependent Quality of Life (ADDQoL) is a widely used instrument for assessing quality of life in Type 2 Diabetes Mellitus (T2DM). However, it does not directly yield health utility values essential for economic evaluations. This study developed mapping algorithms to predict EQ-5D-5L and SF-6Dv2 utility values from ADDQoL scores in T2DM patients in China.

Methods Cross-sectional data from 800 T2DM patients in China, stratified by age, sex, and geographical region, were divided into development (80%) and validation (20%) groups. Pearson correlation analyses were conducted to assess the conceptual overlap between ADDQoL and the EQ-5D-5L and SF-6Dv2. Six predictor sets and six regression methods were explored to map ADDQoL scores to EQ-5D-5L and SF-6Dv2 utility values, respectively. Model performance was evaluated using mean absolute error (MAE), root mean square error (RMSE), and intraclass correlation coefficient (ICC).

Results For the development group, the mean (SD) ADDQoL Average Weighted Impact (AWI) score was − 2.426 (1.052), and the mean (SD) utility values for EQ-5D-5L and SF-6Dv2 were 0.928 (0.092) and 0.791 (0.133), respectively. Among all 36 alternative mapping models each for EQ-5D-5L and SF-6Dv2, the best performance was consistently observed in the two-part models that included the ADDQoL AWI, the first overview item, and their squared terms. For the algorithm mapping to EQ-5D-5L utility values, it achieved a MAE of 0.067, a RMSE of 0.095, and an ICC of 0.414; For the algorithm mapping to SF-6Dv2 utility values, the corresponding metrics were an MAE of 0.099, an RMSE of 0.120, and an ICC of 0.517.

Conclusions This study provides a mapping framework to estimate EQ-5D-5L and SF-6Dv2 utility values from ADDQoL scores. These algorithms could be used to support economic evaluations, specifically tailored for Chinese T2DM populations.

[†]Haoran Fang and Tianqi Hong contributed equally to this work.

*Correspondence:
Yuanyuan Hou
helenhou86@163.com
Shitong Xie
xiest@tju.edu.cn

Full list of author information is available at the end of the article

Highlights

- The Audit of Diabetes-Dependent Quality of Life (ADDQoL) is a widely recognized, disease-specific instrument for assessing the quality of life in individuals with Type 2 Diabetes Mellitus (T2DM).
- While despite its widespread use, the ADDQoL does not directly produce health utility values, which are critical for economic evaluations in healthcare decision-making.
- This study developed mapping algorithms to predict EQ-5D-5L and SF-6Dv2 utility values from ADDQoL scores among T2DM patients in China.

Keywords Mapping, Type 2 diabetes mellitus, ADDQoL, EQ-5D-5L, SF-6Dv2, HRQoL

Introduction

Type 2 Diabetes Mellitus (T2DM), one of the most prevalent chronic metabolic diseases, is characterized by persistent hyperglycemia caused by insulin resistance and inadequate insulin secretion [1]. Between 1990 and 2022, the global prevalence of diabetes among adults doubled from 7 to 14%, affecting over 830 million people [1, 2]. In China, the prevalence of T2DM has increased more than tenfold over the past 40 years, driven by a large population base, rapid socioeconomic development, and urbanization, resulting in over 140.9 million cases — the highest worldwide [2, 3]. Diabetes contributes to numerous complications, including cardiovascular diseases, inflammation, and liver disorders, leading to reduced quality of life, premature mortality, and significant financial risk for patients [4]. Managing T2DM complications also imposes substantial costs on China's healthcare system [5], with projections indicating that total diabetes-related costs will rise from $250.2 billion in 2020 to $460.4 billion in 2030, reflecting an annual growth rate of 6.32%. Meanwhile, the per capita economic burden is expected to increase from $231 to $414, with a 6.02% annual growth rate [6].

To address these challenges, health systems may consider implementing cost-effective strategies to optimize healthcare intervention allocation and alleviate the disease burden. Cost-utility analysis (CUA) is commonly employed in health economic evaluations to assess the values of healthcare interventions [7]. The core of CUA lies in calculating quality-adjusted life years (QALYs) to compare the health gains of interventions within or across conditions. The measurement of health utility values is critical for the calculation of QALYs [8]. Health utility values, lying on a scale of 0 (death) to 1 (full health), can reflect health-related quality of life (HRQoL), a multidimensional concept assessing health status through physical, psychological, and social functioning [9]. Health utility is typically measured by generic preference-based measures (GPBM), with the EQ-5D and SF-6D being the most commonly applied measures, These measures have been extensively validated and adopted in utility measurement and economic evaluations worldwide [10, 11, 12]. China Guidelines for Pharmacoeconomic Evaluations (2020 edition) also recommends using the EQ-5D and SF-6D for assessing health utilities among the Chinese population [13, 14].

However, previous studies have indicated that GPBMs tend to be more responsive to co-morbid conditions than to the specific HRQoL factors associated with particular diseases [15, 16]. Furthermore, several studies have shown that the sensitivity of GPBMs may decrease in populations, which affected by more severe diseases or conditions having particular dimensions not adequately captured by GPBMs [17, 18]. Clinical trials and empirical studies focusing on specific diseases commonly include disease-specific measures to assess the HRQoL of patients to collect more sensitive results [16]. While non-preference-based disease-specific HRQoL measures do not link health states to utility values. A feasible solution is to use mapping (or "crosswalk") functions to convert these disease-specific HRQoL data into utility values compatible with GPBMs [16, 19]. Studies on mapping in the field of diabetes remains relatively limited. In 2014, Vokó et al. mapped the Nottingham Health Profile (NHP), a non-diabetes-specific instrument for assessing HRQoL across physical and emotional dimensions, directly to the EQ-5D-3 L based on data from T2DM patients ($N = 943$) in Hungary [20]. The best-fitting models included all the NHP statements as predictors, demonstrating mapped utility values ranging from − 0.23 to 1.05 (adjusted $R^2 = 0.68$, RMSE = 0.174). In China, the only existing mapping study, developed by Wu et al. in 2022, mapped the Diabetes-Specific Quality of Life (DSQL) to the EQ-5D-3 L utility values ($N = 493$) [21]. Its optimal algorithm was constructed through the Censored Least Absolute Deviations (CLAD) model, which achieved an adjusted R^2 of 0.4296, a mean absolute error (MAE) of 0.0345, and mapped utility values ranging from 0.63 to 1.

While there are limited studies evaluating the measurement properties of DSQL. The Audit of Diabetes-Dependent Quality of Life (ADDQoL) is another instrument widely used around the world for assessing the impact of diabetes on patients' quality of life, integrating general quality of life assessments with diabetes-specific effects [22, 23, 24, 25, 26, 27]. It has demonstrated strong measurement properties in Chinese populations [28, 29, 30,

31, 32]. Nevertheless, there is a lack of empirical studies that developing a mapping function to link ADDQoL scores to GPBMs among T2DM population.

Therefore, this study aims to develop mapping algorithms to directly predict EQ-5D-5L and SF-6Dv2 utility values based on ADDQoL scores among patients with T2DM in China.

Methods

The methods used in this study adhered to the guidelines outlined by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Good Practices for Outcomes Research Task Force Report on mapping, as well as the technical support document from the National Institute for Health and Care Excellence (NICE) on using mapping methods to estimate utility values [19, 33]. This study also applied the Mapping onto Preference-based Measures (MAPS) checklist [34], detailed in Appendix Table 1. The study protocol was approved by the Academic Ethics Committee of Tianjin University (No. TJUE-2023-206), and all participants provided informed consent.

Data collection

This analysis is based on data collected from a national survey conducted between October 2023 and January 2024, which aimed to assess the health status and treatment preferences of T2DM patients in China. The target sample size was 800, recruited from eight cities (Beijing, Guangzhou, Shanghai, Baoding, Chongqing, Jiujiang, Mianyang and Wuhan) and surrounding rural areas across China to ensure geographic diversity and representation of varying levels of economic development. A quota sampling method was used to recruit the sample stratified by age and sex of patients with T2DM in China [35]. Online recruitment and street intercept interviews were used for the data collection, with both methods complemented by face-to-face interviews to ensure sample representativeness and data validity. Street intercept interview improved accessibility for random respondent recruitment, allowing offline face-to-face participation for individuals with low literacy or those unfamiliar with/lacking access to online platforms, while enabling easier large-scale sample collection at a lower cost. However, street intercept interviews might be limited to certain types of patients, such as those with severe conditions and mobility impairments. Therefore, we also utilized face-to-face online recruitment to balance potential selection bias [36].

All interviews were conducted by trained interviewers, whose responsibilities included: (1) introducing the study purpose and questionnaire structure, (2) obtaining informed consent, (3) providing scripted clarifications based on pre-defined protocols only when participants explicitly struggled to understand a question's wording or intent, and (4) supervising questionnaire completion to ensure data validity. Inclusion criteria required participants to be (1) aged ≥ 18 years, (2) diagnosed with T2DM for at least three months, (3) free from cognitive burdens and able to use online devices, and (4) willing to provide informed consent.

Eligible participants completed a questionnaire specifically designed for individuals with T2DM, with the help provided by interviewers if needed. The questionnaire first collected sociodemographic characteristics (e.g., age, sex, and education), and general information related to lifestyle and health behaviors (e.g., BMI). Self-reported answers of the EQ-5D-5L and SF-6Dv2 were then obtained with a randomized order. Additional diabetes-related information, such as duration since diagnosis, fasting blood glucose levels, and the presence of complications, was recorded. Finally, participants completed the ADDQoL. Data were subjected to quality control criteria, including (1) no missing responses and (2) provided correct answer to a question regarding identifying T2DM medications from several options of medications for various diseases. Data not meeting these criteria were excluded from the study.

Source and target measures

ADDQoL serves as the source measure, and both the EQ-5D-5L and SF-6Dv2 serve as the target measures in this study. The EQ-5D-5L has been demonstrated to have better measurement properties than the EQ-5D-3 L among the Chinese and international populations [37, 38, 39]. For the SF-6D, only the SF-6Dv2 has the officially validated Chinese version and corresponding utility value set in mainland China [40, 41]. Therefore, these two measures were chosen in this study.

The ADDQoL is a specialized instrument designed for adults and adolescents aged 16 and above with Type 1 or Type 2 Diabetes, focusing on those who are no longer in school [29, 32, 42]. The instrument consists of two sections: the first assesses general quality of life (QoL) and overall diabetes-specific impact using two overview items (OI1, OI2), while the second examines diabetes' effects on 19 specific domains, including "yes/no" applicability filters for certain domains, to prevent respondents from having no life experience in that area. Each respondent needs to report the impact and importance ratings of diabetes on each domain. Weighted impact (WI) scores are derived by combining impact and importance ratings, ranging from −9 (maximum negative impact) to +3 (maximum positive impact). The Average Weighted Impact (AWI) score is subsequently calculated as the mean of applicable WI scores [42].

The EQ-5D-5L is a widely used GPBM that assesses HRQoL across five dimensions: Mobility (MO), Self-care

(SC), Usual activities (UA), Pain/discomfort (PD), and Anxiety/depression (AD), each with five response levels, creating 3,125 possible health states. It also includes a Visual Analogue Scale ranging from 0 to 100, representing the respondent's subjective health evaluation [43]. The utility value set has been developed in China using the time trade-of (TTO) approach, with the range of −0.391 (55,555) to 1 (11,111) [44].

The SF-6Dv2, derived from the SF-36, measures HRQoL across six dimensions: Physical functioning (PF), Role limitations (RL), Social functioning (SF), Pain (PN), Mental health (MH), and Vitality (VT). While the Pain dimension includes six response levels, others have five, allowing for 18,750 possible health states [45]. The validated Chinese version ensures conceptual equivalence to the English version [40], and the Chinese SF-6Dv2 value set has been developed based on the time trade-of (TTO) approach, range from − 0.277 to 1 [41].

Data analysis

The total sample was divided into a development group (80%) and a validation group (20%) through random allocation [46]. T-tests and Fisher's exact tests were conducted to compare demographic characteristics, health-related information, diabetes-related features, and scores from ADDQoL AWI, EQ-5D-5L, and SF-6Dv2 between the two groups. Continuous variables were reported as means with standard deviations (SD), while categorical variables were presented as frequencies and percentages [47, 48]. Histograms were created to visualize the distributions of the EQ-5D-5L and SF-6Dv2 responses. Data normality was also assessed. When normality assumptions were not met, the Shapiro-Wilk test was applied to evaluate deviations from normality [49].

Before developing mapping algorithms, the conceptual overlap between the source and target measures was assessed. Pearson correlation coefficients (r) were calculated to examine the associations between ADDQoL scores and EQ-5D-5L or SF-6Dv2 utility values [50]. Correlation strength was categorized as strong ($r \geq 0.5$), moderate ($0.3 \leq r < 0.5$), or weak ($r < 0.3$) [51].

Model specification

This study applied a direct mapping method to convert ADDQoL onto EQ-5D-5L and SF-6Dv2 utility values, where predictor sets and regression methods serve as the two key components for direct mapping.

This study developed six predictor sets to systematically explore the optimal mapping models. Predictor set 1 used only the ADDQoL AWI score as a continuous independent variable. Predictor set 2 employed stepwise regression to select key variables from the WI scores of 19 items. Predictor set 3 incorporated the two overview items, OI1 and OI2 alongside the AWI score, integrating them for analysis. To account for potential nonlinear effects, Predictor set 4 added squared terms to Predictor set 3, preserving the positive or negative characteristics of OI1, OI2, and AWI through a sign function. Expanding on this, Predictor set 5 included cubic terms for more detailed modeling of nonlinear relationships. Lastly, Predictor set 6 combined OI1, OI2, and the critical WI items identified through stepwise regression. All predictor sets were refined using backward stepwise regression at a significance level of 0.05 to ensure explanatory strength [52]. Predictors with unexpected directional coefficients (e.g., higher QoL correlating with lower utility values) were excluded to maintain logical consistency and model robustness [53].

Six regression methods were also applied for analysis: Ordinary Least Squares (OLS), Tobit model, Censored Least Absolute Deviations (CLAD), Generalized Linear Model (GLM), Two-Part Model (TPM), and Beta Regression Mixture Model (BM). OLS, the most common method, minimizes squared errors between observed and predicted values but requires strict assumptions, including linearity, homoscedasticity, and normal residuals [54, 55]. The Tobit model accounts for the right-censoring of utility value at 1, capturing underlying distribution characteristics suitable for censored linear data [56, 57]. CLAD provides robust conditional median estimates by minimizing absolute errors, free from normality or homoscedasticity assumptions, making it ideal for censored or truncated data [58]. GLM offers flexibility with link functions to model nonlinear relationships and variance functions to handle heteroscedasticity [47]. Based on modified Park tests and Box-Cox methods, GLM used a Gamma distribution and power link function for EQ-5D-5L and SF-6Dv2 to ensure consistent estimates [59]. TPM excelled with ceiling effects by using logistic regression for predicting full health and OLS for partial health, combining these for overall predictions [19, 60]. Lastly, BM, comprising a beta mixture model for continuous data and a multinomial logit for boundary data, demonstrated notable advantages in utility values mapping by accommodating variable distributions and the 0–1 utility values range, aligning with utility values characteristics [61, 62].

Model performance

36 candidate models were developed each for predicting EQ-5D-5L and SF-6Dv2 utility value by combining six regression methods with six predictor sets. The optimal models for the EQ-5D-5L and SF-6Dv2 were selected in both the development and validation groups based on three criteria. First, predicted utility values had to fall within a reasonable range, with maximum value not greater than 1, and minimum value approximating the lowest assessable values of the EQ-5D-5L (-0.184) and

SF-6Dv2 (-0.179). Second, the model has satisfied prediction accuracy, which was measured using mean absolute error (MAE) and root mean square error (RMSE), with lower values indicating better models; as well as intraclass correlation coefficients (ICC), with higher values indicating better models [63]. ICC was calculated using a bidirectional random-effects model with absolute agreement to assess the consistency between predicted and observed values [46]. For each model, the MAE, RMSE, and ICC values were ranked separately in both the development and validation groups. The average ranking orders for each of these three metrics was then calculated, and an overall average ranking (AR) was derived based on these average orders [64]. Third, among models with crossover predictive performance in the development group and validation group, the simplest model was preferred.

It is important to note that Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are only comparable within models sharing the same likelihood function framework and dataset [46, 56]. Since AIC and BIC could not be applied to CLAD due to its lack of likelihood estimation, and differences in model assumptions, such as the Tobit model's handling of right-censoring and GLM's ability to accommodate non-normal distributions, which further limited direct comparisons. For TPM, separate AIC and BIC values were computed for each component, complicating comparisons with single-framework models [46, 56]. Consequently, this study emphasized MAE, RMSE, ICC, and AR as the primary criteria for model evaluation.

For more visualized comparisons, scatter plots, bar charts, and Bland-Altman plots were also employed. Scatter plots and bar charts were used to explore trends, biases, and agreement between predicted and observed utility values [41, 65]. Bland-Altman plots provided insights into the average differences between predicted and observed values, emphasizing the distribution of bias and its potential impact [66].

All analyses were conducted using Stata 16.0 (StataCorp LLC, College Station, TX, USA).

Results

Descriptive statistics

Of the 1353 respondents invited to participate in the survey, 1045 respondents agreed to participate, with a response rate of 77.2%. 237 respondents were excluded due to ineligibility for the quotas ($N = 205$), or could not provide proof of diagnosis or medication ($N = 32$). Therefore, 808 respondents met the inclusion criteria. Eight respondents were excluded because they did not complete the interview. Finally, a total of 800 respondents were included in this study, with demographic details presented in Table 1. The mean (SD) age was 50.4 (11.9) years, with males accounting for 52.8% ($N = 422$). The mean (SD) BMI was 24.4 (3.8), falling within the interval for overweight, and the mean (SD) disease duration was 5.7 (4.9) years. Additionally, 57.6% ($N = 461$) of respondents reported a family history of diabetes. No significant differences were observed between the development group ($N = 640$) and validation group ($N = 140$) in terms of demographic characteristics, health-related life information, diabetes features, ADDQoL scores, and EQ-5D-5L and SF-6Dv2 utility values.

The mean (SD) ADDQoL AWI score was −2.426 (1.052), ranging from −5.375 to 0. The mean (SD) utility values for EQ-5D-5L and SF-6Dv2 were 0.928 (0.092) and 0.791 (0.133), respectively, with ranges of 0.365–1 and 0.136–1. Figure 1 presents the distributions of EQ-5D-5L and SF-6Dv2 utility values, which were found to deviate from normality according to the Shapiro-Wilk test ($P < 0.001$).

Conceptual overlap

Table 2 presents the results of Pearson correlation analyses between the source and target measures. Moderate correlations ($0.3 \le r < 0.5$) were observed between the utility values of EQ-5D-5L and SF-6Dv2 and the OI1 of ADDQoL. Additionally, significant negative correlations (-0.026 to -0.368) were observed between the dimension scores of EQ-5D-5L and SF-6Dv2 and both the AWI and OI1 scores of ADDQoL.

Mapping ADDQoL onto EQ-5D-5L utility values

The performance of 36 regression models mapping ADDQoL to EQ-5D-5L utility value was summarized in Table 3. In the development group, Predictor set 4 consistently emerged as the best-performing predictor set across all regression methods based on AR. In the validation group, models using Predictor set 4 were optimal for OLS, GLM, TPM, and BM, while inconsistencies were observed for Tobit (Predictor set 6) and CLAD (Predictor set 5). Considering the ARs of MAE, RMSE, and ICC, TPM4 demonstrated superior predictive accuracy and consistency in both the development (MAE = 0.056, RMSE = 0.079, ICC = 0.463, R^2 = 0.181, adjusted R^2 = 0.180) and validation groups (MAE = 0.067, RMSE = 0.095, ICC = 0.414, R^2 = 0.183, adjusted R^2 = 0.178), with predicted utility values range meeting expectations. Figure 2 illustrates the performance of the TPM4 in the validation set. Both the bar chart and Bland–Altman plot reveal that TPM4 consistently outperforms other models, demonstrating greater predictive accuracy and smaller discrepancies between predicted and observed utility values (Appendix Fig. 1).

Table 1 Socio-demographic characteristics of study sample

	Total sample (*n* = 800)	Development group(*n* = 640)	Validation group (*n* = 160)	*P* value
		N (%)		
Age				
Mean (SD)	50.4 (11.9)	50. 5 (11.7)	50.0 (12.6)	0.683
Age groups				
18–27	15 (1.9%)	11 (1.7%)	4 (2.5%)	0.096
28–37	99 (12.4%)	72 (11.3%)	27 (16.9%)	
38–47	248 (31.0%)	209 (32.7%)	39 (24.4%)	
48–57	166 (20.8%)	130 (20.3%)	36 (22.5%)	
58–67	216 (27.0%)	177 (27.7%)	39 (24.4%)	
>=68	56 (7.0%)	41 (6.4%)	15 (9.4%)	
Sex				
Male	422 (52.8%)	336 (52.5%)	86 (53.8%)	0.791
Regions				
Northern China	206 (25.8%)	160 (25.0%)	46 (28.8%)	0.489
Eastern China	112 (14.0%)	86 (13.4%)	26 (16.3%)	
Central China	184 (23.0%)	148 (23.1%)	36 (22.5%)	
Western China	184 (23.0%)	149 (23.3%)	35 (21.9%)	
Southern China	114 (14.3%)	97 (15.2%)	17 (10.6%)	
Education				
Primary school/below	9 (1.1%)	8 (1.3%)	1 (0.6%)	0.788
Junior high school	75 (9.4%)	58 (9.1%)	17 (10.6%)	
High school	270 (33.8%)	213 (33.3%)	57 (35.6%)	
University/above	446 (55.8%)	361 (56.4%)	85 (53.1%)	
Family history of diabetes				
Yes	461 (57.6%)	374 (58.4%)	87 (54.4%)	0.372
No	339 (42.4%)	266 (41.6%)	73 (45.6%)	
Disease duration				
Mean (SD)	5.7 (4.9)	5.6 (4.9)	5.8 (4.8)	0.775
Blood glucose				
Mean (SD)	7.6 (1.9)	7.5 (1.8)	7.8 (2.1)	0.058
BMI				
Mean (SD)	24.4 (3.8)	24.3 (3.8)	24.7 (3.7)	0.317
SF-6Dv2 utility				
Mean (SD)	0.791 (0.133)	0.793 (0.133)	0.782 (0.133)	0.361
Range (min, max)	0.136,1	0.136,1	0.395,1	
EQ-5D-5L utility				
Mean (SD)	0.928 (0.092)	0.931 (0.088)	0.916 (0.104)	0.070
Range (min, max)	0.365,1	0.365,1	0.431,1	
ADDQoL AWI				
Mean (SD)	-2.426 (1.052)	-2.441 (1.048)	-2.370 (1.065)	0.446
Range (min, max)	-5.375,0	-5.375,0	-4.529,0	

P value came from two independent sample t-test or Fisher's exact test between development group and validation group

SD standard deviation; BMI body mass index; ADDQoL AWI The average weighted impact score of the Audit of Diabetes-Dependent Quality of Life

Mapping ADDQoL onto SF-6Dv2 utility values

Table 4 outlines the performance of 36 regression models mapping ADDQoL to SF-6Dv2 utility values. In the development group, Predictor set 4 consistently emerged as the optimal predictor set across all regression methods. However, in the validation group, only TPM identified Predictor set 4 as the best-performing model, while other methods demonstrated inconsistencies. Considering the ARs of MAE, RMSE, and ICC, TPM4 achieved the highest predictive accuracy and consistency across both the development group (MAE = 0.094, RMSE = 0.117, ICC = 0.537, R^2 = 0.225, adjusted R^2 = 0.224) and the validation group (MAE = 0.099, RMSE = 0.120, ICC = 0.517, R^2 = 0.197, adjusted R^2 = 0.192), with predicted utility ranges aligning with expectations. Visual analyses, including bar plots and Bland–Altman plots, further highlight that TPM4 consistently provided predictions more closely aligned with observed

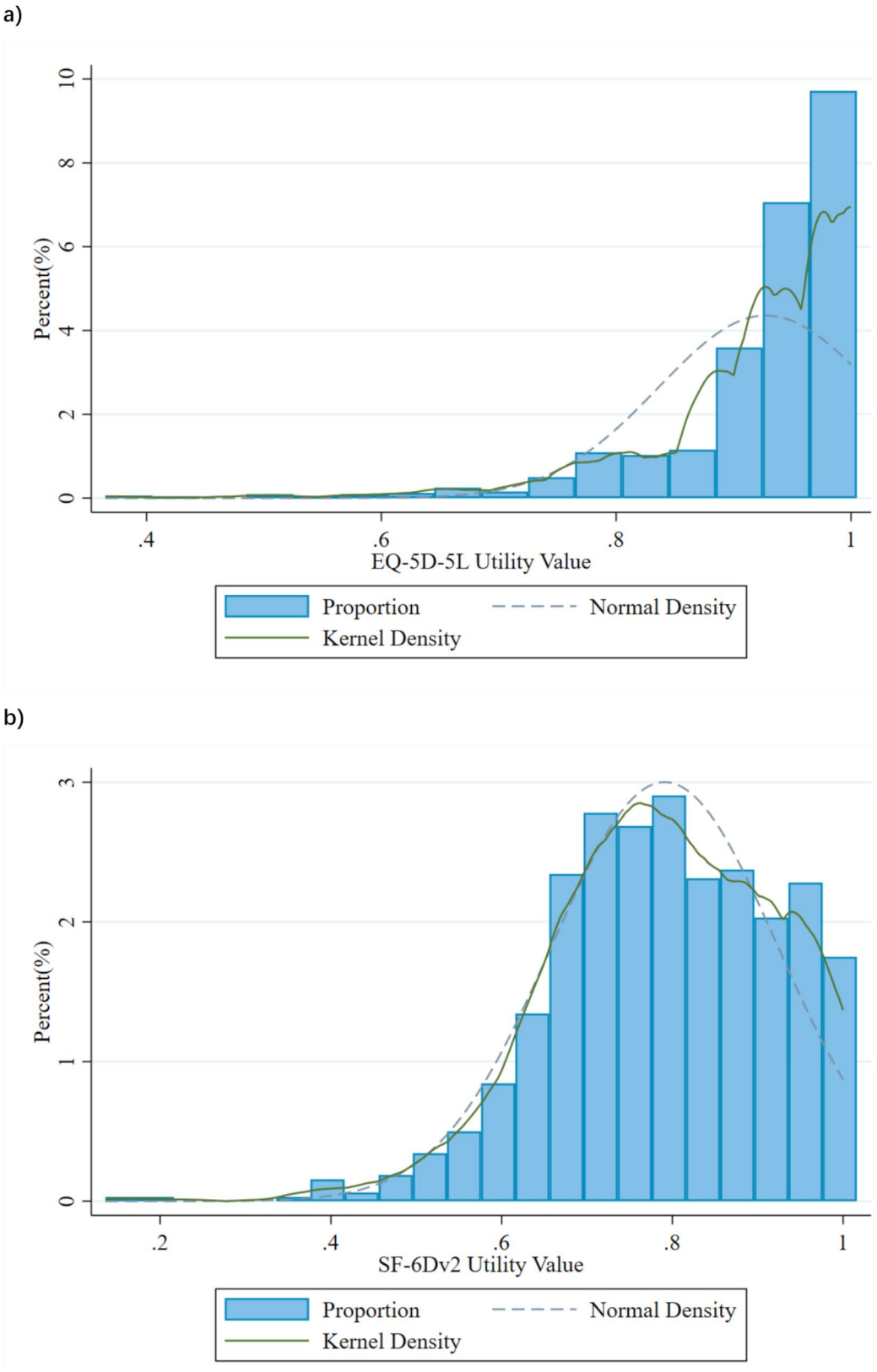

Fig. 1 Distribution of the EQ-5D-5L and the SF-6Dv2 utility value ($N = 800$). (**a**) The distribution of the EQ-5D-5L utility value ($N = 800$). (**b**) The distribution of the SF-6Dv2 utility value ($N = 800$)

Table 2 Pearson correlation coefficients between ADDQoL and EQ-5D-5L/SF-6Dv2

		ADDQoL OI1	ADDQoL OI2	ADDQoL AWI
SF-6Dv2	PF	-0.279	-0.011	-0.057
	RL	-0.368	-0.026	-0.095
	SF	-0.351	-0.036	-0.152
	PN	-0.317	0.003	-0.059
	MH	-0.299	-0.012	-0.076
	VT	-0.343	-0.108	-0.075
	Utility	**0.433**	0.044	0.106
EQ-5D-5L				
	MO	-0.229	-0.024	-0.092
	SC	-0.178	0.024	-0.026
	UA	-0.279	0.029	-0.074
	PD	-0.317	0.008	-0.059
	AD	-0.310	0.016	-0.045
	Utility	**0.375**	-0.002	0.088

Abbr: ADDQoL OI1, Overview item1; OI2, Overview item2; AWI, The average weighted impact score of the Audit of Diabetes-Dependent Quality of Life; MO, Mobility; SC, Self-care; UA, Usual Activities; PD, Pain/Discomfort; AD, Anxiety/Depression; PF, Physical Functioning; RL, Role Limitations; SF, Social Functioning; PN, Pain; MH, Mental Health; VT, Vitality

values (Fig. 2), outperforming other regression models in terms of accuracy and agreement in SF-6Dv2 mapping (Appendix Fig. 2).

Use of optimal mapping algorithms

The mapping algorithms from ADDQoL to EQ-5D-5L and SF-6Dv2 utility values both selected AWI, OI1, and their squared terms as the predictor set, with two-part models employed as the regression methods. The final mapping algorithms are defined as follows:

Defining sign function:

$$sign(x) = \begin{cases} 1, x > 0 \\ 0, x = 0 \\ -1, x < 0 \end{cases}$$

Mapping ADDQoL onto EQ-5D-5L.

(1) Prediction probability function:

$$logit(Pr(\text{DisEQ05D05L})) = 0.4895366 - 1.193757 \times OI1 - 0.7301487 \times AWI + 0.2285856 \times \left(\text{sign}\,(OI1) \times (OI1)^2\right) + 0.1427682 \times \left(\text{sign}\,(AWI) \times (AWI)^2\right)$$

$$\Pr\,(DisEQ - 5D - 5L > 0) = \frac{1}{1 + e^{-\log it(\Pr(DisEQ - 5D - 5L > 0))}}$$

(2) Expected value prediction of DisEQ-5D-5L:

$$E(Dis\text{EQ05D05L}|Dis\text{EQ05D05L} > 0) = 0.1055585 - 0.0820005 \times OI1 - 0.0297864 \times AWI + 0.0279237 \times (sign(OI1) \times (OI1)^2) + 0.0050551 \times (sign\,(AWI) \times \left(AWI\right)^2)$$

(3) Transformed EQ-5D-5L utility values:

$$Dis\,EQ - 5D - 5L = \Pr\,(Dis\,EQ - 5D - 5L > 0) \times E\,(Dis\;\; EQ - 5D - 5L|\,Dis\,EQ - 5D - 5L > 0)$$

$$Dis\,EQ - 5D - 5LUtility = 1 - Dis\,EQ - 5D - 5L$$

Mapping ADDQoL onto SF-6Dv2.

(1) Prediction probability function:

$$logit(Pr(\text{DisSF06Dv2})) = 3.629165 - 2.230134 \times OI1 - 0.8203494 \times AWI + 0.5761989 \times \left(\text{sign}\,(OI1) \times (OI1)^2\right) + 0.2054717 \times \left(\text{sign}\,(AWI) \times (AWI)^2\right)$$

$$Pr(DisSF06Dv2 > 0) = \frac{1}{1 + e^{-logit(Pr(DisSF06Dv2 > 0))}}$$

(2) Expected value prediction of DisSF-6Dv2:

Table 3 Model performance of six regression methods for mapping the ADDQoL to the EQ-5D-5L utility scores

Mapping Methods	Development group ($n=640$)						Validation group ($n=160$)					
	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR
OLS1	0.931 (0.006)	0.914,0.945	0.063	0.088	0.020	28	0.931 (0.006)	0.919,0.945	0.073	0.104	0.023	27
OLS2	0.931 (0.010)	0.905,0.952	0.062	0.087	0.050	25	0.930 (0.010)	0.909,0.943	0.073	0.104	0.030	26
OLS3	0.793 (0.059)	0.570,0.903	0.095	0.119	0.501	24	0.792 (0.062)	0.627,0.904	0.097	0.120	0.505	24
OLS4	**0.931 (0.037)**	**0.792,0.980**	**0.056**	**0.080**	**0.461**	**2**	**0.928 (0.037)**	**0.817,0.997**	**0.067**	**0.095**	**0.418**	**1**
OLS5	0.931 (0.036)	0.812,0.960	0.057	0.080	0.440	3	0.928 (0.037)	0.821,0.960	0.067	0.095	0.414	6
OLS6	0.931 (0.034)	0.802,0.989	0.058	0.081	0.414	16	0.929 (0.035)	0.831,0.984	0.067	0.096	0.390	8
Tobit1	0.967 (0.009)	0.943,0.987	0.062	0.095	-0.138	36	0.967 (0.009)	0.950,0.987	0.076	0.115	-0.175	36
Tobit2	0.967 (0.014)	0.931,0.996	0.062	0.095	-0.090	35	0.966 (0.014)	0.937,0.984	0.075	0.115	-0.160	35
Tobit3	0.958 (0.040)	0.777,1.000	0.056	0.085	0.408	17	0.956 (0.042)	0.830,1.000	0.066	0.102	0.369	12
Tobit4	**0.960 (0.048)**	**0.791,1.000**	**0.055**	**0.086**	**0.459**	**8**	0.957 (0.048)	0.810,1.000	0.066	0.103	0.384	14
Tobit5	0.961 (0.047)	0.789,1.000	0.057	0.087	0.435	19	0.958 (0.049)	0.816,1.000	0.067	0.103	0.394	15
Tobit6	0.957 (0.041)	0.771,1.000	0.056	0.085	0.414	11	**0.955 (0.043)**	**0.817,1.000**	**0.066**	**0.102**	**0.376**	**10**
CLAD1	0.947 (0.003)	0.940,0.954	0.059	0.089	-0.027	29	0.948 (0.003)	0.942,0.954	0.071	0.108	-0.068	31
CLAD2	0.947 (0.003)	0.939,0.954	0.059	0.089	-0.018	27	0.947 (0.003)	0.941,0.951	0.071	0.108	-0.070	30
CLAD3	0.945 (0.041)	0.797,1.000	0.055	0.083	0.426	8	0.944 (0.043)	0.848,1.000	0.064	0.100	0.386	8
CLAD4	**0.952 (0.043)**	**0.796,1.000**	**0.054**	**0.084**	**0.440**	**5**	0.952 (0.042)	0.838,1.000	0.065	0.103	0.332	13
CLAD5	0.946 (0.042)	0.794,1.000	0.055	0.083	0.436	6	**0.944 (0.044)**	**0.843,1.000**	**0.064**	**0.100**	**0.396**	**4**
CLAD6	0.962 (0.024)	0.871,1.000	0.056	0.088	0.230	21	0.962 (0.025)	0.893,1.000	0.067	0.107	0.168	23
GLM1	0.950 (0.003)	0.942,0.956	0.059	0.090	-0.037	32	0.950 (0.003)	0.944,0.956	0.071	0.109	-0.081	34
GLM2	0.950 (0.004)	0.938,0.959	0.059	0.089	-0.022	31	0.949 (0.005)	0.939,0.955	0.071	0.109	-0.076	33
GLM3	0.947 (0.027)	0.792,0.988	0.056	0.083	0.338	11	0.947 (0.028)	0.857,0.988	0.066	0.100	0.300	16
GLM4	**0.947 (0.030)**	**0.819,0.995**	**0.055**	**0.081**	**0.397**	**7**	**0.945 (0.030)**	**0.834,0.997**	**0.066**	**0.099**	**0.330**	**10**
GLM5	0.947 (0.028)	0.824,0.970	0.056	0.082	0.365	15	0.945 (0.029)	0.845,0.970	0.067	0.099	0.321	19
GLM6	0.947 (0.027)	0.793,0.986	0.056	0.083	0.340	11	0.946 (0.028)	0.854,0.984	0.066	0.100	0.295	18
PTM1	0.931 (0.007)	0.911,0.946	0.063	0.088	0.021	29	0.931 (0.007)	0.917,0.946	0.073	0.104	0.027	25
PTM2	0.931 (0.010)	0.903,0.951	0.062	0.087	0.051	26	0.930 (0.010)	0.908,0.943	0.073	0.105	0.029	27
PTM3	0.931 (0.032)	0.798,0.979	0.057	0.081	0.406	11	0.930 (0.034)	0.833,0.980	0.066	0.096	0.392	5
PTM4	**0.931 (0.036)**	**0.812,0.984**	**0.056**	**0.079**	**0.463**	**1**	**0.929 (0.037)**	**0.814,0.986**	**0.067**	**0.095**	**0.414**	**1**
PTM5	0.931 (0.035)	0.814,0.960	0.057	0.080	0.440	3	0.928 (0.036)	0.818,0.960	0.067	0.095	0.415	3
PTM6	0.931 (0.033)	0.792,0.980	0.057	0.081	0.419	10	0.929 (0.035)	0.821,0.977	0.067	0.096	0.397	6
BM1	0.926 (0.002)	0.922,0.930	0.064	0.088	0.004	34	0.927 (0.002)	0.923,0.930	0.074	0.104	0.006	32
BM2	0.927 (0.003)	0.920,0.932	0.064	0.088	0.012	32	0.926 (0.003)	0.920,0.929	0.074	0.104	0.009	29
BM3	0.926 (0.016)	0.841,0949	0.060	0.083	0.248	23	0.926 (0.017)	0.872,0.949	0.069	0.098	0.241	21
BM4	**0.927 (0.023)**	**0.835,0.954**	**0.057**	**0.081**	**0.349**	**17**	**0.925 (0.023)**	**0.836,0.955**	**0.068**	**0.096**	**0.319**	**17**

Table 3 (continued)

Mapping Methods	Development group ($n = 640$)						Validation group ($n = 160$)					
	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR
BM5	0.926 (0.022)	0.839,0.941	0.058	0.081	0.323	20	0.924 (0.022)	0.843,0.941	0.068	0.097	0.310	19
BM6	0.926 (0.016)	0.840,0.949	0.060	0.083	0.251	21	0.925 (0.017)	0.868,0.948	0.069	0.098	0.239	22

SD standard deviation, Min minimum predicted value, Max. maximum predicted value, AIC Akaike information criteria, BIC Bayesin formation criteria, MAE mean absolute error, RMSE root mean square error, ICC interclass correlation, AR Average rank: Converted from the average ranking score coefficient

1/2/3/4/5/6: Regression model using independent variable set1/2/3/4/5/6

Predictor set in bold indicated the best results among the regression method

$$E(DisSF06Dv2|DisSF06Dv2 > 0) = 0.2090763 - 0.1087293\times OI1 - 0.0527731\times AWI + 0.0251571\times (sign(OI1) \times (OI1)^2) + 0.0076547 \times (sign\,(AWI) \times \left(AWI\right)^2)$$

(3) Transformed SF-6Dv2 utility values:

$$DisSF06Dv2 = Pr(DisSF06Dv2 > 0)\times E(DisSF06Dv2|DisSF06Dv2 > 0)$$

$$SF06Dv2Utility = 1 - DisSF06Dv2$$

For example, when the ADDQoL AWI score is -0.125, OI1 score is 1, the probability that the EQ-5D-5L negative utility value is greater than 0 is 0.405, the expected EQ-5D-5L negative utility value is 0.055, and the predicted EQ-5D-5L utility value derived from the conversion is 0.978; The probability that the SF-6Dv2 negative utility value is greater than 0 is 0.888, the expected SF-6Dv2 negative utility value is 0.132, and the predicted SF-6Dv2 utility value derived from the conversion is 0.883.

Discussion

The ADDQoL is a widely utilized instrument for assessing HRQoL in individuals with Type 2 Diabetes Mellitus. However, to date, no mapping algorithms based on ADDQoL for T2DM populations have been developed. This study represents the first effort to map ADDQoL to EQ-5D-5L and SF-6Dv2 utility values among Chinese T2DM population. The findings indicate that the two-part regression model using ADDQoL's AWI scores, OI1 scores, and their squared terms as predictors demonstrated the most accurate predictions for EQ-5D-5L and SF-6Dv2 utility values.

As no existing mapping algorithms link ADDQoL to the EQ-5D-5L or SF-6Dv2, direct comparisons with prior studies were not feasible. Nonetheless, this study demonstrated modest associations between ADDQoL's OI1, AWI scores, and both EQ-5D-5L and SF-6Dv2 utility values. Pearson correlation coefficients indicated moderate to weak associations between ADDQoL's OI1 and AWI and the dimensions and utility values of the EQ-5D-5L and SF-6Dv2. These findings align logically and are consistent with previous studies, in a past study using the Chinese version of ADDQoL, the AWI scores obtained also showed a weak correlation with these two GPBMs (Spearman's rank correlation coefficient = 0.164 for EQ-5D-3 L, and 0.281 for the short form-36) [67, 68]. The moderate to weak correlations may stem from limited structural overlap between ADDQoL and the target measures, confined primarily to psychological health

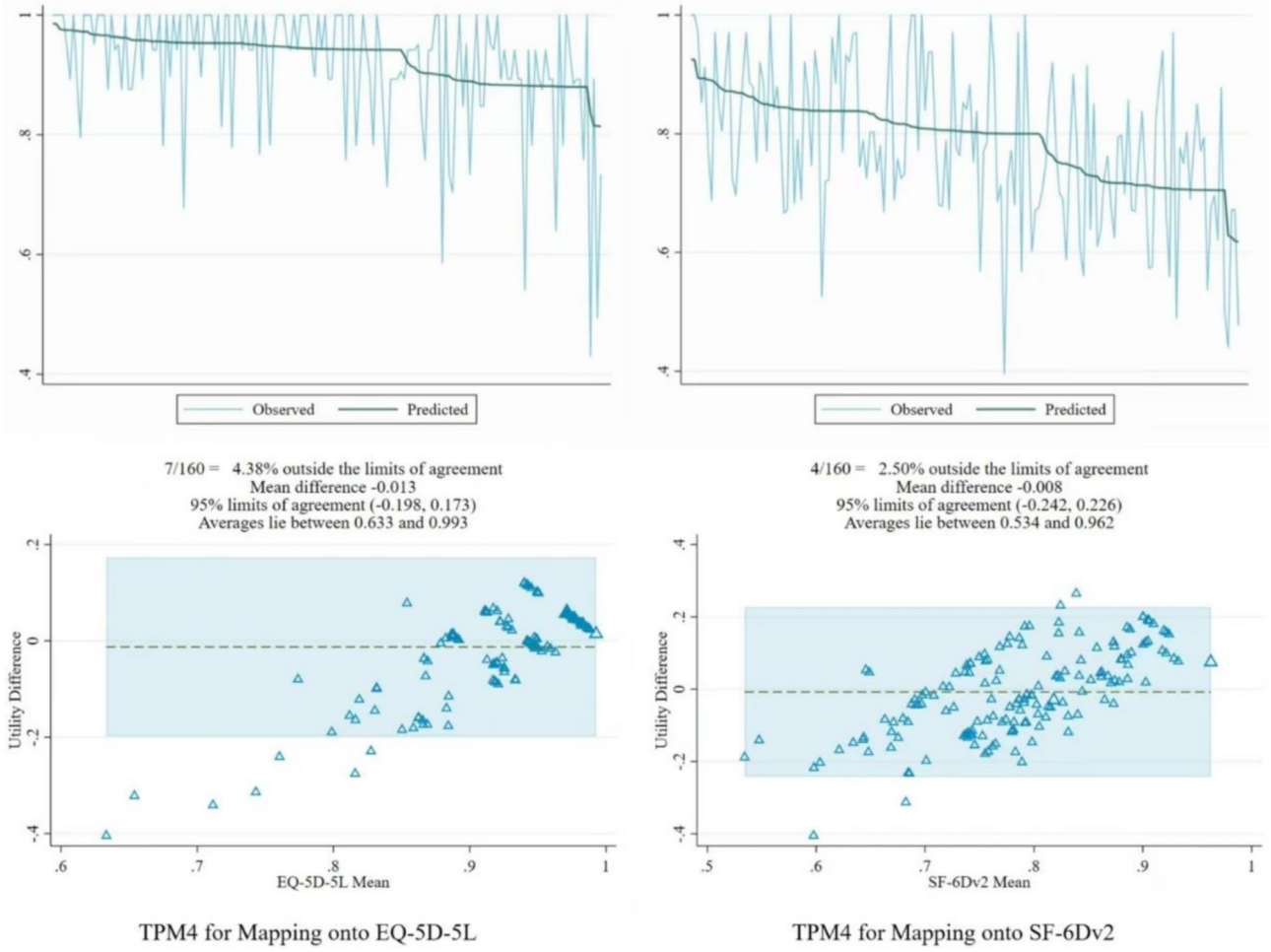

Fig. 2 The optimal model performance is mapped onto the utility values of EQ-5D-5L and SF-6Dv2 in validation group ($N = 160$)

dimensions, failing to capture broader HRQoL aspects [22, 26, 67]. This highlights the limitations of GPBMs in representing the nuanced HRQoL of T2DM patients. Consequently, mapped utility values derived from ADDQoL may better reflect the specific health states of T2DM patients in the absence of GPBM data. Notably, OI1 exhibited the strongest correlations with EQ-5D-5L ($r = 0.375$) and SF-6Dv2 ($r = 0.433$) utility values, underscoring its sensitivity and relevance to mapping studies. This robust association likely arises from OI1's focus on "how you feel about your life right now," aligning closely with the conceptual framework of general quality of life (QoL) and health utility values [42].

This study systematically evaluated different predictor sets, confirming previous findings that incorporating polynomial terms enhances model fit [69, 70]. The optimal predictor set identified in this research included ADDQoL's OI1, AWI scores, and their squared terms, forming the basis of the best-performing polynomial model. Model selection adopted an AR across all three indicators to avoid bias and ensure that the chosen model demonstrated optimal accuracy, consistency, and stability [64], which relied on three key metrics: mean absolute error (MAE) and root mean square error (RMSE) directly reflect predictive accuracy, and the intraclass correlation coefficient (ICC), which assesses consistency between predicted and observed values, reflecting stability and reproducibility. Among six regression methods tested in this study, the two-part model TPM4 consistently exhibited superior performance in predicting health utility values for EQ-5D-5L and SF-6Dv2 in both development and validation groups. This result aligns with Yi Jing Tan et al.'s findings, where TPM was also identified as the best regression approach for mapping the Assessment of Quality of Life-6 Dimensions (AQoL-6D) to EQ-5D-5L in a Malaysian population [71]. Furthermore, when compared to a previously developed mapping function for EQ-5D-3 L in a Hungarian T2DM patients (RMSE = 0.174), the optimal function in this study demonstrated significantly higher accuracy (EQ-5D-5L: RMSE = 0.079; SF-6Dv2: RMSE = 0.117) and a more reasonable utility values prediction range [20]. Additionally,

Table 4 Model performance of six regression methods for mapping the ADDQoL to the SF-6Dv2 utility scores

Mapping Models	Development group($n=640$)						Validation group ($n=160$)					
	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR
OLS1	0.793 (0.014)	0.754,0.826	0.106	0.132	0.045	27	0.794 (0.014)	0.765,0.826	0.108	0.133	0.043	28
OLS2	0.793 (0.014)	0.761,0.811	0.106	0.132	0.047	27	0.793 (0.015)	0.761,0.811	0.108	0.133	0.035	30
OLS3	0.793 (0.059)	0.570,0.903	0.095	0.119	0.501	12	**0.792 (0.062)**	**0.627,0.904**	**0.097**	**0.120**	**0.505**	2
OLS4	**0.793 (0.063)**	**0.580,0.924**	**0.094**	**0.117**	**0.535**	4	0.791 (0.064)	0.613,0.930	0.100	0.120	0.514	7
OLS5	0.793 (0.060)	0.585,0.854	0.095	0.119	0.502	18	0.789 (0.062)	0.623,0.854	0.100	0.121	0.493	18
OLS6	0.793 (0.059)	0.573,0.888	0.095	0.119	0.496	21	0.791 (0.061)	0.629,0.888	0.097	0.120	0.495	5
Tobit1	0.798 (0.014)	0.760,0.829	0.106	0.132	0.042	30	0.798 (0.014)	0.771,0.829	0.109	0.133	0.035	31
Tobit2	0.798 (0.015)	0.764,0.816	0.106	0.132	0.044	31	0.798 (0.015)	0.764,0.816	0.109	0.133	0.029	33
Tobit3	0.797 (0.062)	0.564,0.910	0.095	0.119	0.506	11	**0.796 (0.065)**	**0.625,0.911**	**0.097**	**0.120**	**0.513**	4
Tobit4	**0.797 (0.066)**	**0.576,0.935**	**0.094**	**0.117**	**0.548**	4	0.795 (0.067)	0.611,0.942	0.100	0.120	0.522	9
Tobit5	0.797 (0.062)	0.579,0.862	0.095	0.119	0.513	13	0.794 (0.065)	0.620,0.862	0.099	0.121	0.502	19
Tobit6	0.797 (0.062)	0.566,0.897	0.095	0.119	0.508	13	0.795 (0.064)	0.625,0.897	0.097	0.121	0.504	10
CLAD1	0.808 (0.017)	0.759,0.848	0.106	0.133	0.043	36	0.809 (0.018)	0.773,0.848	0.111	0.135	0.022	35
CLAD2	0.793 (0.009)	0.772,0.805	0.106	0.132	0.032	32	0.793 (0.009)	0.772,0.805	0.108	0.133	0.023	29
CLAD3	0.781 (0.053)	0.586,0.898	0.096	0.120	0.455	24	0.781 (0.056)	0.621,0.901	0.099	0.120	0.469	20
CLAD4	**0.789 (0.061)**	**0.584,0.910**	**0.094**	**0.117**	**0.523**	6	0.786 (0.062)	0.607,0.915	0.101	0.120	0.502	17
CLAD5	0.805 (0.060)	0.592,0.874	0.096	0.119	0.496	23	0.802 (0.062)	0.644,0.874	0.099	0.122	0.487	23
CLAD6	0.798 (0.069)	0.540,0.912	0.095	0.119	0.532	16	**0.796 (0.071)**	**0.603,0.912**	**0.097**	**0.121**	**0.526**	5
GLM1	0.803 (0.011)	0.770,0.829	0.106	0.132	0.032	34	0.803 (0.012)	0.780,0.829	0.110	0.134	0.018	34
GLM2	0.803 (0.013)	0.773,0.819	0.106	0.132	0.037	35	0.803 (0.013)	0.773,0.819	0.110	0.134	0.014	35
GLM3	0.800 (0.064)	0.530,0.914	0.095	0.119	0.520	7	**0.799 (0.067)**	**0.607,0.915**	**0.098**	**0.121**	**0.517**	7
GLM4	**0.800 (0.064)**	**0.558,0.929**	**0.094**	**0.117**	**0.541**	3	0.797 (0.066)	0.604,0.934	0.100	0.121	0.514	16
GLM5	0.800 (0.060)	0.573,0.861	0.095	0.119	0.502	16	0.796 (0.062)	0.621,0.861	0.100	0.121	0.489	24
GLM6	0.800 (0.062)	0.539,0.896	0.095	0.119	0.510	8	0.798 (0.064)	0.615,0.896	0.097	0.121	0.503	15
PTM1	0.793 (0.015)	0.756,0.829	0.106	0.132	0.050	25	0.794 (0.015)	0.766,0.829	0.108	0.133	0.047	27
PTM2	0.793 (0.014)	0.761,0.811	0.106	0.132	0.047	29	0.793 (0.015)	0.761,0.811	0.108	0.133	0.034	31
PTM3	0.793 (0.059)	0.576,0.901	0.095	0.119	0.500	9	0.792 (0.062)	0.625,0.903	0.097	0.120	0.504	2
PTM4	**0.793 (0.063)**	**0.584,0.919**	**0.094**	**0.117**	**0.537**	**1**	**0.790 (0.064)**	**0.618,0.925**	**0.099**	**0.120**	**0.517**	**1**
PTM5	0.793 (0.060)	0.596,0.854	0.095	0.119	0.501	20	0.789 (0.062)	0.628,0.854	0.100	0.121	0.492	21
PTM6	0.793 (0.059)	0.582,0.888	0.095	0.119	0.493	22	0.791 (0.061)	0.633,0.888	0.097	0.120	0.492	10
BM1	0.788 (0.015)	0.743,0.822	0.106	0.132	0.044	32	0.789 (0.015)	0.757,0.822	0.107	0.132	0.050	26
BM2	0.788 (0.016)	0.745,0.819	0.106	0.132	0.053	26	0.788 (0.016)	0.755,0.808	0.107	0.132	0.050	25
BM3	0.788 (0.054)	0.537,0.878	0.095	0.118	0.487	9	0.786 (0.057)	0.607,0.879	0.098	0.119	0.488	13
BM4	**0.788 (0.061)**	**0.541,0.887**	**0.094**	**0.117**	**0.530**	**2**	0.785 (0.062)	0.585,0.890	0.100	0.120	0.506	14

Table 4 (continued)

Mapping Models	Development group($n=640$)						Validation group ($n=160$)					
	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR	Mean (SD)	Min, Max	MAE	RMSE	ICC	AR
BM5	0.788 (0.058)	0.552,0.843	0.096	0.119	0.498	19	0.784 (0.060)	0.599,0.843	0.100	0.120	0.489	21
BM6	0.788 (0.055)	0.531,0.872	0.095	0.118	0.489	13	**0.785 (0.058)**	**0.599,0.865**	**0.098**	**0.119**	**0.490**	**12**

SD standard deviation, Min minimum predicted value, Max. maximum predicted value, AIC Akaike information criteria, BIC Bayesin formation criteria, MAE mean absolute error, RMSE root mean square error, ICC interclass correlation, AR Average rank: Converted from the average ranking score coefficient

1/2/3/4/5/6: Regression model using predictor set1/2/3/4/5/6

Predictor set in bold indicated the best results among the regression method

existing evidence supports that EQ-5D-5L and SF-6Dv2, as target measurements, exhibit superior psychometric properties over EQ-5D-3 L, better capturing patient health utility values [72].

Research suggests that more flexible modeling approaches, such as Beta regression mixture (BM) models, may effectively address heteroscedasticity in the mapping process and yield more precise predictions across the full utility values range of 0 to 1 [61, 62, 73]. As shown in Appendix Fig. 2, BM4, based on the Beta mixture regression method, marginally outperformed TPM4 in the ADDQoL-to-SF-6Dv2 development group for both MAE (BM4: 0.0937; TPM4: 0.0938) and RMSE (BM4: 0.11680; TPM4: 0.11683), demonstrating notably superior fit for lower utility values. The potential advantages of BM in predictive accuracy warrant attention. However, an examination of Bland–Altman plots for mapping to EQ-5D-5L (Fig. 2) reveals BM models' relatively modest performance in consistency between predicted and observed values compared to TPM. Furthermore, inconsistencies in BM results between the development and validation groups for SF-6Dv2 (Table 4) highlight concerns about the model's stability. Notably, OLS also demonstrated competitive performance, ranking second only to TPM in developing models for EQ-5D-5L utility values (Table 3). This suggests that simple OLS-based mapping models can achieve robust results, as some prior studies have similarly opted for OLS when differences in model performance were negligible [74]. Nonetheless, the lack of standardized criteria for acceptable performance disparities underscores the rationale for selecting TPM4 as the optimal mapping function in this study, given its best performance across metrics. Importantly, the variability in model outcomes is likely more attributable to the development sample than to inherent limitations of these regression approaches, highlighting the need for future research to further explore comparative performance across these methods.

The Bland-Altman analysis revealed that the two optimal models mapping ADDQoL to EQ-5D-5L and SF-6Dv2 utility values slightly underestimated the mean utility values (Mean difference for EQ-5D-5L = -0.013; SF-6Dv2 = -0.008). Nevertheless, the majority of data points fell within the 95% limits of agreement (Outside the limitation: EQ-5D-5L = 4.38%; SF-6Dv2 = 2.50%), indicating good overall consistency between predicted and observed values, which also showed the optimal SF-6Dv2 model had better agreement stability. However, such comparisons need cautious interpretation, it is worth noting that the selection of one of the two mapping algorithms developed in this study typically depends on data availability and the overall study design. For instance, in economic evaluations, the choice of mapping

algorithm should ideally be as consistent as possible with the elicitation methods employed for other utility data.

This study demonstrated that the range of mapped utility values differed between the two measures (0.812–0.984 for the EQ-5D-5L, and 0.584–0.919 for the SF-6Dv2), and the utility values derived from the EQ-5D-5L were systematically higher than those from the SF-6Dv2. Consistent with previous research [75], the mapped EQ-5D-5L utilities in diabetic patients exhibited smaller variations across health states compared to utilities measured directly, potentially affecting the results of economic evaluations. Additionally, due to limited samples in poor health conditions, both the EQ-5D-5L and the SF-6Dv2 systematically overestimate mapped utility values for poor health states, which may reduce the observed QALY gains from health improvements, potentially leading to distorted cost-effectiveness outcomes [76]. Researchers may need to conduct sensitivity analyses to validate the robustness of results generated by these mapping algorithms.

This study has several limitations. First, while the development and validation datasets were entirely independent during model exploration, both were derived from the same population sample, underscoring the lack of external validation. Second, to mitigate potential selection bias due to the overrepresentation of younger, technology-proficient patients, we implemented face-to-face interviews. Specifically, trained interviewers assisted elderly or patients with low digital literacy in completing questionnaires through standardized clarification protocols. However, the partial reliance on online devices in data collection may still have inadvertently excluded populations with limited digital access, thereby skewing the sample toward younger, technologically adept individuals. Third, the study's generalizability is limited by its relatively small sample size and cross-sectional design. Future research should prioritize longitudinal studies with larger and more diverse populations to further validate and refine the proposed mapping algorithms, thereby enhancing their applicability in real-world settings.

Conclusion

This study provides two mapping algorithms, derived from the TPM model combined with the responses of the ADDQoL, both of which exhibit acceptable goodness of fit and precision. These algorithms enable the prediction of EQ-5D-5L and SF-6Dv2 utility values from the ADDQoL among T2DM patients in China. The utilization of these mapping functions not only complements patient data in clinical practice and research but also offers an empirical application for informing economic evaluations related to the T2DM interventions.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12955-025-02371-1.

Supplementary Material 1

Acknowledgements

This study was funded by the National Natural Science Foundation of China (grant No. 72404205) and the Natural Science Foundation of Tianjin, China (grant No. 23JCQNJC01650). We would like to thank all the interviewers and respondents for taking part in this study.

Author contributions

Concept and design: TH and SX. Acquisition of data: TH, XL, YH, and SX. Analysis and interpretation of data: HF, TH, LX, CL, YH, and SX. Drafting of the manuscript: HF and SX. Statistical analysis: HF, TH, YH, and SX. Obtaining funding: SX. Supervision: SX. All authors commented on previous versions of the manuscript and approved the final manuscript.

Funding

This study was funded by the National Natural Science Foundation of China (grant No. 72404205) and the Natural Science Foundation of Tianjin, China (grant No. 23JCQNJC01650).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Role of the funder

The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Conflicts of interest/competing interests

No conflicts of interest were reported by the authors.

Consent to participate

Informed consent was obtained from all individual participants included in the study. Participants were informed about their freedom of refusal. Anonymity and confidentiality were maintained throughout the research process.

Ethical approval

This study was approved by the Academic Ethics Committee at Tianjin University (No. TJUE-2023-206) and was conducted in accordance with the Declaration of Helsinki.

Author details

[1]School of Pharmaceutical Science and Technology, Tianjin University, Tianjin, China
[2]School of Biomedical Engineering, McMaster University, Hamilton, ON, Canada
[3]Center for Social Science Survey and Data, Tianjin University, Tianjin, China

Received: 11 January 2025 / Accepted: 7 April 2025
Published online: 30 April 2025

References

1. Kudabayeva K, Tleumagambetova B, Bazargaliyev Y, Kosmuratova R, Zhylkybekova A. Global Trends in LADA Type Diabetes Research: A Bibliometric Analysis of Publications from Web of Science and Scopus, 1994–2024. J Diabetes Res. 2024;2024:4960075.

2. Worldwide trends in. Diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. Lancet. 2024;404(10467):2077–93.
3. IDF Diabetes Atlas. 10th edn [https://www.diabetesatlas.org]
4. Hong SN, Mak IL, Chin WY, Yu EYT, Tse ETY, Chen JY, Wong CKH, Chao DVK, Tsui WWS, Lam CLK, et al. Age-specific associations between the number of co-morbidities, all-cause mortality and public direct medical costs in patients with type 2 diabetes: A retrospective cohort study. Diabetes Obes Metab. 2023;25(2):454–67.
5. He X, Zhang Y, Zhou Y, Dong C, Wu J. Direct medical costs of incident complications in patients newly diagnosed with type 2 diabetes in China. Diabetes Ther. 2021;12(1):275–88.
6. Liu J, Liu M, Chai Z, Li C, Wang Y, Shen M, Zhuang G, Zhang L. Projected rapid growth in diabetes disease burden and economic burden in China: a spatio-temporal study from 2020 to 2030. Lancet Reg Health West Pac. 2023;33:100700.
7. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, Caulley L, Chaiyakunapruk N, Greenberg D, Loder E, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. Value Health. 2022;25(1):3–9.
8. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. Eur J Health Econ. 2018;19(4):557–70.
9. Karimi M. J Brazier 2016 Health, Health-Related quality of life, and quality of life: what is the difference?? PharmacoEconomics 34 7 645–9.
10. Kangwanrattanakul K, Krägeloh CU. EQ-5D-3L and EQ-5D-5L population norms for Thailand. BMC Public Health. 2024;24(1):1108.
11. Mukuria C, Rowen D, Mulhern B, McDool E, Kharroubi S, Bjorner JB, Brazier JE. The short form 6 dimensions (SF-6D): development and evolution. Appl Health Econ Health Policy. 2025;23(1):19–33.
12. Xie S, Wu J, Chen G. Comparative performance and mapping algorithms between EQ-5D-5L and SF-6Dv2 among the Chinese general population. Eur J Health Econ. 2024;25(1):7–19.
13. Liu GG, Hu S, Wu J. China Guidelines for Pharmacoeconomic Evaluations (2020). Beijing, China; 2020.
14. Xie S, Wang D, Wu J, Liu C, Jiang W. Comparison of the measurement properties of SF-6Dv2 and EQ-5D-5L in a Chinese population health survey. Health Qual Life Outcomes. 2022;20(1):96.
15. Mulhern B, Meadows K. The construct validity and responsiveness of the EQ-5D, SF-6D and diabetes health Profile-18 in type 2 diabetes. Health Qual Life Outcomes. 2014;12:42.
16. Oliveira Gonçalves AS, Werdin S, Kurth T, Panteli D. Mapping studies to estimate Health-State utilities from Nonpreference-Based outcome measures: A systematic review on how repeated measurements are taken into account. Value Health. 2023;26(4):589–97.
17. Mehdipour A, O'Hoski S, Beauchamp MK, Wald J, Kuspinar A. Content validity of preference-based measures for economic evaluation in chronic obstructive pulmonary disease. Health Qual Life Outcomes. 2021;19(1):101.
18. Peters N, Dal Bello-Haas V, Packham T, Chum M, O'Connell C, Johnston WS, MacDermid JC, Turnbull J, Van Damme J, Kuspinar A. Do generic Preference-Based measures accurately capture areas of Health-Related quality of life important to individuals with amyotrophic lateral sclerosis: A content validation study. Patient Relat Outcome Meas. 2021;12:191–203.
19. Longworth L, Rowen D. NICE decision support unit technical support documents. In: *NICE DSU technical support document 10: the use of mapping methods to estimate health state utility values*. edn. London: National Institute for Health and Care Excellence (NICE).
20. Vokó Z, Németh R, Nagyjánosi L, Jermendy G, Winkler G, Hídvégi T, Kalotai Z, Kaló Z. Mapping the Nottingham health profile onto the Preference-Based EuroQol-5D instrument for patients with diabetes. Value Health Reg Issues. 2014;4:31–6.
21. Wu Y, Cheng L, Guo H. Mapping diabetes specific quality of life scale to EQ-5D utility scale. Chin J Health Stat. 2022;39(4):499.
22. Ostini R, Dower J, Donald M. The audit of Diabetes-Dependent quality of life 19 (ADDQoL): feasibility, reliability and validity in a population-based sample of Australian adults. Qual Life Res. 2012;21(8):1471–7.
23. McMillan CV, Honeyford RJ, Datta J, Madge NJ, Bradley C. The development of a new measure of quality of life for young people with diabetes mellitus: the ADDQoL-Teen. Health Qual Life Outcomes. 2004;2:61.
24. Lemon SC, Rosal MC, Welch G. Measuring quality of life in low-income, Spanish-speaking Puerto Ricans with type 2 diabetes residing in the Mainland U.S. Qual Life Res. 2011;20(9):1507–11.
25. Bak E, Marcisz C, Nowak-Kapusta Z, Dobrzyn-Matusiak D, Marcisz E, Krzeminska S. Psychometric properties of the audit of Diabetes-Dependent quality of life (ADDQoL) in a population-based sample of Polish adults with type 1 and 2 diabetes. Health Qual Life Outcomes. 2018;16(1):53.
26. Iversen MM, Espehaug B, Rokne B, Haugstvedt A, Graue M. Psychometric properties of the Norwegian version of the audit of Diabetes-Dependent quality of life. Qual Life Res. 2013;22(10):2809–12.
27. Visockienė Ž, Narkauskaitė-Nedzinskienė L, Puronaitė R, Mikaliūkštienė A. Validation of the LITHUANIAN version of the 19-item audit of diabetes dependent quality of life (ADDQOL - LT) questionnaire in patients with diabetes. Health Qual Life Outcomes. 2018;16(1):206.
28. Zhang XH, Tan K, Tan HH, Thumboo J, Li SC. Are english and Chinese versions of the audit of Diabetes-Dependent quality of life equivalent?? An exploratory study based on the universalist approach. Value Health Reg Issues. 2012;1(1):75–81.
29. Wang HF, Bradley C, Chang TJ, Chuang LM, Yeh MC. Assessing the impact of diabetes on quality of life: validation of the Chinese version of the 19-item audit of Diabetes-Dependent quality of life for Taiwan. Int J Qual Health Care. 2017;29(3):335–42.
30. Chang YW, Shen FC, Chen CY. Investigation of treatment satisfaction and health-related quality of life after add-on to metformin-based therapy in patients with type 2 diabetes. Front Public Health. 2023;11:1152284.
31. Wee HL, Tan CE, Goh SY, Li SC. Usefulness of the audit of Diabetes-Dependent Quality-of-Life (ADDQoL) questionnaire in patients with diabetes in a multiethnic Asian country. PharmacoEconomics. 2006;24(7):673–82.
32. Fung CS, Wan EY, Yu CL, Wong CK. Validity and reliability of the 19-item audit of Diabetes-Dependent quality of life (ADDQoL-19) questionnaire in Chinese patients with type 2 diabetes mellitus in primary care. Qual Life Res. 2016;25(9):2373–8.
33. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, Botteman M, Busschbach J. Mapping to estimate Health-State utility from Non-Preference-Based outcome measures: an ISPOR good practices for outcomes research task force report. Value Health. 2017;20(1):18–27.
34. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, Gray A. Preferred reporting items for studies mapping onto preference-based outcome measures: the MAPS statement. Health Qual Life Outcomes. 2015;13:106.
35. Li Y, Teng D, Shi X, Qin G, Qin Y, Quan H, Shi B, Sun H, Ba J, Chen B, et al. Prevalence of diabetes recorded in Mainland China using 2018 diagnostic criteria from the American diabetes association: National cross sectional study. BMJ. 2020;369:m997.
36. Chuang LH, Zhang H, Hong T, Xie S. Evaluating the preferences and Willingness-to-Pay for oral antidiabetic drugs among patients with type 2 diabetes mellitus in China: A discrete choice experiment. Patient. 2024;17(5):565–74.
37. Weng G, Hong Y, Luo N, Mukuria C, Jiang J, Yang Z, Li S. Comparing EQ-5D-3L and EQ-5D-5L in measuring the HRQoL burden of 4 health conditions in China. Eur J Health Econ. 2023;24(2):197–207.
38. Zhu J, Yan XX, Liu CC, Wang H, Wang L, Cao SM, Liao XZ, Xi YF, Ji Y, Lei L, et al. Comparing EQ-5D-3L and EQ-5D-5L performance in common cancers: suggestions for instrument choosing. Qual Life Res. 2021;30(3):841–54.
39. Zhang AX, Mao ZX, Wang P. [Comparison between EuroQol Five-Dimensional questionnaire and Short-Form 6-Dimension health survey for the populations in China:A scoping review]. Zhongguo Yi Xue Ke Xue Yuan Xue Bao. 2023;45(4):655–65.
40. Wu J, Xie S, He X, Chen G, Brazier JE. The simplified Chinese version of SF-6Dv2: translation, cross-cultural adaptation and preliminary psychometric testing. Qual Life Res. 2020;29(5):1385–91.
41. Wu J, Xie S, He X, Chen G, Bai G, Feng D, Hu M, Jiang J, Wang X, Wu H, et al. Valuation of SF-6Dv2 health States in China using time Trade-off and Discrete-Choice experiment with a duration dimension. PharmacoEconomics. 2021;39(5):521–35.
42. Kong D, Ding Y, Zuo X, Su W, Xiu L, Lin M, Rao S, Yu S. Adaptation of the audit of Diabetes-Dependent quality of life questionnaire to people with diabetes in China. Diabetes Res Clin Pract. 2011;94(1):45–52.
43. Luo N, Li M, Liu GG, Lloyd A, de Charro F, Herdman M. Developing the Chinese version of the new 5-level EQ-5D descriptive system: the response scaling approach. Qual Life Res. 2013;22(4):885–90.
44. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value Health. 2017;20(4):662–9.

45. Brazier JE, Mulhern BJ, Bjorner JB, Gandek B, Rowen D, Alonso J, Vilagut G, Ware JE. Developing a new version of the SF-6D health state classification system from the SF-36v2: SF-6Dv2. Med Care. 2020;58(6):557–65.
46. Huang D, Zeng D, Tang Y, Jiang L, Yang Q. Mapping the EORTC QLQ-C30 and QLQ H&N35 to the EQ-5D-5L and SF-6D for papillary thyroid carcinoma. Qual Life Res. 2024;33(2):491–505.
47. Bilbao A, Martín-Fernández J, García-Pérez L, Arenaza JC, Ariza-Cardiel G, Ramallo-Fariña Y, Ansola L. Mapping WOMAC onto the EQ-5D-5L utility index in patients with hip or knee osteoarthritis. Value Health. 2020;23(3):379–87.
48. Wang J, Hong T, Fang H, Luo C, He X, Xie S. Comparison of the measurement properties and consistency between the EQ-5D-3L and EQ-5D-Y-3L in adolescents aged 15–17 in China. Health Qual Life Outcomes. 2024;22(1):59.
49. Sakoglu N, Donmez T. Does prior lower abdominal surgery prevent laparoscopic hernia repair (Totally extraperitoneal or transabdominal Preperitoneal)? A Prospective Observational Study. J Laparoendosc Adv Surg Tech A 2024.
50. McDool E, Mukuria C, Brazier J. A comparison of the SF-6Dv2 and SF-6D UK utility values in a mixed patient and healthy population. PharmacoEconomics. 2021;39(8):929–40.
51. Cohen J. A power primer. Psychol Bull. 1992;112(1):155–9.
52. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health. 2020;8(1):e000262.
53. Feng Y, Devlin NJ, Shah KK, Mulhern B, van Hout B. New methods for modelling EQ-5D-5L value sets: an application to english data. Health Econ. 2018;27(1):23–38.
54. Lamu AN, Olsen JA. Testing alternative regression models to predict utilities: mapping the QLQ-C30 onto the EQ-5D-5L and the SF-6D. Qual Life Res. 2018;27(11):2823–39.
55. Mukuria C, Rowen D, Harnan S, Rawdin A, Wong R, Ara R, Brazier J. An updated systematic review of studies mapping (or Cross-Walking) measures of Health-Related quality of life to generic Preference-Based measures to generate utility values. Appl Health Econ Health Policy. 2019;17(3):295–313.
56. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. Qual Life Res. 2000;9(8):901–10.
57. Clarke P, Gray A, Holman R. Estimating utility values for health States of type 2 diabetic patients using the EQ-5D (UKPDS 62). Med Decis Mak. 2002;22(4):340–9.
58. Pullenayegum EM, Tarride JE, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? Value Health. 2010;13(4):487–94.
59. Basu A. Extended generalized linear models: simultaneous Estimation of flexible link and variance functions. Stata J. 2005;5(4):501–16.
60. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. Eur J Health Econ. 2015;16(2):141–51.
61. Gray LA, Wailoo AJ, Hernandez Alava M. Mapping the FACT-B instrument to EQ-5D-3L in patients with breast cancer using adjusted limited dependent variable mixture models versus response mapping. Value Health. 2018;21(12):1399–405.
62. Gray LA, Hernández Alava M, Wailoo AJ. Development of methods for the mapping of utilities using mixture models: mapping the AQLQ-S to the EQ-5D-5L and the HUI3 in patients with asthma. Value Health. 2018;21(6):748–57.
63. Li C, Dou L, Fu Q, Li S. Mapping the Seattle angina questionnaire to EQ-5D-5L in patients with coronary heart disease. Health Qual Life Outcomes. 2023;21(1):64.
64. Abdin E, Chong SA, Seow E, Tan KB, Subramaniam M. Mapping the PHQ-8 to EQ-5D, HUI3 and SF6D in patients with depression. BMC Psychiatry. 2021;21(1):451.
65. Ho KM. Scatter plot and correlation coefficient. Anaesth Intensive Care. 2012;40(4):730–1.
66. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb). 2015;25(2):141–51.
67. Soon SS, Goh SY, Bee YM, Poon JL, Li SC, Thumboo J, Wee HL. Audit of Diabetes-Dependent quality of life (ADDQoL) [Chinese Version for Singapore] questionnaire: reliability and validity among Singaporeans with type 2 diabetes mellitus. Appl Health Econ Health Policy. 2010;8(4):239–49.
68. Grigorescu ED, Lăcătușu CM, Crețu I, Floria M, Onofriescu A, Ceasovschih A, Mihai BM, Șorodoc L. Self-Reported satisfaction to treatment, quality of life and general health of type 2 diabetes patients with inadequate glycemic control from North-Eastern Romania. Int J Environ Res Public Health 2021, 18(6).
69. Davison NJ, Thompson AJ, Turner AJ, Longworth L, McElhone K, Griffiths CEM, Payne K. Generating EQ-5D-3L utility scores from the dermatology life quality index: A mapping study in patients with psoriasis. Value Health. 2018;21(8):1010–8.
70. Aghdaee M, Gu Y, Sinha K, Parkinson B, Sharma R, Cutler H. Mapping the Patient-Reported outcomes measurement information system (PROMIS-29) to EQ-5D-5L. PharmacoEconomics. 2023;41(2):187–98.
71. Tan YJ, Ong SC. Direct and indirect mapping of assessment of quality of life – 6 dimensions (AQoL-6D) onto EQ-5D-5L utilities using data from a multicenter, Cross-Sectional study of Malaysians with chronic heart failure. Value Health. 2024;27(12):1762–70.
72. Xu RH, Luo N, Dong D. Measurement properties of the EQ-5D-3L, EQ-5D-5L, and SF-6Dv2 in patients with late-onset Pompe disease. Eur J Health Econ. 2024;25(9):1505–15.
73. Yu L, Yang H, Lu L, Fang Y, Zhang X, Li S, Li C. Developing mapping algorithms to predict EQ-5D health utility values from bath ankylosing spondylitis disease activity index and bath ankylosing spondylitis functional index among patients with ankylosing spondylitis. Health Qual Life Outcomes. 2024;22(1):61.
74. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ. 2010;11(2):215–25.
75. Yang F, Devlin N, Luo N. Cost-Utility analysis using EQ-5D-5L data: does how the utilities are derived matter?? Value Health. 2019;22(1):45–9.
76. Wailoo A, Alava MH, Pudney S, Barton G, O'Dwyer J, Gomes M, Irvine L, Meads D, Sadique Z. An international comparison of EQ-5D-5L and EQ-5D-3L for use in Cost-Effectiveness analysis. Value Health. 2021;24(4):568–74.

Publisher's note