Yu *et al. Health and Quality of Life Outcomes* (2024) 22:61
https://doi.org/10.1186/s12955-024-02276-5

Health and Quality of Life Outcomes

RESEARCH

Open Access

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

Lingjia Yu[1†], Huizhi Yang[2†], Liyong Lu[3,4,5], Yingying Fang[1], Xianyu Zhang[1], Shunping Li[3,4,5] and Chaofan Li[3,4,5*]

Abstract

Background Preference-based measures of health-related quality of life (HRQoL), such as the EQ-5D or the SF-6D, are essential for health economic evaluation. However, they are rarely included in clinical trials of ankylosing spondylitis (AS). This study aims to develop mapping algorithms to predict EQ-5D-3L and EQ-5D-5L health utility scores from the Bath Ankylosing Disease Activity Index (BASDAI) and the Bath Ankylosing Spondylitis Functional Index (BASFI).

Methods Patients with AS were recruited from the largest tertiary hospital in Shandong province, China, between December 2019 and October 2020. Patients were selected by convenience sampling method according to the following criteria: (1) diagnosed with AS according to the New York criteria; (2) aged 18 years and above; and (3) without mental disorders; (4) able to understand the questionnaires; (5) without serious complications. There were 243 patients who completed the face-to-face questionnaire survey, and 5 cases with missing values in key variables were excluded. Ordinary least squares, censored least absolute deviations, Tobit, adjusted limited dependent variable mixture model and beta-mixture model (BM) in the direct approach and ordered logit and multinomial logit (Mlogit) model in the response approach were used to develop mapping algorithms. Mean absolute error, root mean square error, Spearman's correlation coefficient and concordance correlation coefficient were used to access predictive performance.

Results The 238 patients with AS had a mean age of 35.19 (SD = 9.59) years, and the majority (74.47%) were male. The observed EQ-5D-3L and EQ-5D-5L health utility values were 0.88 (SD = 0.12) and 0.74 (SD = 0.27), respectively. The EQ-5D-5L had higher conceptual overlap with the BASDAI and BASFI than the EQ-5D-3L did. The Mlogit was the

[†]Lingjia Yu and Huizhi Yang contributed equally to this work.

*Correspondence:
Chaofan Li
lcfan@sdu.edu.cn

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

best-performing model for the EQ-5D-3L, and the BM showed better performance in predicting EQ-5D-5L than other direct and indirect mapping models did.

Conclusion This study demonstrates that the EQ-5D-5L, rather than EQ-5D-3L, should be selected as the target outcome measure of HRQoL in patients with AS in China, and the BM mapping algorithm could be used to predict EQ-5D-5L values from BASDAI and BASFI for health economic evaluation.

Keywords EQ-5D, Mapping, BASDAI, BASFI, Ankylosing spondylitis

Background

Ankylosing spondylitis (AS) is a common chronic inflammatory rheumatic disease associated with low back pain, stiffness and restriction of lumbar spine motion [1, 2]. The symptoms of AS can lead to severe functional impairment and a marked decrease in patients' health-related quality of life (HRQoL) [2], resulting in a heavy physical burden for patients and a significant impact on their normal daily life, social communication and work engagement [3, 4]. HRQoL refers to patient-reported health status of functioning and well-being in physical, mental, and social domains of life. In addition, the diagnosis and treatment of AS also impose a substantial economic burden on patients' family, society and the whole county [5]. Although the prevalence of AS in mainland China between 2005 and 2019 (0.22–0.35%) was lower than that in the United States (0.9–1.4% from 2009 to 2010) and that in central Europe (0.3–0.5% from 2000 to 2006), it has shown an increasing trend [1, 2, 6]. The increasing epidemiological trend and the enormous economic burden of AS pose great challenges for the governmental decision-making regarding the inclusion of AS health interventions in public health insurance coverage.

Health economic evaluation is recommended as a critical tool to inform policy decisions on the efficient allocation of limited healthcare resources for the prevention and treatment of AS [7]. The Chinese College of Rheumatology included health economic evaluation in the expert consensus and recommendations for the treatment of AS in 2021 [8]. Cost-utility analysis (CUA) is a widely used health economic evaluation method that compares the cost per quality-adjusted life year (QALY) or the incremental cost-effectiveness ratio of different interventions [9, 10]. In order to calculate QALYs, we need to use generic preference-based measures of HRQoL, such as the EQ-5D or SF-6D, to capture health state utilities [11]. However, most clinical trials of AS therapies did not include EQ-5D or SF-6D questionnaires to measure patients' QALY but rather disease-specific questionnaires to measure QALY [12–14]. To bridge this gap, mapping has been widely used to predict EQ-5D health utility values by estimating the relationship between preference-based and disease-specific questionnaires using various regression methods [15, 16].

In recent years, several studies have developed mapping algorithms in patients with AS in European countries. Ara and Mlcoch developed linear regression models to predict EQ-5D-3L utility values using date from a large multicenter RCT conducted in European countries [17, 18]. Wailoo predicted EQ-5D-3L scores from the Bath Ankylosing Disease Activity Index (BASDAI) or the Bath Ankylosing Spondylitis Functional Index (BASFI) using both direct and indirect mapping approaches [19], and Neilson converted BASDAI and BASFI to EQ-5D-5L values using similar models [12]. More recently, Hernandez Alava predicted EQ-5D-3L values from the Ankylosing Spondylitis Disease Activity Score (ASDAS) using bespoke mixture models and compared their predictive performance with BASDAI [20].

In principle, country-specific algorithms are recommended to predict health utility values for health economic evaluations in different countries [21]. However, there has been little research on mapping algorithms based on data and tariffs in China. Considering the increasing prevalence of AS and the increasing role of health economic evaluation in health policy-making in China [22], a country-specific mapping algorithm should be developed to facilitate the implementation of CUA in Chinese clinical trials. In addition, previous studies on mapping algorithms in AS patients included EQ-5D-3L or EQ-5D-5L questionnaires. However, there were significant unpredictable differences between EQ-5D-3L and EQ-5D-5L in health economic evaluation [23]. Therefore, it is necessary to select an appropriate outcome questionnaire and mapping method to generate unbiased QALYs for CUA.

To fill these research gaps, this study aims to develop mapping algorithms for both EQ-5D-3L and EQ-5D-5L in patients with AS in China and select the best-performing model to support health policy decision-making. The mapping practice and reporting followed the guidelines of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the "Mapping onto Preference-based measures reporting Standards (MAPS) statement" developed by Petrou and Wailoo [15, 24, 25].

Methods

Study sample

The samples were recruited from a tertiary hospital in Shandong province, China, between December 2019 and October 2020. This hospital, which has more than 4000 ward beds, is the largest and best-performing tertiary hospital in Shandong province [26]. In 2022, it delivered 2.68 million ambulatory or outpatient visits, 187.6 thousand hospital visits, and 148 thousand surgeries. The rheumatology department of this hospital employs 20 medical staff members and trains rheumatism specialists from all three provinces of Shandong, Xinjiang and Ningxia. Furthermore, the department is the preferred choice for the diagnosis and treatment of complex rheumatism cases from all 16 cities in Shandong provinces. This study included patients with AS disease who met the following criteria: (1) Diagnosed as AS according to the New York criteria [27]; (2) Aged 18 years and above; and (3) No mental disorders; (4) Able to understand the questionnaires; (5) Without serious complications. The AS patients were selected using a convenience sampling method and given their written informed consent prior to the survey. Trained interviewers conducted face-to-face questionnaire surveys. A total of 243 patients completed the survey during the study period. After deleting five cases with missing values in key variables, the remaining 238 patients were included in the development and validation of the mapping algorithm.

This study obtained ethical approval from Ethics Committee of Scholl of Healthcare Management, Shandong University (ECSHCMSDU20181102).

Outcome measures

The EQ-5D is a generic preference-based questionnaire designed to measure health-related quality of life. It comprises five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). The instrument was initially developed using three levels (no problem, some problem and extreme problem) for each dimension in the 1990s [28] and then adapted to five levels (no problems, slight problems, moderate problems, severe problems, and extreme problems) in 2011 [29]. The EQ-5D-5L has been demonstrated to outperform EQ-5D-3L in aspects of feasibly, ceiling effect, discriminatory power, and convergent validity [30–32]. Both the EQ-5D-3L and the EQ-5D-5L have been translated into Chinese and have demonstrated excellent reliability and validity in measuring and valuing health status [33]. The Chinese value set was employed to estimate health utility values. The values calculated from the EQ-5D-3L and EQ-5D-5L tariffs for China range from −0.149 to 1 [34], and from −0.391 to 1, respectively [35]. The higher values of EQ-5D indicate better health, with 1 representing perfect health, 0 corresponding to death and negative values representing health status worse than death.

Source measures

In accordance with previous studies on AS mapping [12, 17–20], predictive variables included BASDAI, BASFI and age. The BASDAI, developed by Garrett, is a self-administered questionnaire comprising six items attributed to five major symptoms of the AS disease. Each item was recorded on an 11-point numerical rating scale, with 0 representing no problem and 10 representing the most serious problem, in order to assess the severity of the five symptoms. The BASDAI score ranged from 0 to 10, with equal weight being given to each symptom. The BASDAI questionnaire has been translated into Chinese and has been demonstrated to be reliable and valid for the assessment of the status and activity of patients with AS in China [36].

The BASFI is also a self-assessment questionnaire comprising of 10 items. Eight of these items reflect impairment of function, while the remaining two items related to the ability of daily life. Each item is rated on a horizontal visual analogue scale and with a length of 10 cm, and scored from 0 (easy) to 10 (impossible). The BASFI score was calculated as the mean of 10 items, with scores ranging from 0 to 10. Higher scores indicated worse function or status. Lin also translated the BASFI questionnaire into Chinese and verified its good reliability and validity among Chinese patients with AS [36].

Statistical analysis

The data analysis strategy followed the "Good Practice for Mapping Studies" developed by ISPOR [15]. The frequency and percentage were employed to describe the fundamental characteristics of categorical variables. The distribution of continuous variables was described using means, standard deviations, minimum, and maximum. Additionally, a histogram was employed to graphically display the distribution of EQ-5D values.

The normality of the data was tested using the Shapiro-Wilk test. Due to the non-normal distribution of continuous variables, Spearman rank correlation coefficients were employed to assess the conceptual overlap between BASDAI scores, BASFI scores and EQ-5D values. Furthermore, the relationship between each item of BASDAI with BASFI and each EQ-5D dimension score was tested using Spearman rank correlation coefficient.

We employed both direct and indirect approaches to develop the mapping algorithm. In the direct approach, statistical models were selected according to the distributional characteristics of EQ-5D values, including inflation at 1, skewness, multimodality, and gaps in the range of feasible values [15]. The initial approach involved the use of ordinary least squares (OLS) regression, the most common method, to predict EQ-5D values. However, a previous review demonstrated that the predicted values of OLS exhibited systematic bias, including values outside the feasible range, underestimation for patients with good health, and overestimation for patients with severe health conditions [37]. Secondly, in order to address the inflation of EQ-5D values at 1, we explored censored least absolute deviations (CLAD) and Tobit models [16]. Thirdly, the EQ-5D utility score was frequently dramatically skewed, bounded at bottom and top values, and inflated at the upper limit. Previous studies have demonstrated that bespoke regression models, such as the adjusted limited dependent variable mixture model (ALDVMM) [38] and the beta-mixture (BM) model [39], exhibit superior performance to traditional models in predicting EQ-5D scores. The ALDVMM and BM outperformed the basic regression model in several aspects. These included the ability to capture the multimodality of EQ-5D values, to account for gaps between 1 (full health) and the nearest feasible value, and to deal with boundary values [39]. The Stata commands "aldvmm" and "betamix" were employed to fit the ALDVMM and BM models, respectively [40, 41].

In terms of indirect mapping approach, both ordered logit (Ologit) and multinomial logit (Mlogit) methods were employed. The procedures described by Gray [42] and Hernández Alava [43] were followed to estimate five Ologit and Mlogit models for each EQ-5D dimension. These models were used to predict the probability of a given response level in the first step. Subsequently, the EQ-5D values were calculated based on the probabilities of each level of each dimension and the Chinese EQ-5D tariffs.

In the present study, no suitable external dataset could be identified as a validation sample. Consequently, we employed in-sample five-fold cross-validation to validate the mapping algorithm, following the methodology previously described by [44]. Firstly, the primary data were divided into five distinct categories. Secondly, four subsamples were employed as estimation data, while the remaining subsample was retained as validation data in order to evaluate predictive performance. Thirdly, the aforementioned process was repeated five times, with each subsample serving as validation data on one occasion.

In accordance with good mapping best practices [15], the mean absolute error (MAE) and root mean squared error (RMSE) were employed to assess the predictive performance of each model type and specification. Furthermore, Spearman's correlation coefficient and concordance correlation coefficient (CCC) were employed as performance indicators to assess the concordance between observed and predicted EQ-5D values. The CCC was a robust indicator for evaluating the correlation between observed and predicted values from uniform or other non-uniform distributions [45]. The model with the lowest MAE and RMSE, as well as the highest Spearman's correlation coefficient and CCC values, demonstrated superior predictive performance in comparison to other models. The models were ranked according to four performance indicators, and the model with the lowest averaging rank was selected as the best-performing one [46]. Furthermore, the mean, standard deviation, minimum and maximum of observed and predicted EQ-5D values were compared. A scatter plot between observed and predicted EQ-5D values was drawn to identify any potential bias. A variance-covariance matrix was constructed to conduct a probabilistic sensitivity analysis for future economic evaluation.

All data analysis was conducted in Stata 15.0 (Stata Corp LP, College Station, Texas, USA) [47]. The threshold for statistical significance was set at $P<0.05$.

Results

Patient characteristics

The basic characteristics of the 238 observations are reported in Table 1. The mean age of the patients was 35.19 years (SD=9.59), with the majority being male (74.47%). The mean (SD) of the EQ-5D-3L value was 0.88 (0.12), while that of the EQ-5D-5L value was 0.74 (0.27). The maximum of EQ-5D-3L and EQ-5D-5L values was 1, whereas the minimum of EQ-5D-5L value (-0.35) was much smaller than that of EQ-5D-3L (0.17). The mean (SD) score for BASDIA and BADFI was 3.60 (2.38) and 3.19 (2.62), respectively. Fig. 1 shows the distribution of EQ-5D values. Both the EQ-5D-3L and EQ-5D-5L values exhibited a left-skewed distribution, with the majority of values concentrated between 0.7 and 1.0.

Conceptual overlap

Table 2 displayed the Spearman correlation coefficients between BASDAI and BASFI scores and EQ-5D total scores and five dimensions. In general, the BASDAI and BASFI scores were moderately negatively significantly correlated with the EQ-5D-3L total scores (r=-0.4071 and r=-0.4144, respectively, $P<0.001$), while they were strongly negatively significantly correlated with EQ-5D-5L total scores (-0.8056 and -0.8310, respectively, $P<0.001$). The Spearman correlation coefficients between BASDAI and BASFI scores and EQ-5D dimension scores were all significantly positive ($P<0.001$), with values ranging from 0.4750 to 0.7853. In general, there was a high degree of conceptual overlap between the BASDAI and BASFI questionnaires and the EQ-5D questionnaires.

Table 1 Basic characteristics of the study samples ($N=238$)

	N (%)	Range (min, max)
Age, year (Mean ± SD)	35.19 ± 9.59	18, 66
Sex		
Male	182 (74.47)	
Female	56 (23.53)	
Region of residence		
Urban	121 (50.84)	
Rural	117 (49.16)	
Marital status		
Never married	58 (24.37)	
Married	170 (71.43)	
Divorce or widowed	4 (1.68)	
Other	6 (2.52)	
Years of schooling		
< 6	8 (3.36)	
6–9	51 (21.43)	
9–12	61 (25.63)	
> 12	118 (49.58)	
Household income per year, Yuan (RMB)		
< 30,000	57 (23.95)	
30,000–80,000	104 (43.70)	
80,000-150,000	68 (28.57)	
> 150,000	9 (3.78)	
Duration of disease, year		
< 1	25 (10.50)	
1–2	42 (17.65)	
3–5	63 (26.47)	
6–8	47 (19.75)	
> 9	61 (25.63)	
EQ-5D-3L index value	0.88 (0.12)	0.17, 1.00
EQ-5D-5L index value	0.74 (0.27)	-0.35,1.00
BASDAI	3.60 (2.38)	0, 9.20
BASFI	3.19 (2.62)	0, 10.00

EQ-5D-3L model development and predictive performance

Table 3 displays the predictive performance of five direct mapping models and two indirect mapping models for EQ-5D-3L. Among the seven models, the mean of the predictive EQ-5D-3L value of OLS was equal to that of the observed, while the CLAD and Tobit exhibited higher predictive mean values, and the other four models exhibited lower predictive mean values. Notwithstanding, the minimums of predictive values of all the seven models were much higher than those of the observed EQ-5D-3L values (0.170). Moreover, the maximum of predictive values of CLAD and Tobit were beyond the feasible range of EQ-5D-3L, which were higher than 1. (Fig. 2) displays the scatter plot between observed EQ-5D-3L values and predictions captured from seven models. The Mlogit exhibited the lowest MAE value (0.084) and the highest Spearman correlation coefficient (0.440), while the OLS exhibited the lowest RMSE values (0.113) and the Ologit model exhibited the highest CCC index (0.364). Consequently, the Mlogit model demonstrated superior performance in predicting EQ-5D-3L from BASDAI, BASFI and age. The coefficients and their variance-covariance matrix for each mapping model used to predict EQ-5D-3L scores are reported in Supplementary Material 1.

EQ-5D-5L model development and predictive performance

Table 4 reports the predictive performance of seven models for EQ-5D-5L values. The means of predictive EQ-5D-5L values from the OLS (0.7396) and BM (0.7312) models were much closer to the observed EQ-5D-5L value (0.7388) than those from the other five models. The BM and Ologit model exhibited negative predictive values, while the minimum of predictive values from other models were all greater than zero. (Fig. 3) displays the scatter plot between observed EQ-5D-5L values and predictions captured from seven models. With regard to the four predictive performance indicators, the Mlogit model exhibited the lowest MAE values (0.1006), while the BM demonstrated the lowest RMSE (0.1533) and the highest Spearman correlation coefficient (0.8570) and CCC index (0.8130). The average rank of the four predictive performance indicates that the BM is the optimal fit model for predicting EQ-5D-5L from BASDAI, BASFI and age, and the ALDVMM was a sub-optimal model. The coefficients and their variance-covariance matrix of seven regression models to predict EQ-5D-5L value are reported in the Supplementary Material 2.

The best-performing algorithms for predicting EQ-5D-3L and EQ-5D-5L values were provided in Supplementary Material 3 and can be easily implemented in Stata software.

Discussion

In order to facilitate the use of CUA in health resource policy making in AS disease, this study developed both direct and indirect mapping algorithms to predict EQ-5D-3L and EQ-5D-5L utility values from two widely used AS disease-specific questionnaires (BASDAI and BASFI). According to the ranks of five predictive performance indicators, Mlogit was identified as the optimal model for estimating EQ-5D-3L scores, while BM demonstrated the highest performance for estimating EQ-5D-5L values. Given the increasing prevalence trend of AS in China and the lack of preference-based questionnaires in clinical studies, the mapping algorithms provide a reliable and convenient tool for converting BASDAI and BASFI to EQ-5D-3L and EQ-5D-5L health utility values. This could facilitate the application of CUA in the future.

Our algorithms are compared to existing algorithms developed in European countries. Firstly, consistent with previous studies [12, 17, 19], we also included BASDAI, BASFI and age into the regression models to estimate

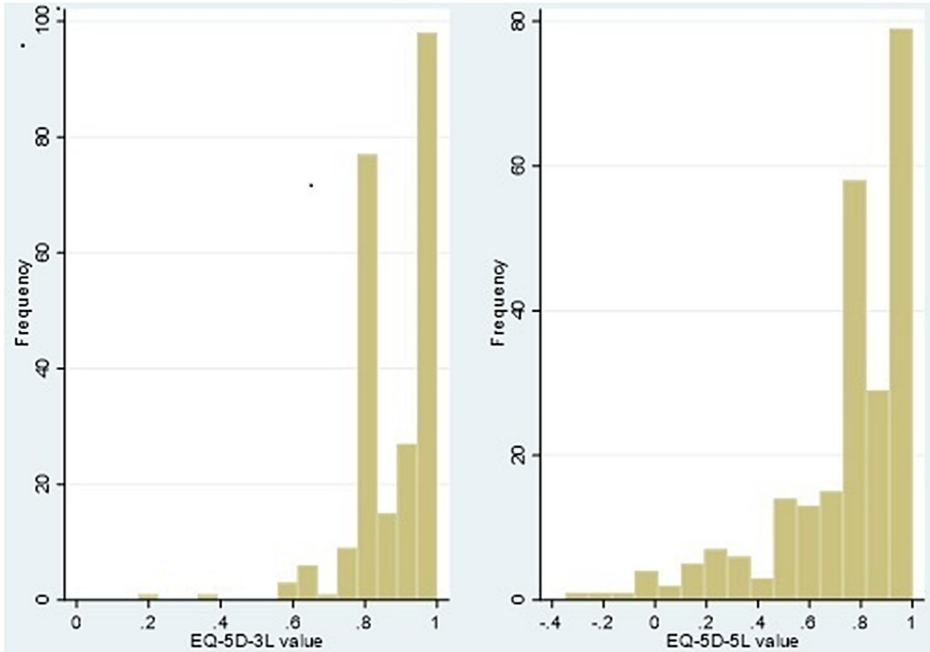

Fig. 1 Distribution of EQ-5D value

Table 2 Spearman's correlation coefficients between EQ-5D-5L, EQ-5D-3L, BASDAI, and BASFI scores

	Total score	MO	SC	UA	PD	AD	Utility
EQ-5D-3L							
BASDAI	-0.4071	0.5897	0.4750	0.6840	0.7181	0.6471	-0.4071
BASFI	-0.4144	0.6916	0.5737	0.7265	0.6275	0.5767	-0.4144
EQ-5D-5L							
BASDAI	-0.8056	0.6521	0.5687	0.6901	0.7614	0.7430	-0.8056
BASFI	-0.8310	0.7643	0.7131	0.7853	0.7034	0.7138	-0.8310

All correlation coefficients were statistically significant at $p < 0.001$ level. *EQ-5D* EuroQol five-dimensional questionnaire, *BASDAI* bath ankylosing spondylitis disease activity index, *BASFI* bath ankylosing spondylitis functional index

Table 3 Predictive performance of direct and indirect mapping models for the EQ-5D-3L ($N = 238$)

Model type	Mean	SD	Min	Max	MAE	RMSE	Spearman	CCC	Average rank
EQ-5D-3L value	0.8811	0.1207	0.1702	1.0000					
Full sample									
OLS	0.8811	0.0466	0.7590	0.9444					
CLAD	0.8788	0.0737	0.6788	1.0002					
Tobit	0.9100	0.0680	0.7317	1.0107					
ALDVMM	0.8799	0.0468	0.7463	0.9427					
BM	0.8944	0.0469	0.7254	0.9491					
Ologit	0.8791	0.0896	0.5966	0.9868					
Mlogit	0.8795	0.0881	0.6573	0.9882					
Cross validation									
OLS	0.8810	0.0473	0.7383	0.9499	0.0869	**0.1130**	0.4059	0.2410	5
CLAD	0.9100	0.0687	0.7077	1.0266	0.0850	0.1197	0.4057	0.2890	5
Tobit	0.8958	0.0745	0.6847	1.0207	0.0856	0.1174	0.4158	0.3230	3
ALDVMM	0.8788	0.0501	0.7224	0.9483	0.0896	0.1168	0.4114	0.2360	7
BM	0.8677	0.0553	0.7175	0.9569	0.0858	0.1132	0.4195	0.2810	3
Ologit	0.8786	0.0901	0.5758	0.9885	0.0848	0.1202	0.4365	**0.3640**	2
Mlogit	0.8786	0.0893	0.6124	0.9905	**0.0836**	0.1199	**0.4395**	0.3630	**1**

CLAD censored least absolute deviations, *ALDVMM* adjusted limited dependent variable mixture model, *BM* beta-mixture model, *Ologit* ordered logit, *Mlogit* multinomial logit. Bold indicates the best model in this performance indicator

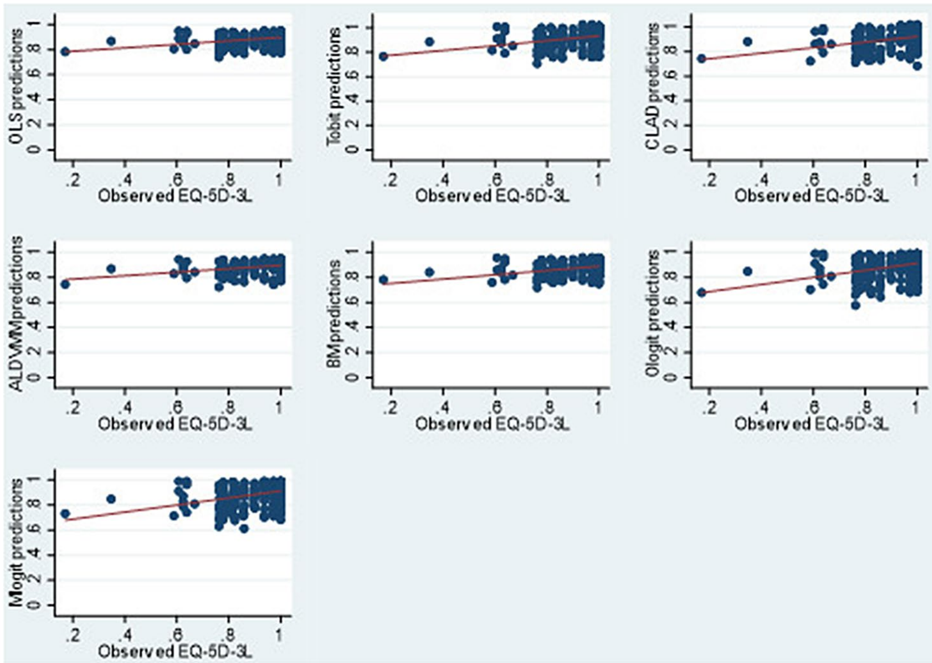

Fig. 2 Mean predicted vs. mean observed EQ-5D-3L values

Table 4 Predictive performance of direct and indirect mapping models for the EQ-5D-5L (*N* = 238)

Model type	Mean	SD	Min	Max	MAE	RMSE	Spearman	CCC	Average rank
EQ-5D-5L value	0.7388	0.2694	-0.3480	1.0000					
Full sample									
OLS	0.7388	0.2220	0.1635	1.0253					
CLAD	0.8042	0.1808	0.3349	1.0419					
Tobit	0.7716	0.2568	0.1083	1.1165					
ALDVMM	0.7450	0.2081	0.1617	0.9805					
BM	0.7575	0.2003	0.1513	0.9611					
Ologit	0.7561	0.2235	-0.0254	0.9745					
Mlogit	0.7573	0.2149	0.0200	0.9788					
Cross validation									
OLS	0.7396	0.2230	0.1472	1.0286	0.1091	0.1567	0.8471	0.7990	3
CLAD	0.7729	0.2581	0.0865	1.1269	0.1225	0.1666	0.8550	0.8020	6
Tobit	0.7928	0.2270	0.1099	1.1310	0.1182	0.1759	0.8491	0.7560	7
ALDVMM	0.7457	0.2078	0.1484	0.9835	0.1044	0.1567	0.8523	0.7880	2
BM	0.7312	0.2300	-0.0171	0.9791	0.1027	**0.1533**	**0.8570**	**0.8130**	**1**
Ologit	0.7567	0.2250	-0.0874	0.9773	0.1012	0.1588	0.8551	0.7960	5
Mlogit	0.7569	0.2172	0.0069	0.9830	**0.1006**	0.1578	0.8485	0.7930	4

CLAD censored least absolute deviations, *ALDVMM* adjusted limited dependent variable mixture model, *BM* beta-mixture model, *Ologit* ordered logit, *Mlogit* multinomial logit. Bold indicates the best model in this performance indicator

EQ-5D scores. The results confirmed that BASDAI and BASFI scores were significant predictors of health utility values. Secondly, the results indicated that the indirect mapping approach using Mlogit model exhibited the highest predicting performance in estimating EQ-5D-3L scores among patients with AS in China. Conversely, Wailoo's findings demonstrated that the bespoke mixture model was the optimal model in the UK [19]. This may be attributed to the characteristics of the patients included in this study. As illustrated in Table 1, they reported lower age, BASDAI, BASFI and EQ-5D-5L scores than those from the UK [19]. It is therefore probable that the AS patients in this study exhibited better health status, which may have resulted in the selection of a more suitable mapping algorithm for them than that employed in other studies. Furthermore, the majority of the observed EQ-5D-5L scores in our study were unevenly concentrated between 0.8 and 1, which may limit the

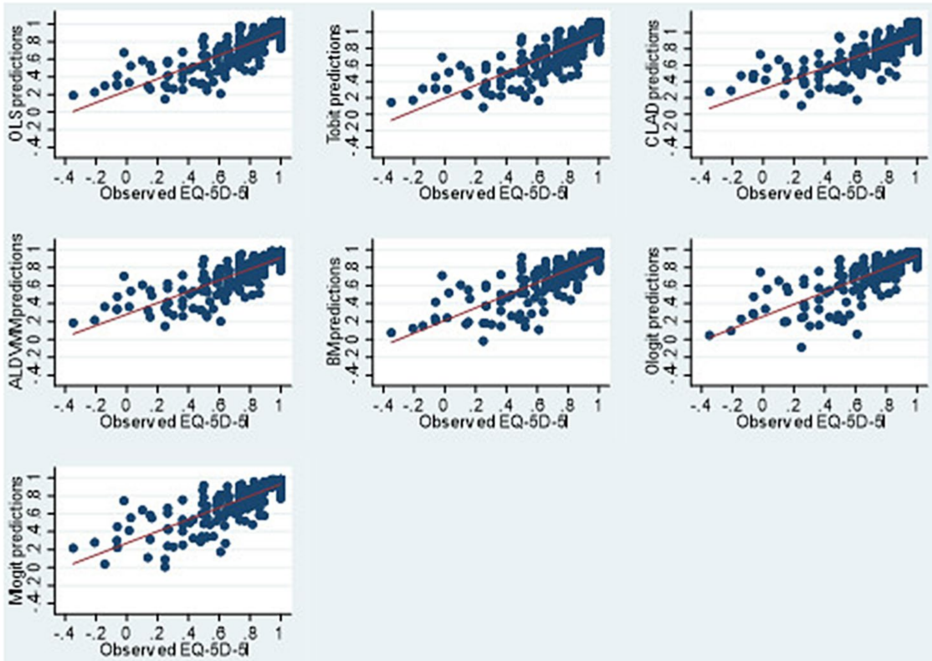

Fig. 3 Mean predicted vs. mean observed EQ-5D-5L values

ability of the direct mapping approach to fit a centralized distribution.

Thirdly, our study found that the BM was the best-performing model and the ALDVMM was the sub-optimal model in predicting EQ-5D-5L scores. Similar with our findings, Neilson also found the ALDVMM performed better than the indirect mapping approach and linear OLS models in converting BASDAI and BASFI to EQ-5D-5L scores [12]. These findings corroborate the advantage of bespoke mixture models in addressing the multimodality of EQ-5D-5L data distribution [39].

Fourthly, we found that the mapping algorithms for EQ-5D-3L and EQ-5D-5L are different even when using the same explanatory variables and patient data. This demonstrated that the EQ-5D-3L and EQ-5D-5L could not be employed interchangeably in AS clinical studies. On the one hand, both the observed and predicted EQ-5D-3L scores exhibited higher means and minimum values but lower standard deviation than the EQ-5D-5L did. This indicated that the distribution of EQ-5D-3L was more concentrated than that of EQ-5D-5L, based on the same patients with AS. Consequently, the MAE and RMSE of mapping models for EQ-5D-3L were slightly smaller than those of EQ-5D-5L. On the other hand, the EQ-5D-5L scores exhibited a stronger relationship with BASDAI and BASFI than EQ-5D-3L scores did. Furthermore, the Spearman's correlation coefficient and CCC between observed and predicted EQ-5D-5L values were much higher than those of EQ-5D-3L, indicating that EQ-5D-5L exhibited a higher conceptual overlap with BASDAI and BASFI than EQ-5D-3L did. Consequently, we recommended that EQ-5D-5L should be broadly employed in health economic evaluation and BM could be selected to predict health utility values from BASDAI and BASFI from clinical research among patients with AS.

Our study made two aspects of contribution. Firstly, we developed country-specific mapping algorithms for converting BASDIA and BADFI to EQ-5D health utility values based on patients with AS from China. The mapping algorithms based on BM model could be applied in future health economic evaluations of AS pharmacies and therapies in clinical research conducted in China. Secondly, we compared the conceptual overlap between two AS disease-specific instruments (BASDAI and BASFI) and two versions of the EQ-5D questionnaire, and compared the predictive performances of EQ-5D-3L and EQ-5D-5L. The findings indicated that EQ-5D-5L exhibited a higher conceptual overlap and predictive performance than EQ-5D-3L. Consequently, it is implied that mapping studies of AS in other countries should select EQ-5D-5L values as target outcome measures. The novelty in this research lies in the development of mapping algorithms to predict EQ-5D values from BASDAI and BASFI among Asian patients, whereas previous studies were conducted among the UK or Czech patients with AS. Furthermore, we also employed an indirect mapping approach and provided corresponding mapping algorithms, which enabled researchers in other countries to predict EQ-5D values in country-specific tariffs.

Nevertheless, it should be noted that this study had several potential limitations. Firstly, the patients with AS were obtained from a single hospital in Shandong province, which were not representative samples of Chinese patients. Nevertheless, Zhao found that there was no statistically significant difference in the prevalence of AS by sampling resource and district in China [6], indicating that the mapping algorithms could be generally used in health economic evaluation studies among Chinese populations. However, there may be still selection bias in our study, given that this hospital is regarded as the leading medical facility in Shandong province and typically admits patients with severe AS. Secondly, there was no external samples to validate the mapping algorithms, which may limit the generalizability of results. Although five-fold cross-validation was conducted, future research should validate the mapping algorithms when large-scale, representative external samples are available. Thirdly, the mapping algorithms demonstrated an overestimation of EQ-5D health utility values for patients with very poor health status, particularly for EQ-5D-3L scores .

Conclusions

To the best of our knowledge, this is the first study to develop Chinese country-specific mapping algorithms from BASDAI and BASFI to EQ-5D health utility values in patients with AS. The results demonstrated that direct mapping algorithms based on the Mlogit model exhibited superior performance in predicting EQ-5D-3L values, and BM was the optimal model for predicting EQ-5D-5L scores. Given the higher conceptual overlap between EQ-5D-5L and BASDAI and BASFI, we propose that EQ-5D-5L should be used as the target outcome measure for conducting CUA in AS, and BM mapping algorithms should be employed to generate health utility values from disease-specific questionnaires in clinical research.

Abbreviations

AS	Ankylosing spondylitis
BASDAI	Bath Ankylosing Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BM	Beta-mixture model
HRQoL	Health-related quality of life
CUA	Cost-utility analysis
QALYs	Quality-adjusted life years
MAPS	Mapping onto Preference-based measures reporting Standards
CLAD	Censored least absolute deviations
ALDVMM	Adjusted limited dependent variable mixture model
Ologit	Ordered logit
Mlogit	Multinomial logit
MAE	Mean absolute error
RMSE	Root mean squared error
CCC	Concordance correlation coefficient
SD	Standard deviation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12955-024-02276-5.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Author contributions

All authors contributed to the study conception and design. Material preparation and data collection were conducted by Yingying Fang, Xianyu Zhang and Shunping Li. Data analysis were performed by Chaofan Li, Liyong Lu and Huizhi Yang. The first draft of the manuscript was written by Lingjia Yu and Chaofan Li and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Code availability

The codes in this article will be shared on reasonable request to the corresponding author.

Ethical approval

This study was approved by the Ethics Review Board of the School of Health Care Management, Shandong University (No. ECSHCMSDU20181102). This research adhered to the tenets of the Declaration of Helsinki.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Competing interests

The authors declare no competing interests.

Author details

[1]Nursing Department, Rheumatology department, Qilu hospital of Shandong University, Jinan 250012, China
[2]Shunyi District Center for Disease Control and Prevention, Beijing 101300, China
[3]Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Wenhua Xi Road 44, Jinan 250012, China
[4]NHC Key Lab of Health Economics and Policy Research, Shandong University, Wenhua Xi Road 44, Jinan 250012, China
[5]Center for Health Preference Research, Shandong University, Wenhua Xi Road 44, Jinan 250012, China

Received: 8 December 2023 / Accepted: 19 July 2024
Published online: 07 August 2024

References

1. Taurog JD, Chhabra A, Colbert RA. Ankylosing spondylitis and Axial Spondyloarthritis. N Engl J Med. 2016;374(26):2563–74. https://doi.org/10.1056/NEJMra1406182.
2. Braun J, Sieper J. Ankylosing spondylitis. Lancet. 2007;369(9570):1379–90. https://doi.org/10.1016/S0140-6736(07)60635-7.
3. Boonen A, van der Linden SM. The burden of ankylosing spondylitis. J Rheumatol. 2006;33:4–11.
4. Korotaeva T, Dina O, Holdsworth E, Fallon L, Milligan G, Meakin S, et al. Investigating diagnosis, treatment, and burden of disease in patients with ankylosing spondylitis in Central Eastern Europe and the United States: a real-world study. Clin Rheumatol. 2021;40(12):4915–26. https://doi.org/10.1007/s10067-021-05864-8.

5. Zhu W, He X, Cheng K, Zhang L, Chen D, Wang X, et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. Bone Res. 2019;7:22. https://doi.org/10.1038/s41413-019-0057-8.
6. Zhao J, Huang C, Huang H, Pan J, Zeng L, Luo M, et al. Prevalence of ankylosing spondylitis in a Chinese population: a systematic review and meta-analysis. Rheumatol Int. 2020;40(6):859–72. https://doi.org/10.1007/s00296-020-04537-0.
7. Chen Y, Dong H, Wei Y, Yang Y, Ming J, Yu H. Using health technology assessment to inform insurance reimbursement of high technology medicines in China: an example of cancer immunotherapy. BMJ. 2023. https://doi.org/10.1136/bmj-2022-069963. e069963.
8. Committee BAOI. Expert Consensus on Long-Term Management of ankylosing spondylitis (2021). Chin J Integr Traditional Western Med. 2021;41(12):1426–34.
9. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C et al. (2022). Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *BMJ*, e067975. https://doi.org/10.1136/bmj-2021-067975.
10. UK Office For Health Improvement and Disparities. (2020). Guidance: Cost utility analysis: health economic studies. https://www.gov.uk/guidance/cost-utility-analysis-health-economic-studies. Accessed 09–25, 2023.
11. Arnold D, Girling A, Stevens A, Lilford R. Comparison of direct and indirect methods of estimating health state utilities for resource allocation: review and empirical analysis. BMJ. 2009;339:b2688. https://doi.org/10.1136/bmj.b2688.
12. Neilson AR, Jones GT, Macfarlane GJ, Pathan EM, McNamee P. Generating EQ-5D-5L health utility scores from BASDAI and BASFI: a mapping study in patients with axial spondyloarthritis using longitudinal UK registry data. Eur J Health Econ. 2022;23(8):1357–69. https://doi.org/10.1007/s10198-022-01429-x.
13. Wei JC, Tsou H, Leong P, Chen C, Huang J. Head-to-Head comparison of Etanercept vs. Adalimumab in the treatment of Ankylosing spondylitis: an open-label randomized controlled crossover clinical trial. Front Med. 2020;7. https://doi.org/10.3389/fmed.2020.566160.
14. Emery P, Van Keep M, Beard S, Graham C, Miles L, Jugl SM, et al. Cost effectiveness of Secukinumab for the treatment of active ankylosing spondylitis in the UK. PharmacoEconomics. 2018;36(8):1015–27. https://doi.org/10.1007/s40273-018-0675-9.
15. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. 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. https://doi.org/10.1016/j.jval.2016.11.006.
16. Ara R, Rowen D, Mukuria C. The Use of Mapping to Estimate Health State Utility values. PharmacoEconomics. 2017;35(S1):57–66. https://doi.org/10.1007/s40273-017-0548-7.
17. Ara RM, Reynolds AV, Conway P. The cost-effectiveness of etanercept in patients with severe ankylosing spondylitis in the UK. Rheumatology. 2007;46(8):1338–44. https://doi.org/10.1093/rheumatology/kem133.
18. Mlcoch T, Sedova L, Stolfa J, Urbanova M, Suchy D, Smrzova A, et al. Mapping the relationship between clinical and quality-of-life outcomes in patients with ankylosing spondylitis. Expert Rev Pharmacoecon Outcomes Res. 2017;17(2):203–11. https://doi.org/10.1080/14737167.2016.1200468.
19. Wailoo A, Hernandez M, Philips C, Brophy S, Siebert S. Modeling Health State Utility Values in Ankylosing spondylitis: comparisons of direct and indirect methods. Value Health. 2015;18(4):425–31. https://doi.org/10.1016/j.jval.2015.02.016.
20. Hernandez Alava M, Wailoo A, Chrysanthou G, Barcelos F, van Gaalen FA, Santos H, et al. Measuring quality of life of patients with axial spondyloarthritis for economic evaluation. RMD Open. 2022;8(1):e001955. https://doi.org/10.1136/rmdopen-2021-001955.
21. Chen G, McKie J, Khan MA, Richardson JR. Deriving health utilities from the MacNew Heart Disease Quality of Life Questionnaire. Eur J Cardiovasc Nurs. 2015;14(5):405–15. https://doi.org/10.1177/1474515114536096.
22. Chen Y, Zhao K, Liu G, Chen W. Health technology assessment to inform decision making in China: progress, challenges, and sustainability. BMJ. 2023;e068910. https://doi.org/10.1136/bmj-2021-068910.
23. Wailoo A, Alava MH, Pudney S, Barton G, O'Dwyer J, Gomes M, et al. An International comparison of EQ-5D-5L and EQ-5D-3L for use in cost-effectiveness analysis. Value Health. 2021;24(4):568–74. https://doi.org/10.1016/j.jval.2020.11.012.
24. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. Preferred reporting items for studies mapping onto preference-based outcome measures: the MAPS Statement. Appl Health Econ Health Policy. 2015;13(5):437–43. https://doi.org/10.1007/s40258-015-0191-4.
25. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. The MAPS Reporting Statement for studies Mapping onto generic preference-based outcome measures: explanation and elaboration. PharmacoEconomics. 2015;33(10):993–1011. https://doi.org/10.1007/s40273-015-0312-9.
26. Qilu Hospital. (2024). Introduction to Qilu Hospital of Shandong University. https://www.qiluhospital.com/en/index.php?m=content&c=index&a=show&catid=135&id=626. Accessed 06–12, 2024.
27. Linden SVD, Valkenburg HA, Cats A. Evaluation of Diagnostic Criteria for Ankylosing Spondylitis. Arthr Rhuem. 1984;27(4):361–8. https://doi.org/10.1002/art.1780270401.
28. Rabin R, Charro FD. EQ-5D: a measure of health status from the EuroQol Group. Annals Med (Helsinki). 2001;33(5):337–43. https://doi.org/10.3109/07853890109002087.
29. 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. https://doi.org/10.1007/s11136-012-0200-0.
30. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, Present and Future. Appl Health Econ Health Policy. 2017;15(2):127–37. https://doi.org/10.1007/s40258-017-0310-5.
31. Buchholz I, Janssen MF, Kohlmann T, Feng Y. A systematic review of studies comparing the Measurement Properties of the three-level and five-level versions of the EQ-5D. PharmacoEconomics. 2018;36(6):645–61. https://doi.org/10.1007/s40273-018-0642-5.
32. Christiansen ASJ, Møller MLS, Kronborg C, Haugan KJ, Køber L, Højberg S, et al. Comparison of the three-level and the five-level versions of the EQ-5D. Eur J Health Econ. 2021;22(4):621–8. https://doi.org/10.1007/s10198-021-01279-z.
33. Wang H, Patrick DL, Edwards TC, Skalicky AM, Zeng H, Gu W. Validation of the EQ-5D in a general population sample in urban China. Qual Life Res. 2012;21(1):155–60. https://doi.org/10.1007/s11136-011-9915-6.
34. Liu GG, Wu H, Li M, Gao C, Luo N. Chinese Time Trade-Off Values for EQ-5D Health states. Value Health. 2014;17(5):597–604. https://doi.org/10.1016/j.jval.2014.05.007.
35. 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. https://doi.org/10.1016/j.jval.2016.11.016.
36. Lin Z, Gu J, He P, Gao J, Zuo X, Ye Z, et al. Multicenter validation of the value of BASFI and BASDAI in Chinese ankylosing spondylitis and undifferentiated spondyloarthropathy patients. Rheumatol Int. 2011;31(2):233–8. https://doi.org/10.1007/s00296-009-1313-9.
37. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE Health Technology assessments. Value Health. 2013;16(1):202–10. https://doi.org/10.1016/j.jval.2012.10.010.
38. 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. https://doi.org/10.1016/j.jval.2018.06.006.
39. 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. https://doi.org/10.1016/j.jval.2017.09.017.
40. Hernández Alava M, Wailoo AJ, Ara R. Tails from the Peak District: Adjusted Limited Dependent Variable mixture models of EQ-5D Questionnaire Health State Utility values. Value Health. 2012;15(3):550–61. https://doi.org/10.1016/j.jval.2011.12.014.
41. Gray LA, Alava MH. A command for fitting mixture regression models for bounded dependent variables using the Beta distribution. Stata J. 2018;18(1):51–75. https://doi.org/10.1177/1536867X1801800105.
42. Gray AM, Rivero-Arias O, Clarke PM. Estimating the Association between SF-12 responses and EQ-5D utility values by response mapping. Med Decis Making. 2006;26(1):18–29. https://doi.org/10.1177/0272989X05284108.
43. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. A comparison of direct and indirect methods for the Estimation of Health Utilities from clinical outcomes. Med Decis Making. 2014;34(7):919–30. https://doi.org/10.1177/0272989X13500720.
44. Payakachat N, Summers KH, Pleil AM, Murawski MM, Thomas J, Jennings K, et al. Predicting EQ-5D utility scores from the 25-item National Eye Institute Vision Function Questionnaire (NEI-VFQ 25) in patients with age-related

macular degeneration. Qual Life Res. 2009;18(7):801–13. https://doi.org/10.1007/s11136-009-9499-6.
45. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989;45(1):255–68.
46. Yang F, Wong CKH, Luo N, Piercy J, Moon R, Jackson J. Mapping the kidney disease quality of life 36-item short form survey (KDQOL-36) to the EQ-5D-3L and the EQ-5D-5L in patients undergoing dialysis. Eur J Health Econ. 2019;20(8):1195–206. https://doi.org/10.1007/s10198-019-01088-5.
47. StataCorp. (2017). Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC

Publisher's Note