ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Patient-Reported Outcomes

Mapping the Lequesne Functional Index Into the EQ-5D-5L Utility Index in Patients With Knee Osteoarthritis

Nadia Dardenne, MSc, Anne-Françoise Donneau, PhD, Olivier Bruyère, PhD

ABSTRACT

Objective: This study aimed to map the Lequesne index onto the EuroQol 5 Dimension (EQ-5D-5L) utility index for patients with knee osteoarthritis.

Methods: Baseline data from a previous randomized controlled trial were used; 461 patients were involved in the mapping development, and 230 in the validation phase. Various modeling techniques, including generalized linear models, tobit, and beta regression, were used. Factors such as age, sex, and body mass index were considered as covariates. Model selection was based on criteria such as Akaike and Bayesian information criteria, adjusted R^2, mean absolute error, and root mean squared error. Validation involved assessing the preselected models using mean absolute error, root mean squared error, and intraclass correlation coefficient. This study follows the Mapping Onto Preference-Based Measures Reporting Standards statement.

Results: Five models were developed, with 2 incorporating age, sex, with or without body mass index along with the Lequesne index showing the best fit across regressions. Validation results were similar for the 3 regressions, with beta regression models exhibiting wider ranges closer to the validation data set. Intraclass correlation coefficient values were better for beta regression models. Both models tended to overpredict for lower EQ-5D-5L values and underpredict for better health status.

Conclusion: These mapping functions, the first of their kind, effectively translate the Lequesne index to EQ-5D-5L values in patients with knee osteoarthritis. They demonstrate satisfactory fit and precision, providing valuable tools for clinicians and researchers, particularly in situations where generic preference-based health-related quality of life instruments are inaccessible for utility derivation in cost-effectiveness studies.

Keywords: arthritis, Lequesne, mapping, utility.

VALUE HEALTH. 2024; ■(■):■–■

Highlights

- Addressing the burden of knee osteoarthritis (OA), this study pioneers the development of mapping functions to translate Lequesne index scores to EQ-5D-5L values, providing a novel tool for future investigations in health outcomes assessment for patients with knee OA.
- Using rigorous modeling techniques, including generalized linear model, tobit, and beta regressions, the research establishes 2 models with age, sex, body mass index, and Lequesne index as key factors, showing good goodness-of-fit indexes, widening the scope for accurate health predictions.
- These mapping functions provide clinicians and researchers with valuable tools to accurately assess the health status of patients with knee OA, especially in situations where generic preference-based instruments are inaccessible, thereby facilitating progress in cost-effectiveness studies.

Introduction

Knee osteoarthritis (OA) stands as one of the most common degenerative joint diseases among older individuals, causing functional disability and chronic pain. In addition, OA places a significant strain on social and health resources, creating a substantial economic burden for patients, healthcare providers, and society as a whole.[1] This underscores the increasing importance of cost-effectiveness evidence for decision makers at various levels, given the necessity of efficient allocation of limited healthcare resources.[2] Cost-effectiveness studies often use quality-adjusted life-years (QALYs) as a key measure. QALYs consider both the quality and quantity of life, enabling comprehensive comparisons among different treatment approaches, patient groups, and clinical contexts.[3] To enable the calculation of a QALY measure, the EuroQol group introduced the EQ-5D, a health-related quality of life (HRQoL) instrument.[4] The EQ-5D comprises 5 health-related questions or dimensions, which can be transformed into a single index serving as a utility measure. This EQ-5D utility index offers a straightforward, universal measure applicable in both clinical and economic assessments. Not only is it widely used in this type of evaluation, but its adoption is also recommended in all economic analyses to ensure consistency and comparability across studies.

In cases where EQ-5D data are unavailable, studies necessitate the use of mapping or crosswalking techniques to estimate EQ-5D values based on the available outcome measures. These measures might encompass clinical symptoms; nonpreference-based quality of life assessments, both generic and condition specific; or preference-based measures other than EQ-5D. In OA, several studies have endeavored to map the Western Ontario and McMaster Universities Arthritis Index (WOMAC) questionnaire or the Oxford questionnaire onto EQ-5D values, facilitating the calculation of QALYs from these specific health assessment tools.[5-10] Notably, the WOMAC questionnaire is likely the most widely used tool for this purpose, although it requires payment of a fee for its usage. In contrast, the Lequesne index, another frequently used measure in clinical trials, is available for use without any cost.[11] Despite the prevalence of the WOMAC questionnaire,

1098-3015/$36.00 - see front matter Copyright © 2024, International Society for Pharmacoeconomics and Outcomes Research, Inc. Published by Elsevier Inc.

the availability of the Lequesne index as a free tool highlights its accessibility and affordability, underscoring its significance in research and clinical settings. Importantly, there is a notable gap in research given that no studies have yet explored mapping the Lequesne index into EQ-5D values, indicating a potential avenue for future investigations in health outcomes assessment.

This study aimed to develop mapping functions that translate the Lequesne index scores to EQ-5D-5L values, a well-validated 5-level version of the EQ-5D,[12,13] in patients with knee OA. This involved using various statistical strategies to develop these functions and subsequently comparing and, if possible, validating them.

Methods

Study Population

We used the baseline data of a 24-week randomized, double-blind, placebo-controlled, parallel-group trial (ClinicalTrials.gov identifier NCT03200288) evaluating the clinical efficacy and safety of a single intra-articular injection of high- and low- molecular weight hyaluronic acid formulation in comparison with a single intra-articular injection of placebo (saline) for managing pain in individuals with symptomatic knee OA.[14] Participants were enrolled from outpatient facilities in public, private, and university clinics, as well as hospitals across Belgium, Germany, Hungary, Italy, and Poland. Eligible participants included both female and male subjects aged between 40 and 80 years, diagnosed of primary knee OA based on the American College of Rheumatology criteria. They exhibited Kellgren and Lawrence grade 2 to 3 radiographic evidence of OA, experienced symptoms for at least 3 months, and had moderate-to-severe pain at the time of enrolment. Screening pain intensity in the affected knee, measured using a 100 mm visual analog scale, needed to be equal to or greater than 40 mm on the visual analog scale (and less than 20 mm in the contralateral knee). These criteria were confirmed at randomization after discontinuation of analgesics/nonsteroidal anti-inflammatory drugs.

This study was written according to the Mapping Onto Preference-Based Measures Reporting Standards statement.

Measurements

The EQ-5D-5L questionnaire evaluates health status in 5 domains: mobility, self-care, routine activities, pain/discomfort, and anxiety/depression.[15] Each response level ranges between 1 and 5. The scores from the descriptive component can be reported as a 5-digit number, known as profile scores. The profile score can be converted into a utility index using a country-specific value set. All countries' value sets were obtained from the Euro-Qol.org website. The EQ-5D-5L has shown good psychometric properties in patients with hip or knee OA.[12] For this particular study, the EQ-5D-5L values for each patient were transformed into utility values (EQ-5D-5L index values) using the data from 5 major European countries for which utility value sets are currently available (Denmark, France, Germany, The Netherlands, and United Kingdom).[16-20]

The Lequesne index is a disease-specific questionnaire designed for joint-related assessments. It comprises 3 dimensions: pain or discomfort (assessed by 5 items), the maximum walking distance (evaluated through 2 items), and activities of daily living (assessed using 4 items).[11] Each dimension has a maximum total score of 8, leading to a total score range of 0 to 24. In particular, the scores for pain and activities of daily living scales vary from 0 (indicating no pain or functional limitation) to 8 (representing extreme pain or functional limitation). The "maximum distance walked" section is graded from 0 (unlimited) to 6 (less than 100 meters). In addition, the score is adjusted upward by 1 point if the patient uses 1 walking stick or crutch or by 2 points if 2 walking sticks or crutches are used. The Lequesne index consolidates symptoms and function into a single global index score, ranging from 0 to 24, where higher scores indicate poorer health status.

Statistical Analysis

Results were expressed as numbers and frequencies for qualitative parameters and as mean and standard deviation, median (P50) and interquartile range (P25-P75), and range for quantitative parameters. The normality of the distribution of the quantitative parameters was investigated using the mean-median comparison, the histogram, and quantile-quantile plot and tested with the Shapiro-Wilk hypothesis test.

The mapping functions were developed using two-thirds of the data, whereas the remaining third was used to validate them. No missing data were observed in our database.

Linear association between the EQ-5D-5L and Lequesne scores was investigated using the nonparametric Spearman correlation coefficient ($R_{Spearman}$) and graphically. A Student t test was applied to test the significance of this correlation.

To derive mapping functions, 3 types of regression were computed: the generalized linear model (GLM), the tobit regression, and the beta regression. As mentioned in Bilbao et al,[21] the GLM supposes continuous outcome and is based on the strong assumptions of homoscedasticity and normality of residuals. Given that EQ-5D-5L is a discrete variable whose values lie within a given interval, this type of regression is generally not appropriate.[22] The tobit regression can be seen as an alternative even if it assumed normality of residuals too.[23,24] Finally, the beta regression was also considered, this method being largely recommended in the literature to model this type of outcome.[25,26] Given that this methodology assumed that the outcome must be defined in the interval (0, 1), values equal to 1 were changed to 0.99. No values equal to 0 were observed in the data.

Furthermore, generalized additive models (GAMs)[27] were applied to investigate whether the relation between the Lequesne and EQ-5D-5L score could be considered as linear or whether second or third power association between these 2 scores could be assumed.

For these 3 types of regression, 5 models were considered: model 1 with the Lequesne score as the only predictor; models 2 and 3 with, respectively, a second and a third power of Lequesne score as predictors; model 4, with the Lequesne score as predictor and the covariates age, body mass index (BMI), and sex; and finally model 5, which considered only covariates with $P < .10$. In fact, for models 4 and 5, several demographic (age and sex) and clinical (BMI) variables were considered important in predicting the quality of life of patients with OA and were therefore included in the prediction model.[28] Selection of a best model was based on goodness-of-fit measures such as the Akaike information criterion (AIC) and Bayesian information criterion (BIC) and the adjusted R-squared (R^2). The following measures of predictive accuracy were also calculated: the mean absolute error (MAE) and the root mean squared error (RMSE).[29,30] The 95% predictable intervals were calculated using the same methodology described in Bilbao et al[21] and using bootstrapping method. Actually, 1000 samples were generated, and for each sample, the MAE and the RMSE were calculated. The 95% predictable intervals were then derived from the 2.5th and 97.5th percentiles of the distribution of the bootstrapped values. In conclusion, the lower the AIC, BIC, MAE, and

RMSE and the higher R^2, the better the goodness of fit. Distribution of residuals was also investigated graphically. Coefficients and standard errors of the selected best models were then reported.

In the second step, preferred mapping functions were validated by means MAE, RMSE, and 2-way fixed intraclass correlation coefficient (ICC) (A, 1)[31] with 95% CI. Values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability.[32] Agreement between predicted and observed values for the score of utility was also represented by using a Bland-Altman plot. Mean differences and limits of agreement were reported.

Results were significant at the 5% critical level ($P < .05$). The statistical analysis were conducted using SAS (version 9.4 for Windows; SAS Institute Inc, Cary, NC) statistical package and R (version 4.3; R Core Team 2023) with particular packages VGAM[33,34] and irr.[35]

Results

The 691 subjects were allocated so that 66.7% (n = 461) were in the training data set and 33.3% (n = 230) were in the validation data set. The median EQ-5D-5L score was 0.64 (0.55-0.73) in the training set and was 0.63 (0.53-0.72) in the validation set. The median Lequesne score was also similar for the training and validation set with median value equal to 11.5 (9.00-14.0) and 12.0 (9.50-14.0), respectively. Distribution of the covariates, namely age, BMI, and gender, showed no differences in the training and validation data set (Table 1).

EQ-5D-5L and Lequesne scores were significantly correlated ($R_{Spearman} = -0.55$, $P < .0001$ in the training data set; $R_{Spearman} = -0.54$, $P < .0001$, in the validation data set) (Appendix Fig. 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.06.017). For the training data set, GAMs confirmed this linear relationship with no significant second or third power ($P > .05$) association between these 2 scores.

Five models using the 3 types of regression described earlier, namely GLM, tobit, and beta regression, were investigated in terms of fit measures, predictive accuracy and index, and residuals. Results were presented in Table 2. Whatever the methodology used, all models highlighted a statistically significant linear association between the Lequesne score and the EQ-5D-5L score ($P < .0001$).

For the GLM regressions, lower AIC values were found for models 4 and 5, the 2 models where covariates were included. Even if a lower BIC value was found for the model with only a linear relationship between the 2 scores, the model 5, where age and BMI were included and significantly associated with EQ-5D-5L ($P < .05$), presented a close result. The R^2 values remained constant across all models. None of the 5 models presented better values for predictive accuracy and index. For these reasons, the models 4 and 5 were preferred, models for which, moreover, the homoscedasticity of the residuals and the assumption of normality were checked (Appendix Fig. 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.06.017).

Same observations could be made for the tobit regression models, with lower AIC values for the models 4 and 5, lower BIC values for the models 1 and 5, and constant R^2, predictive index, and accuracy values across all models. In addition, the covariates age and BMI were also significantly associated with EQ-5D-5L ($P < .05$). Investigation of the homogeneity and normality of residuals led to similar conclusions as for GLM regressions (Appendix Fig. 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.06.017).

Finally, the results of the beta regression models, which led to the same conclusions as those of the tobit and GLM regression models, did not clearly point to a better model. The values of R^2, predictive accuracy, and index remained similar. However, for model 5, only age was statistically significant whereas P value was $< .10$ for BMI. Residuals were also plotted for information (Appendix Fig. 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2024.06.017).

Whatever the type of regression, all models overestimated the observed mean and underestimated the standard deviation. All models also showed similar MAE and RMSE values although these values were slightly better for GLM and tobit regression. In contrast, the ranges obtained using the beta regression models were wider and closer to those of the validation data set (Table 3). ICC values were also better for the beta regression models and indicated moderate reliability (ICC > 0.50) whereas the other methods indicated poor reliability. Mean differences and limits of agreement remained stables according the model. The Bland-Altman plot (Fig. 1) showed that, in absolute values, prediction errors were greater for lower values of EQ-5D-5L, whatever the model and the type of regression. Furthermore, the models tented to overpredict for lower EQ-5D-5L values whereas they tended to underpredict for better health status, and this trend was more marked for GLM models.

Discussion

In this study, using the baseline data of a large randomized controlled trial (RCT) in patients with knee OA, we have developed a mapping function that accurately predicts EQ-5D-5L utility

Table 1. Description of the training and validation data set.

Variables		All (N = 691)	Training set (n = 461)	Validation set (n = 230)
EQ-5D-5L (0-1), P50 (P25-P75)		0.64 (0.54-0.73)	0.64 (0.55-0.73)	0.63 (0.53-0.72)
Mean (SD)		0.62 (0.16)	0.62 (0.16)	0.60 (0.16)
Range		0.0032-1.00	0.0032-1.00	0.11-1.00
Lequesne (0-24), P50 (P25-P75)		12.0 (9.0-14.0)	11.5 (9.00-14.0)	12.0 (9.50-14.0)
Age (years), P50 (P25-P75)		64.0 (58.0-63.7)	64.0 (58.0-70.0)	64.0 (58.0-70.0)
BMI (kg/m^2), P50 (P25-P75)		29.1 (26.3-30.9)	29.0 (26.2-30.8)	29.4 (26.6-30.9)
Gender, n (%)	Women	461 (66.7)	310 (67.2)	151 (65.7)
	Men	230 (33.3)	151 (32.8)	79 (34.3)

BMI indicates body mass index.

Table 2. Assessment of fit measures, predictive accuracy and index for different models predicting EQ-5D-5L as a function of Lequesne score in the training data set (n = 461).

Variables		Without covariates			With covariates	
		Model 1	Model 2	Model 3	Model 4	Model 5
						Model 4 with $P < .10$
GLM	Variables	L	L + L^2	L + L^2 + L^3	L + age + sex + BMI	L + age + BMI
Fit measures						
	AIC	−541.07	−539.42	−539.42	−547.21	−547.20
	BIC	−528.67	−522.89	−518.76	−522.41	−526.53
	R^2	0.31	0.31	0.31	0.32	0.32
Predictive accuracy						
	MAE (95% PI)	0.103 (0.095-0.111)	0.102 (0.095-0.111)	0.103 (0.095-0.111)	0.101 (0.093-0.109)	0.102 (0.094-0.110)
	RMSE (95% PI)	0.133 (0.124-0.143)	0.133 (0.123-0.143)	0.133 (0.123-0.143)	0.131 (0.122-0.141)	0.131 (0.121-0.141)
Predictive index						
	Mean (SD)	0.62 (0.090)	0.62 (0.090)	0.62 (0.090)	0.62 (0.093)	0.62 (0.092)
	Range	0.38-0.87	0.36-0.85	0.43-0.80	0.40-0.87	0.38-0.88
Tobit	Variables	L	L + L^2	L + L^2 + L^3	L + age + sex + BMI	L + age + BMI
Fit measures						
	AIC	−524.81	−523.04	−522.84	−530.99	−530.89
	BIC	−512.41	−506.50	−502.17	−506.20	−510.22
Predictive accuracy						
	MAE (95% PI)	0.103 (0.096-0.111)	0.103 (0.095-0.111)	0.103 (0.095-0.111)	0.102 (0.094-0.109)	0.102 (0.094-0.109)
	RMSE (95% PI)	0.134 (0.124-0.142)	0.134 (0.123-0.143)	0.133 (0.123-0.142)	0.132 (0.121-0.141)	0.132 (0.121-0.142)
Predictive index						
	Mean (SD)	0.62 (0.091)	0.62 (0.091)	0.62 (0.091)	0.62 (0.093)	0.62 (0.093)
	Range	0.38-0.87	0.36-0.86	0.43-0.80	0.40-0.87	0.38-0.88
Beta	Variables	L	L + L^2	L + L^2 + L^3	L + age + sex + BMI	L + age + BMI
Fit measures						
	AIC	−531.28	−529.34	−528.03	−534.99	−534.58
	BIC	−518.88	−512.81	−507.36	−510.19	−513.91
	R^2	0.31	0.31	0.31	0.33	0.32
Predictive accuracy						
	MAE (95% PI)	0.104 (0.095-0.112)	0.103 (0.095-0.112)	0.104 (0.095-0.112)	0.102 (0.094-0.110)	0.102 (0.094-0.110)
	RMSE (95% PI)	0.134 (0.123-0.144)	0.133 (0.123-0.143)	0.133 (0.123-0.142)	0.132 (0.121-0.141)	0.132 (0.121-0.141)
Predictive index						
	Mean (SD)	0.62 (0.093)	0.62 (0.093)	0.62 (0.093)	0.62 (0.095)	0.61 (0.095)
	Range	0.35-0.84	0.36-0.84	0.40-0.82	0.38-0.83	0.36-0.84

AIC indicates Akaike information criterion; BIC, Bayesian information criterion; BMI, body mass index; GLM, generalized linear model; L, Lequesne score; MAE, mean absolute error; PI, predictable interval; RMSE, root mean squared error.

based on the Lequesne index. To the best of our knowledge, this is the first attempt of an algorithm mapping the Lequesne index to any of the EQ-5D scales in patients with OA. The derived mapping functions demonstrated a satisfactory fit, providing a means of estimating utilities for cost-effectiveness studies in scenarios where generic HRQoL preference-based questionnaires are not available (Fig. 2).

In the area of hip or knee OA research, several studies have linked the EQ-5D to disease-specific measures.[36] Although the WOMAC index has been the primary focus of many investigations,

Table 3. Predictive accuracy and index for the preferred models predicting EQ-5D-5L as a function of Lequesne score in the validation data set (n = 230).

Variables		Model 4	Model 5 Model 4 with $P < .10$
GLM		L + age + sex + BMI	L + age + BMI
Parameters, coefficient ± SE			
Intercept		0.876 ± 0.0740*	0.884 ± 0.074*
L		−0.0251 ± 0.00178*	−0.0253 ± 0.00177*
Age		−0.00161 ± 0.000740†	−0.00172 ± 0.00074†
Sex		0.01882 ± 0.0133	/
BMI		0.00453 ± 0.00213†	0.00478 ± 0.00213†
Predictive accuracy			
	MAE	0.107	0.107
	RMSE	0.138	0.138
	ICC (95% CI)	0.471 (0.365-0.566)	0.468 (0.361-0.562)
	MD (LOA)	0.015 (−0.255 to 0.285)	0.015 (−0.255 to 0.284)
Predictive index			
	Mean (SD)	0.616 (0.096)	0.615 (0.095)
	Range	0.350-0.888	0.35-0.88
Tobit		L + age + sex + BMI	L + age + BMI
Parameters, coefficient ± SE			
Intercept		0.880 ± 0.074*	0.888 ± 0.074*
		−2.020 ± 0.0332*	−2.017 ± 0.0332*
L		−0.0252 ± 0.00178*	−0.0254 ± 0.00178*
Age		−0.00162 ± 0.000740†	−0.00174 ± 0.00074†
Sex		0.0194 ± 0.0133	/
BMI		0.00451 ± 0.00213†	0.00476 ± 0.00213†
Predictive accuracy			
	MAE	0.107	0.107
	RMSE	0.138	0.138
	ICC (95% CI) MD (LOA)	0.473 (0.367-0.567) 0.015 (−0.254 to 0.285)	0.469 (0.363-0.564) 0.015 (−0.255 to 0.285)
Predictive index			
	Mean (SD)	0.616 (0.0966)	0.616 (0.095)
	Range	0.348−0.891	0.351-0.880
Beta		L + age + sex + BMI	L + age + BMI
Parameters, coefficient ± SE			
Intercept		1.692 ± 0.337*	1.730 ± 0.336*
L		−0.114 ± 0.0082*	−0.115 ± 0.0082*
Age		−0.0063 ± 0.0034	−0.00676 ± 0.00334†
Sex		0.0936 ± 0.0604	/
BMI		0.0165 ± 0.0096	0.0178 ± 0.0096
Predictive accuracy			
	MAE	0.111	0.110
	RMSE	0.144	0.143
	ICC (95% CI) MD (LOA)	0.536 (0.438-0.622) 0.020 (−0.259 to 0.299)	0.536 (0.437-0.622) 0.020 (−0.258 to 0.298)
Predictive index			

continued on next page

Table 3. Continued

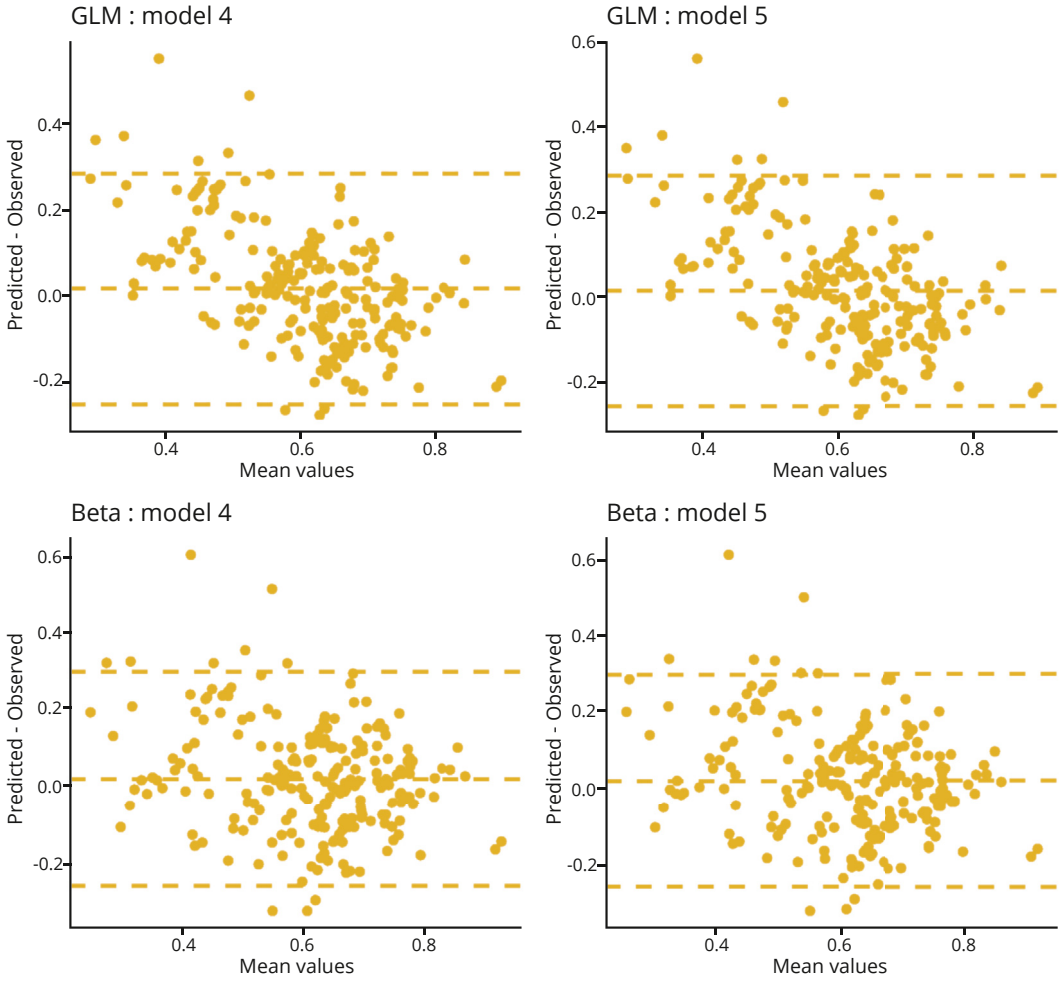
Variables		Model 4	Model 5 Model 4 with $P < .10$
	Mean (SD)	0.621 (0.132)	0.620 (0.10)
	Range	0.244-0.907	0.247-0.899

BMI indicates body mass index; GLM, generalized linear model; ICC, intraclass correlation coefficient; L, Lequesne score; LOA, limits of agreement; MAE, mean absolute error; MD, mean differences; RMSE, root mean squared error; SE, standard error.
*$P < .001$.
†$P < .05$.

there have also been cases where the Oxford index has been used for this purpose.[5-10] Despite the challenges inherent in comparing these studies due to statistical variations, the performance of our mapping model using the Lequesne index showed notable similarities to previous mapping models. For example, our model showed MAEs in the range of what has been observed from previous mapping efforts in OA with value ranging from 0.001 to 0.20.[37]

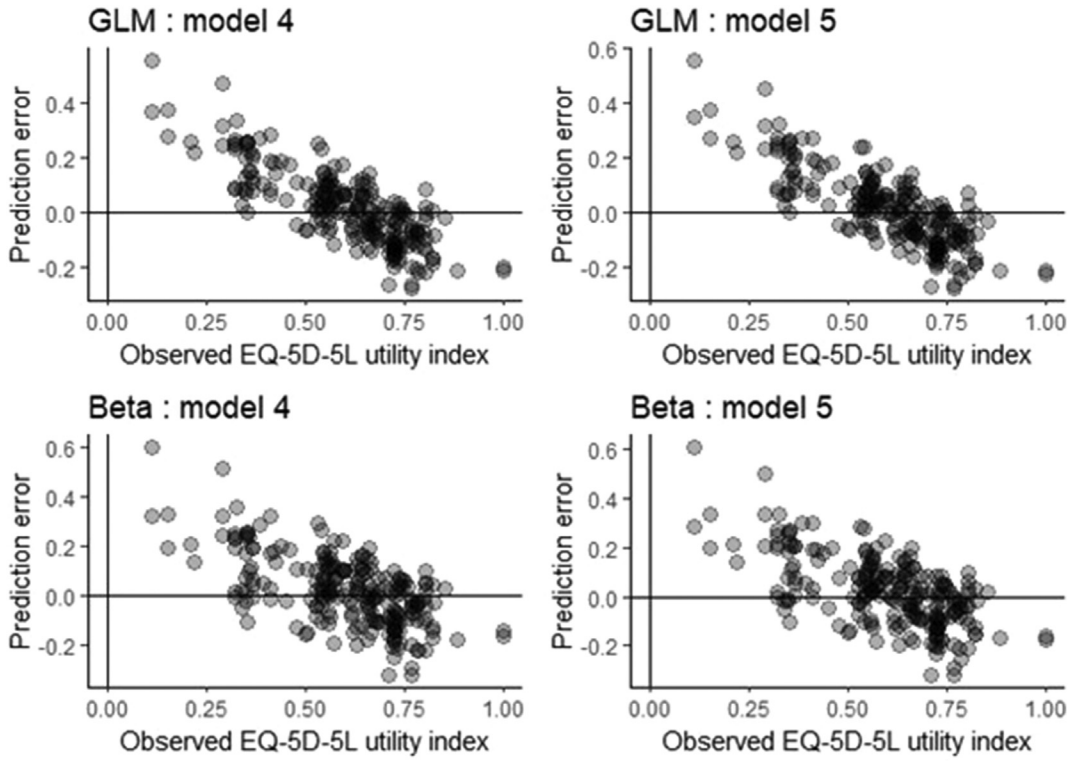
Within the framework of our study, which included the evaluation of 5 different models, a notable observation emerged with regard to the models that included age, sex, and/or BMI as explanatory variables whatever the type of regression, GLM, tobit, or beta regression. Remarkably, these specific models showed good fit and were consequently selected as the preferred mapping models. This finding highlights the importance of incorporating sociodemographic data into such models. The inclusion of age, sex, and BMI as contributing factors not only slightly improved the accuracy and fit of the models but also highlighted the crucial role that sociodemographic variables may play in refining and improving the predictive ability of mapping

Figure 1. Bland-Altman plots for predicted and observed utility values based on the Lequesne score.



GLM indicates generalized linear model.

Figure 2. Prediction errors by observed EQ-5D-5L utility index of models 4 and 5 for GLM and Beta regression in the validation dataset.



algorithms. This emphasizes the importance of considering a wider range of patient characteristics, beyond purely disease-specific measures, when constructing effective mapping models.[38]

Regarding the statistical methodology used in our study, both types of regression, GLM and tobit regression, exhibited comparable MAE and RMSE, although slightly better values were observed than beta regression. However, the beta regression models showed wider ranges, more aligned with the validation data set. In addition, the ICC values favored the beta regression models with value > 0.5 that indicated moderate reliability. In a WOMAC mapping study using the EQ-5D-5L, the tobit and beta models were not better than the GLMs, like in our study, confirming that the GLM models can be appropriate at least when the 5-item EQ-5D is used.[21] In our study, despite disparities in predictions and coefficients across data sets, these differences are unlikely to be substantial to affect the outcomes of an economic evaluation. Indeed, our models provided precise predictions in both internal and external validation data sets, indicating its potential to perform effectively in similar populations. At last, an advantage of our study was the investigation of the appropriate powers for increasing the Lequesne index using GAMs. This approach was crucial given that it confirmed the linear relationship between the response and Lequesne index, thereby increasing the accuracy of our analysis.

The identified preferred mapping models selected during the validation analysis consistently showed a tendency to overpredict lower EQ-5D-5L scores and underpredict better health status. This striking systematic bias may be due to the use of the 5-item EQ-5D scale and the reliance on linear regression methods, even if for the latter it is has been investigated through GAM. Recognizing this inherent limitation, alternative methods, particularly Bayesian methods, emerge as potential remedies to mitigate these challenges.[5]

This study has some limitations. First, our study population coming from an RCT may not be fully representative of all patients with knee OA and caution should be exercised when applying this mapping model to estimate utilities of other patient populations. Second, the Lequesne index and EQ-5D-5L measures do not match perfectly, which can lead to increased uncertainty when using the mapping function to estimate utilities. Therefore, it is important to note that this mapping function cannot fully replace the EQ-5D-5L in economic evaluations. Third, we did not use the best-validated sample because the training and validation data sets were from the same RCT and were quite homogeneous. It would be advisable to use an independent data set to assess the external validity of a mapping model, but unfortunately we do not have access to such a database. Finally, although this mapping model is applicable to both individual- and group-level predictions, it is recommended that it be used at the group level. This preference stems from the fact that, in economic evaluations, effectiveness is typically compared at the group level.

In conclusion, to the best of our knowledge, the mapping functions we have developed represent pioneering efforts in the conversion of the Lequesne index to EQ-5D-5L scores in people diagnosed of knee OA. These established functions have demonstrated a sufficient level of accuracy and precision, providing resources for both practitioners and researchers. They serve as key tools, particularly in scenarios requiring cost-effectiveness assessments, where access to generic preference-based tools for deriving HRQoL utilities may be limited. By bridging this gap, these functions provide a feasible and reliable means of estimating patient-reported outcomes, thereby contributing significantly to informed decision making in clinical practice and health services research in the area of knee OA.

ARTICLE IN PRESS

Author Disclosures

Author disclosure forms can be accessed below in the Supplemental Material section.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2024.06.017.

Article and Author Information

Accepted for Publication: June 21, 2024

Published Online: xxxx

doi: https://doi.org/10.1016/j.jval.2024.06.017

Author Affiliations: Biostatistics Center (B-STAT), University Hospital and University of Liège, Liège, Belgium (Dardenne, Donneau); WHO Collaborating Center for Public Health Aspects of Musculoskeletal Health and Ageing, Research Unit in Public Health, Epidemiology and Health Economics, and Department of Sport and Rehabilitation Sciences, University of Liège, Liège, Belgium (Bruyère).

Correspondence: Olivier Bruyère, PhD, Research Unit in Public Health, Epidemiology and Health Economics, University of Liège, Liège, Belgium. Email: olivier.bruyere@uliege.be

Author Contributions: *Concept and design:* Dardenne, Bruyère
Acquisition of data: Bruyère
Analysis and interpretation of data: Dardenne, Donneau, Bruyère
Drafting of the article: Dardenne, Bruyère
Critical revision: Dardenne, Donneau, Bruyère
Obtaining funding: Bruyère
Administrative, technical, or logistic support: Dardenne, Donneau, Bruyère
Supervision: Donneau, Bruyère

Funding/Support: This study was supported by a research grant from IBSA (grant QMS-SOP008).

Role of the Funder/Sponsor: IBSA has no role in the design of the study.

Acknowledgment: We thank IBSA for providing access to their database.

REFERENCES

1. Collaborators GBDO. Global, regional, and national burden of osteoarthritis, 1990-2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Rheumatol*. 2023;5(9):e508–e522.
2. Hiligsmann M, Cooper C, Arden N, et al. Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). *Semin Arthritis Rheum*. 2013;43(3):303–313.
3. Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and healthy-years equivalents. *Med Decis Mak*. 1989;9(2):142–149.
4. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, present and future. *Appl Health Econ Health Policy*. 2017;15(2):127–137.
5. Ayala A, Forjaz MJ, Ramallo-Farina Y, Martin-Fernandez J, Garcia-Perez L, Bilbao A. Response mapping methods to estimate the EQ-5D-5L from the western Ontario McMaster universities osteoarthritis in patients with hip or knee osteoarthritis. *Value Health*. 2021;24(6):874–883.
6. Dakin H, Gray A, Murray D. Mapping analyses to estimate EQ-5D utilities and responses based on Oxford Knee Score. *Qual Life Res*. 2013;22(3):683–694.
7. Fawaz H, Yassine O, Hammad A, Bedwani R, Abu-Sheasha G. Mapping of disease-specific Oxford Knee Score onto EQ-5D-5L utility index in knee osteoarthritis. *J Orthop Surg Res*. 2023;18(1):84.
8. Martin-Fernandez J, Morey-Montalvo M, Tomas-Garcia N, et al. Mapping analysis to predict EQ-5D-5 L utility values based on the Oxford Hip Score (OHS) and Oxford Knee Score (OKS) questionnaires in the Spanish population suffering from lower limb osteoarthritis. *Health Qual Life Outcomes*. 2020;18(1):184.
9. Wailoo A, Hernandez Alava M, Escobar Martinez A. Modelling the relationship between the WOMAC osteoarthritis Index and EQ-5D. *Health Qual Life Outcomes*. 2014;12:37.
10. Xie F, Pullenayegum EM, Li SC, Hopkins R, Thumboo J, Lo NN. Use of a disease-specific instrument in economic evaluations: mapping WOMAC onto the EQ-5D utility index. *Value Health*. 2010;13(8):873–878.
11. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation–value in comparison with other assessment tests. *Scand J Rheumatol Suppl*. 1987;65:85–89.
12. Bilbao A, Garcia-Perez L, Arenaza JC, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual Life Res*. 2018;27(11):2897–2908.
13. Feng YS, Kohlmann T, Janssen MF, Buchholz I. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. *Qual Life Res*. 2021;30(3):647–673.
14. Migliore A, Blicharski T, Plebanski R, et al. Knee osteoarthritis pain management with an innovative high and low molecular weight hyaluronic acid formulation (HA-HL): A randomized clinical trial. *Rheumatol Ther*. 2021;8(4):1617–1636.
15. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–1736.
16. Jensen CE, Sørensen SS, Gudex C, Jensen MB, Pedersen KM, Ehlers LH. The Danish EQ-5D-5L value set: A hybrid model using cTTO and DCE data. *Appl Health Econ Health Policy*. 2021;19(4):579–591.
17. Andrade LF, Ludwig K, Goni JMR, Oppe M, de Pouvourville G. A French value set for the EQ-5D-5L. *Pharmacoeconomics*. 2020;38(4):413–425.
18. Ludwig K, Graf von der Schulenburg JM, Greiner W. German value set for the EQ-5D-5L. *Pharmacoeconomics*. 2018;36(6):663–674.
19. MAAE S, de Wit GA, Prenger R, E, Dutch AS, E A, StolkM MV. K MV Tariff for the Five-Level Version of EQ-5D. *Value Health*. 2016;19(4):343–352.
20. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ*. 2018;27(1):7–22.
21. Bilbao A, Martin-Fernandez J, Garcia-Perez L, et al. Mapping WOMAC onto the EQ-5D-5L utility index in patients with hip or knee osteoarthritis. *Value Health*. 2020;23(3):379–387.
22. Arostegui I, Nunez-Anton V, Quintana JM. Statistical approaches to analyse patient-reported outcomes as response variables: an application to health-related quality of life. *Stat Methods Med Res*. 2012;21(2):189–214.
23. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. *Qual Life Res*. 2000;9(8):901–910.
24. Austin PC. A comparison of methods for analyzing health-related quality-of-life measures. *Value Health*. 2002;5(4):329–337.
25. Hunger M, Baumert J, Holle R. Analysis of SF-6D index data: is beta regression appropriate? *Value Health*. 2011;14(5):759–767.
26. Silvia Ferrari FC-N, Cribari-Neto F. Beta regression for modelling rates and proportions. *J Appl Stat*. 2004;31(7):799–815.
27. Hastie T, Tibshirani R. Generalized additive models for medical research. *Stat Methods Med Res*. 1995;4(3):187–196.
28. Costa D, Cruz EB, Silva C, et al. Factors associated with clinical and radiographic severity in people with osteoarthritis: A cross-sectional population-based study. *Front Med (Lausanne)*. 2021;8:773417.
29. Ameri H, Yousefi M, Yaseri M, Nahvijou A, Arab M, Akbari Sari A. Mapping the cancer-specific QLQ-C30 onto the generic EQ-5D-5L and SF-6D in colorectal cancer patients. *Expert Rev Pharmacoecon Outcomes Res*. 2019;19(1):89–96.
30. Dai Z, He Q, Zhu L, et al. Automatic prediction model for online diaphragm motion tracking based on optical surface monitoring by machine learning. *Quant Imaging Med Surg*. 2023;13(4):2065–2080.
31. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30–46.
32. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
33. TW Y. *Vector Generalized Linear Addit Models*. 2023:1–589.
34. YY. *Machine Learning Evaluation Metrics*; 2022.
35. MG. *Various Coefficients of Interrater Reliability and Agreement*. 2019.
36. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health Qual Life Outcomes*. 2018;16(1):31.
37. 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):215–225.
38. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes*. 2013;11:151.