


# Converting Parkinson-Specific Scores into Health State Utilities to Assess Cost-Utility Analysis

Gang Chen<sup>1</sup> · Miguel A. Garcia-Gordillo<sup>2,3</sup>  · Daniel Collado-Mateo<sup>4,5</sup> · Borja del Pozo-Cruz<sup>6</sup> · José C. Adsuar<sup>4</sup> · José Manuel Cordero-Ferrera<sup>7</sup> · José María Abellán-Perpiñán<sup>3</sup> · Fernando Ignacio Sánchez-Martínez<sup>3</sup>

Published online: 7 June 2018  
© Springer International Publishing AG, part of Springer Nature 2018

## Abstract

**Objectives** The aim of this study was to compare the Parkinson's Disease Questionnaire-8 (PDQ-8) with three multi-attribute utility (MAU) instruments (EQ-5D-3L, EQ-5D-5L, and 15D) and to develop mapping algorithms that could be used to transform PDQ-8 scores into MAU scores. **Methods** A cross-sectional study was conducted. A final sample of 228 evaluable patients was included in the analyses. Sociodemographic and clinical data were also collected. Two EQ-5D questionnaires were scored using Spanish tariffs. Two models and three statistical techniques were used to estimate each model in the direct mapping framework for all three MAU instruments, including the most widely used ordinary least squares (OLS), the robust

MM-estimator, and the generalized linear model (GLM). For both EQ-5D-3L and EQ-5D-5L, indirect response mapping based on an ordered logit model was also conducted. Three goodness-of-fit tests were employed to compare the models: the mean absolute error (MAE), the root-mean-square error (RMSE), and the intra-class correlation coefficient (ICC) between the predicted and observed utilities.

**Results** Health state utility scores ranged from 0.61 (EQ-5D-3L) to 0.74 (15D). The mean PDQ-8 score was 27.51. The correlation between overall PDQ-8 score and each MAU instrument ranged from  $-0.729$  (EQ-5D-5L) to  $-0.752$  (EQ-5D-3L). A mapping algorithm based on PDQ-8 items had better performance than using the overall score. For the two EQ-5D questionnaires, in general, the indirect mapping approach had comparable or even better performance than direct mapping based on MAE.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s40271-018-0317-5>) contains supplementary material, which is available to authorized users.

✉ Miguel A. Garcia-Gordillo  
miguelgarciaordillo@gmail.com

Gang Chen  
gang.chen@flinders.edu.au

Daniel Collado-Mateo  
danicolladom@gmail.com

Borja del Pozo-Cruz  
b.delpozocruz@auckland.ac.nz

José C. Adsuar  
carmelo.adsuar@gmail.com

José Manuel Cordero-Ferrera  
jmcordero@unex.es

José María Abellán-Perpiñán  
dionisos@um.es

Fernando Ignacio Sánchez-Martínez  
fernando@um.es

<sup>1</sup> College of Medicine and Public Health, Flinders University, Adelaide, Australia

<sup>2</sup> Facultad de Administración y Negocios, Universidad Autónoma de Chile, sede Talca, Chile

<sup>3</sup> Department of Applied Economics, Faculty of Economics and Business, University of Murcia, Murcia, Spain

<sup>4</sup> Faculty of Sport Sciences, University of Extremadura, Cáceres, Spain

<sup>5</sup> Instituto de Calidad de Vida, Actividad Física y Salud, Universidad Autónoma de Chile, sede Talca, Chile

<sup>6</sup> Department of Sport and Exercise Science, University of Auckland, Auckland, New Zealand

<sup>7</sup> Department of Economics, Faculty of Economics and Business, University of Extremadura, Badajoz, Spain

**Conclusions** Mapping algorithms developed in this study enable the estimation of utility values from the PDQ-8. The indirect mapping equations reported for two EQ-5D questionnaires will further facilitate the calculation of EQ-5D utility scores using other country-specific tariffs.

### Key Points

Regulatory authorities generally recommend the use of cost-utility analysis, which implies the need to use health state utility as the outcome measure.

The current study aims to enable cost-utility analysis using data from a widely used Parkinson-disease-specific questionnaire.

Additionally, the current study is the first that attempts to predict the utilities of 15D and EQ-5D-5L based on PDQ-8 dimensions and total scores of patients with Parkinson's disease.

## 1 Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder around the world. Despite much research on this disease, its specific cause remains unknown. Its diagnosis is subject to a combination of bradykinesia with at least one of the remaining clinical manifestations: muscular rigidity, tremor or postural instability [1, 2]. The multiple motor and non-motor symptoms of such a severe chronic condition as PD have a strong impact on health-related quality of life (HRQoL) and patient well-being. A gradual deterioration of quality of life (QoL) involves a significant financial burden for the public healthcare system [3, 4]. Health programs and treatments aimed to alleviate the symptoms of PD patients may reduce HRQoL losses but also impose costs on the healthcare budget. Health economic evaluation can then become an important tool for decision making by healthcare managers. By comparing the costs and results of health interventions, one can assess if emerging technologies or even those currently used are 'value for money.'

The economic evaluation may not end all issues regarding the allocation of scarce resources across healthcare interventions, but it can help to reduce political arbitrariness in making such decisions. There exist different economic evaluation techniques, but the one recommended by the National Institute for Health and Care Excellence (NICE) in the UK and other health regulatory agencies around the world is cost-utility analysis (CUA) [5]. This method may be regarded as a particular case of cost-effectiveness analysis,

which represents the outcome (i.e., effectiveness) in quality-adjusted life-years (QALYs). A QALY includes the two main dimensions of health outcomes, namely, QoL (morbidity) and life-years (mortality) in a single index [6]. In essence, QALYs are obtained by weighting life-years with a QoL index, ranging from 0 (death) to 1 (full health).

QoL weights are also known as 'utilities' (hence the term cost-utility analysis) and are designed to reflect preferences of the general population about the values of different health states. There are two ways to obtain these utilities: via direct preference valuation techniques, such as time trade-off (TTO) and standard gamble (SG), or indirectly via HRQoL measurement instruments, also known as multi-attribute utility (MAU) instruments [7]. These MAU instruments include a multi-attribute descriptive system and an algorithm that converts the description of a health state into a single QoL index. Among existing MAU instruments, the EQ-5D-3L has been the most widely used [8]. For all of them, a value set or 'tariff' of utilities is available, and in some cases, there are several country-specific 'tariffs.'

Some disease-specific instruments are often preferred to generic instruments because of their higher sensitivity to changes of HRQoL and greater ability to detect clinically important differences in a particular condition. That is why these instruments are commonly used in the context of clinical trials [9–11]. However, most of these disease-specific measures do not have weights based on peoples' preferences, and do not have a scoring algorithm that allows utility indexes to be obtained. A major advantage of generic preference-based HRQoL instruments is that they generate utilities that can be used to compare QALYs gained for interventions across patient groups and diseases to aid broad health-service resource allocation decisions.

Since disease-specific instruments may be superior in terms of sensitivity and are also widely used in clinical practice, it would be useful to find a way to assign utilities to condition-specific measures. One way consists of using mapping algorithms, which often include demographic characteristics, in particular age and gender, to improve the predictive performance of the models [12].

The main purpose of this study is to develop a mapping algorithm using HRQoL data from the Parkinson's Disease Questionnaire-8 (PDQ-8) to impute utility values into three MAU instruments selected in this study (EQ-5D-3L, EQ-5D-5L, and 15D).

## 2 Methods

### 2.1 Study Sample and Data Collection

A cross-sectional study was conducted. Data were obtained between May 2012 and August 2013. Patients were

recruited from 23 local PD associations, belonging to 13 out of the 17 different regions of Spain. All participants in this study were over 30 years of age (range 34–86 years) and had been diagnosed with PD. Exclusion criteria included patients with the highest degree of severity possible (Hoehn and Yahr stage V) because of a lack of autonomy in completing the questionnaire set by themselves. Likewise, those diagnosed with other similar disorders were excluded, such as other subtypes of Parkinsonism like vascular cause, drug-induced, dementia with parkinsonism, multiple systems atrophy (MSA) [13]. A total of 280 patients who were registered in the study received a set of questionnaires and detailed information about the aims and procedures of the study. They were included after giving their written informed consent. However, 52 patients were subsequently excluded because of missing data on either the MAU instruments or the PDQ-8. A final sample of 228 evaluable patients was included in the analyses. They answered the set of questionnaires under the supervision of trained interviewers. The study was approved by the Ethics Committee of the University of Extremadura and was developed following the ethical guidelines of the Declaration of Helsinki as revised in Seoul in October 2008.

## 2.2 Health-Related Quality of Life and Clinical Measures

The EQ-5D-3L [14] is one of the most widely used MAU instruments to assess HRQoL. It consists of five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and three response levels for each dimension. A new five-level version EQ-5D (EQ-5D-5L) was developed to increase the sensitivity of the original three-level version [15]. This new questionnaire maintains the five dimensions from its original version but includes five response levels for each dimension. Both the EQ-5D-3L and EQ-5D-5L questionnaires were scored using the official Spanish tariffs [16, 17].

The 15D is a MAU instrument that was originally developed in Finland [18]. It has 15 dimensions: mobility, vision, hearing, breathing, sleeping, eating, speech, excretion, usual activities, mental function, discomfort and symptoms, depression, distress, vitality, and sexual activity. However, in comparison, the EQ-5D-5L seems to have better efficiency and greater sensitivity in detecting clinical HRQoL changes than the 15D [19]. The valuation system of the 15D is based on an application of the MAU theory and derived by using a combination of rating scale and magnitude estimation methods from a Finnish population sample [18].

The PDQ-8 is a clinical measure, which was developed from the PDQ-39 [20]. The PDQ-8 requires less time for self-completion than the PDQ-39. The PDQ-8 is a short-

form questionnaire, made up of eight items, representing different dimensions: mobility, activities of daily living, emotional well-being, social support, cognition, communication, bodily discomfort, and stigma. Each item is scored on an ordinal scale from 0 (never or not at all) to 4 (always or cannot do at all). The summary index of all items is standardized on a scale with a range of 0–100, where the lower index represents better HRQoL.

In all cases, Spanish versions of the instruments were used in the questionnaires. A total of 35 min was the mean duration of the interviews.

## 2.3 Statistical Analyses

### 2.3.1 Descriptive Statistics

The Shapiro–Wilk test was used to determine whether QoL scores were normally distributed. The degree of correlation between the health state utility and PDQ-8 (item and overall) scores was calculated using the Spearman rank correlation.

### 2.3.2 Mapping Analyses

A direct mapping approach using regression techniques was first used to develop a series of mapping algorithms from the PDQ-8 onto each of the three MAU instruments. In each regression, the dependent (left hand) variable was a utility score calculated according to the choice of MAU instrument. Two sets of independent (right hand) variables were considered. Model I used the overall PDQ-8 score, its square term (to detect non-linear effects), age, and gender (a dummy variable). The independent variables of Model II replaced the overall PDQ-8 score (and its squared term) with the item scores (and corresponding squared terms) of the PDQ-8. Secondly, for both EQ-5D-3L and EQ-5D-5L, an indirect response mapping was also conducted. In the indirect mapping, the responses of each EQ-5D dimension were predicted separately from PDQ-8 using the ordered logit model. Then, applying country-specific tariffs, the country-specific EQ-5D utilities can be calculated. This indirect mapping approach provides an advantage of producing a set of mapping algorithms that can potentially be used for other country-specific tariffs, not just the Spanish tariff that was used for this study. This indirect mapping was not applied to the 15D, which contains broader dimensions not included in the PDQ-8, since large prediction errors would be expected.

Three statistical techniques were used to estimate each model in the direct mapping approach, including the most widely used ordinary least squares (OLS), the robust MM-estimator (which is designed to deal with the presence of potential outliers) [21, 22] and the generalized linear model

(GLM) (which allows for the non-normal distribution of dependent variables) [23]. For the GLM estimator, among different combinations of commonly used family (Gaussian and gamma) and link (log) functions in the crosswalk literature, the optimal combination was chosen (and reported in Table 3) for each mapping function using the goodness-of-fit criteria described below (detailed results not reported but available from the authors). The ordered logit model was used to estimate the EQ-5D response in the indirect mapping [24]. Except for the constant, independent variables were retained only when they were statistically significant (i.e.,  $p < 0.05$ ), based on a forward stepwise regression technique [22, 25]. For logical reasons, the squared terms of the overall score of PDQ-8 and the square of the item scores were only included when the linear terms were significant.

### 2.3.3 Validation and Goodness-of-Fit

Without an external database, this study adopts a commonly applied internal validation ‘hold-out’ approach [22, 25]. The full data were randomly divided into two mutually exclusive groups using Stata’s random-number generator: 75% of the full sample was used as an ‘estimation’ sample, whilst the remaining 25% served as a ‘validation’ sample.

Three goodness-of-fit tests were employed to compare models, the mean absolute error (MAE), the root-mean-square error (RMSE), and the intra-class correlation coefficient (ICC) between the predicted and observed utilities. The ICC, which shows the agreement between the absolute magnitude of two variables, was calculated using a two-way mixed model [26]. Since there is no consensus on the choice of criteria, the optimal regression technique was selected, which achieved the best result with respect to the majority of the above three criteria. For the OLS estimates, an additional popular goodness-of-fit measure,  $R^2$ , is also reported (although it was not used as an evaluation criterion since this statistic is not available for all econometric methods). The optimal method for estimating each mapping function was identified through evaluating goodness-of-fit measures calculated on the 25% validation sample. The ICC was calculated in SPSS version 21.0 (IBM Corp, Armonk, NY, USA). Other analyses were conducted in Stata version 14.1 (StataCorp LP, College Station, Texas, USA).

## 3 Results

### 3.1 Descriptive Statistics

Summary statistics for patients’ characteristics are reported in Table 1. The mean (SD) age was 66.2 (9.5) years old,

with the majority of patients being male (66.2%). Health state utility scores assessed using the MAU instruments ranged from 0.61 (EQ-5D-3L) to 0.74 (15D). The mean (SD) PDQ-8 score was 27.51 (18.65). The distributions of the patients’ responses in the three MAU instruments are presented in the Appendix (see electronic supplementary material) and, as can be seen, health state utilities are not normally distributed. There were no statistically significant differences in patients’ characteristics between the estimation and validation samples.

The bivariate correlation between the PDQ-8 domains and each MAU instrument is reported in Table 2. Generally, it can be seen that the lowest correlations (ranged from insignificant  $-0.154$  to  $-0.208$ ,  $p > 0.05$ ) were found between the social support domain (of PDQ-8) and all the MAU instruments. The strongest correlations varied in each MAU instrument: between 15D and the ‘communication’ domain ( $r = 0.578$ ), and between EQ-5D-3L/EQ-5D-5L and the ‘activities of daily living’ domain ( $r = -0.671/-0.618$ ). The correlations between the overall PDQ-8 score and each MAU instrument ranged from  $-0.729$  (EQ-5D-5L) to  $-0.752$  (EQ-5D-3L).

### 3.2 Mapping Results

Goodness-of-fit statistics for each of the mapping algorithms are reported in Table 3 for both the estimation sample (Panel A) and the validation sample (Panel B). The better goodness-of-fit is indicated by lower MAE/RMSE and higher ICC. Focusing on the validation sample, it can be seen that firstly, in the majority of cases, mapping algorithms performed better using the PDQ-8 item scores than using the overall score. Secondly, the optimal statistical method identified (bold in table) varied in each mapping function. Among the six mapping functions, the direct mapping framework, the robust MM-estimator, was chosen four times, followed by the OLS (twice). Among all mapping equations, better goodness-of-fit was found when mapping PDQ-8 onto 15D, judging by the ICC and supported by the scatter plots between observed and predicted utilities (Fig. 1). In addition, goodness-of-fit results in the validation sample are not necessarily worse than the estimation sample. In fact, for 15D, all three indicators reported in the validation sample are better than their counterparts in the estimation sample.

Regarding the indirect response mapping, results show that the performance was comparable or even better than the direct mapping approach based on MAE but was slightly worse based on RMSE/ICC. For example, the optimal MAE of mapping PDQ-8 item scores onto EQ-5D-3L was 0.1681 based on the MM-estimator, whilst the indirect response mapping had an MAE of 0.1606. Similarly, for the EQ-5D-5L scenario, the MAE of indirect

**Table 1** Patient characteristics

	Full simple N = 228	Estimation simple N = 172	Validation simple N = 56	p value*
Age, mean (SD)	66.2 (9.5)	65.8 (9.8)	67.3 (8.4)	0.32
Gender (%)				0.20
Female	33.77	36.05	26.79	
Male	66.23	63.95	73.21	
Comorbidity (%)				0.24
Only Parkinson's disease	54.80	50.60	67.90	
Besides Parkinson's disease	45.20	49.40	32.10	
Diagnostic date (%)				0.22
1970–2000	26.80	25.20	31.50	
2001–2004	15.00	17.60	7.40	
2005–2008	25.80	23.90	31.50	
2009–2012	32.40	33.30	29.60	
Hoehn and Yahr stage (%)				0.28
Stages I–II (mild to moderate)	68.40	70.2	62.50	
Stages III–IV (severe)	31.60	29.8	37.50	
Quality of life, mean (SD)				
15D	0.74 (0.16)	0.75 (0.15)	0.72 (0.17)	0.20
EQ-5D-3L	0.61 (0.33)	0.63 (0.31)	0.54 (0.39)	0.26
EQ-5D-5L	0.63 (0.23)	0.64 (0.22)	0.60 (0.26)	0.33
Parkinson's Disease Questionnaire (PDQ)	27.51 (18.65)	26.44 (18.04)	30.80 (20.21)	0.16

H&YS Hoehn and Yahr Stages

\*Chi-squared test used for categorical data, and the Wilcoxon rank-sum test used for age and quality of life outcomes

**Table 2** Correlation coefficients of Parkinson's Disease Questionnaire (PDQ-8) and multi-attribute utilities

	15D	EQ-5D-3L	EQ-5D-5L
PDQ1 (mobility)	– 0.523	– 0.565	– 0.549
PDQ2 (activities of daily living)	– 0.568	– 0.671	– 0.618
PDQ3 (emotional well-being)	– 0.559	– 0.535	– 0.540
PDQ4 (social support)	– <i>0.154</i>	– <i>0.208</i>	– <i>0.203</i>
PDQ5 (cognition)	– 0.437	– 0.447	– 0.429
PDQ6 (communication)	– 0.578	– 0.472	– 0.481
PDQ7 (bodily discomfort)	– 0.488	– 0.409	– 0.378
PDQ8 (stigma)	– 0.466	– 0.443	– 0.461
PDQ-8 overall score	– 0.750	– 0.752	– 0.729

Except for three coefficients in italics, all other correlation coefficients were statistically significant (Sidak-adjusted significance level,  $p < 0.05$ )

response mapping was 0.1063, slightly worse than the best MAE value identified from direct mapping of 0.1054 (MM-estimator), but much better than the other two methods in which both MAE values were larger than 0.11. Goodness-of-fit statistics of the estimation sample are largely comparable with the validation sample (details available in Table 3).

Mapping equations corresponding to the optimal methods identified above are reported in Table 4 (direct mapping) and Table 5 (indirect response mapping), respectively. In direct mapping, none of the personal characteristics were significant. They were not included in the final models, whilst in indirect response mapping, age and gender were found to be significant in some equations. For all mapping equations, social support and cognition domains of PDQ-8 were found to be consistently insignificant. All other domains were significant in at least one mapping function.

## 4 Discussion

The present study is the first that attempts to predict the utilities from 15D and EQ-5D-5L based on PDQ-8 dimensions and total scores of patients with PD. There are a few published studies that developed mapping algorithms onto EQ-5D-3L in PD [25, 27]. Between all generic MAU instruments, the EQ-5D-3L is the most used measure to predict utilities in this population and other diseases, since this is the MAU instrument preferred by NICE in the UK [28]. On the other hand, both PDQ-8 and PDQ-39 have

**Table 3** Goodness-of-fit results for transformation from PDQ-8 scores to multi-attribute utilities

Overall score					Item scores				
Method	MAE	RMSE	ICC	R <sup>2</sup>	Method	MAE	RMSE	ICC	R <sup>2</sup>
Panel A—estimation sample, N = 172									
(1) 15D									
OLS	0.0800	<b>0.1044</b>	0.815	0.522	<b>OLS</b>	0.0730	<b>0.0956</b>	<b>0.858</b>	0.600
MM	<b>0.0792</b>	0.1052	<b>0.819</b>		MM	<b>0.0713</b>	0.0989	0.856	
GLM <sup>a</sup>	0.0800	0.1048	0.811		GLM <sup>b</sup>	0.0733	0.0965	0.857	
(2) EQ-5D-3L									
OLS	0.1668	<b>0.2234</b>	<b>0.791</b>	0.485	<b>OLS</b>	0.1405	<b>0.1922</b>	<b>0.867</b>	0.619
MM	<b>0.1621</b>	0.2394	0.715		MM	<b>0.1380</b>	0.1986	0.854	
GLM <sup>a</sup>	0.1721	0.2293	0.763		GLM <sup>a</sup>	0.1478	0.2017	0.839	
					OLOGIT	0.1502	0.2384	0.761	
(3) EQ-5D-5L									
OLS	0.1177	<b>0.1546</b>	0.795	0.491	<b>OLS</b>	0.1021	<b>0.1395</b>	<b>0.850</b>	0.585
MM	<b>0.1165</b>	0.1557	<b>0.804</b>		MM	0.1069	0.1478	0.834	
GLM <sup>b</sup>	0.1172	0.1565	0.799		GLM <sup>a</sup>	<b>0.1017</b>	0.1406	0.845	
					OLOGIT	0.1079	0.1506	0.837	
Panel B—Validation sample, N = 56									
(1) 15D									
OLS	0.0717	<b>0.0951</b>	0.883		OLS	0.0723	0.0964	0.878	
MM	<b>0.0707</b>	0.0959	<b>0.886</b>		<b>MM</b>	<b>0.0641</b>	<b>0.0877</b>	<b>0.910</b>	
GLM <sup>a</sup>	0.0736	0.0953	0.878		GLM <sup>b</sup>	0.0736	0.0965	0.878	
(2) EQ-5D-3L									
OLS	<b>0.1760</b>	<b>0.2260</b>	<b>0.861</b>		<b>OLS</b>	0.1801	<b>0.2188</b>	<b>0.886</b>	
MM	0.1988	0.2738	0.762		MM	<b>0.1681</b>	0.2194	0.884	
GLM <sup>a</sup>	0.1869	0.2484	0.811		GLM <sup>a</sup>	0.1952	0.2421	0.841	
					OLOGIT	0.1606	0.2465	0.838	
(3) EQ-5D-5L									
OLS	0.1180	0.1562	0.853		OLS	0.1139	0.1430	0.892	
MM	<b>0.1163</b>	<b>0.1550</b>	<b>0.863</b>		<b>MM</b>	<b>0.1054</b>	<b>0.1376</b>	<b>0.903</b>	
GLM <sup>b</sup>	0.1179	0.1540	0.857		GLM <sup>a</sup>	0.1127	0.1422	0.888	
					OLOGIT	0.1063	0.1519	0.887	

Best results among the three direct mapping methods are in bold type

GLM generalized linear model, ICC intra-class correlation coefficient, MAE mean absolute error, MM robust MM-estimator, OLOGIT ordered logit (indirect response mapping), OLS ordinary least squares, PDQ Parkinson's Disease Questionnaire, RMSE root mean squared error, R<sup>2</sup> R squared

<sup>a</sup>The *Gaussian* family with *log* link was used

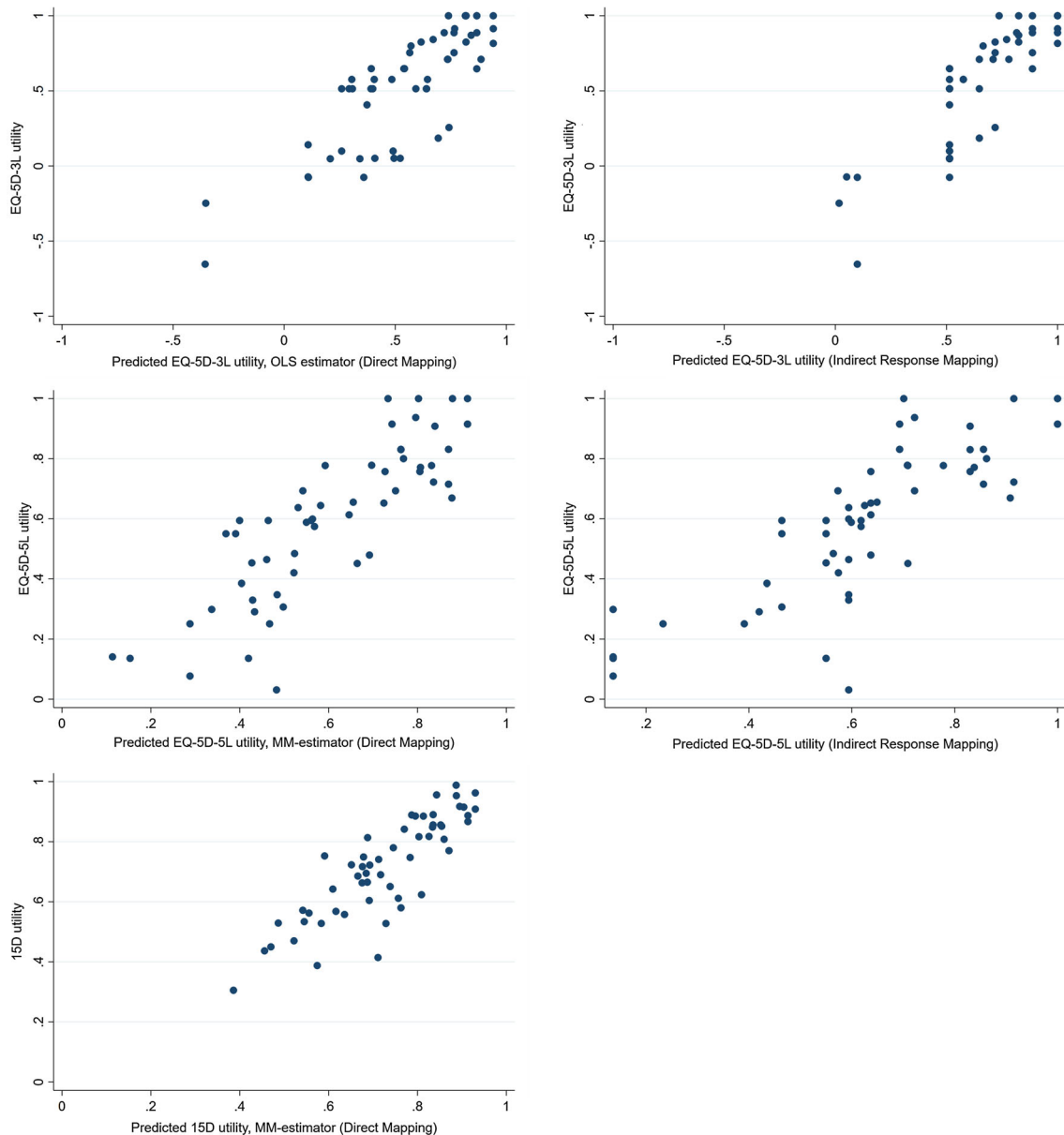
<sup>b</sup>The *Gamma* family with *log* link was used

been used interchangeably in the development of this technique to predict utilities [25, 27, 29, 30].

The mean score (SD) of PDQ-8 [27.51 (18.65)] is within what is expected according to other studies with PD patients. For example, a study conducted in Singapore with 88 patients [31] reported a mean (SD) PDQ-8 score of 31.6 (16.8), whilst another multicenter study with 134 patients from the UK, Germany, and Italy [32] showed a mean (SD) score of 28.19 (17.82). The mean (SD) utilities from MAU instruments differ (from 0.61 for EQ-5D-3L to 0.74 for 15D). Likewise, the SD of patient utilities varied by more

than 100%, from 0.16 for 15D to 0.33 for the EQ-5D-3L. Therefore, it seems necessary to compare the utilities derived from different MAU instruments with caution. All three MAU instruments show similar magnitudes of correlation with the PDQ-8D total score (range from 0.73 to 0.75).

Regarding the mapping performance, the goodness-of-fit results reported in this study are within the ranges of previously published mapping studies [33]. It is difficult to compare the performance of this study with previous mapping literature from PDQ-8 owing to the different



**Fig. 1** Scatter plots between observed and predicted multi-attribute utilities (direct mapping based on item model shown on left and indirect response mapping shown on right),  $N = 56$ . *OLS* ordinary least squares

country-specific tariffs being used. Cheung et al. [25] developed a mapping algorithm onto EQ-5D-3L Japanese value sets based on patients from Singapore. The highest  $R^2$  attained was 0.547 with three significant PDQ items identified (PDQ1, PDQ2, PDQ7), and PDQ3 included as a dummy variable, plus a constant. Dams et al. [27] considered both the Unified Parkinson's Disease Rating Scale and PDQ-8 as predictors for the EQ-5D-3L (scored using a German value set developed using TTO and a European value set developed using the visual analog scale [VAS] method). For the model in which only PDQ-8 items were considered and scored using German specific tariffs, the significant predictors include the squared terms of three

items (PDQ1, PDQ2, and PDQ8) and a constant, with an overall  $R^2$  of 0.603 (slightly lower than the  $R^2$  reported in this study of 0.619). The mapping functions onto a European VAS value set attained a slightly higher  $R^2$  (0.666) with the same items (PDQ1, and the squared terms of PDQ2 and PDQ8).

There are several strengths of this study [34]. Firstly, this study developed mapping algorithms onto three MAU instruments and offers future users a choice of mapping algorithms that are regarded to be more suitable in their studies. Secondly, the optimal mapping algorithms were selected from among different statistical methods. Thirdly, in addition to the more widely used direct mapping, this

**Table 4** Direct mapping equations from PDQ-8 scores to multi-attribute utilities,  $N = 172$ 

	Overall score			Item score		
	15D MM	EQ-5D-3L OLS	EQ-5D-5L MM	15D MM	EQ-5D-3L OLS	EQ-5D-5L MM
PDQ-8 overall score	- 0.633038 [0.047]**	- 1.204545 [0.095]**	- 0.903466 [0.087]**			
PDQ1 (mobility)				- 0.016630 [0.006]**	0.139035 [0.037]**	- 0.042722 [0.015]**
PDQ2 (activities of daily living)				- 0.032457 [0.008]**	- 0.098250 [0.015]**	- 0.058275 [0.014]**
PDQ3 (emotional well-being)				- 0.043158 [0.007]**	- 0.048426 [0.014]**	- 0.035174 [0.012]**
PDQ6 (communication)				- 0.035125 [0.010]**	- 0.051234 [0.015]**	- 0.033793 [0.010]**
PDQ7 (bodily discomfort)				- 0.025980 [0.008]**		
PDQ8 (stigma)						- 0.038236 [0.012]**
PDQ1 (mobility) squared					- 0.064723 [0.011]**	
Constant	0.929969 [0.011]**	0.944152 [0.030]**	0.896904 [0.020]**	0.930228 [0.011]**	0.867057 [0.028]**	0.912667 [0.018]**

The PDQ-8 overall score included in the regression model was calculated as original scores divided by 100

MM robust MM-estimator, OLS ordinary least squares, PDQ Parkinson's Disease Questionnaire

\*\* $p < 0.01$ , \*  $p < 0.05$ . Standard errors in parentheses

study has also conducted indirect response mapping to predict the responses to each of the EQ-5D dimensions. One of the key advantages of this approach is that by applying different country-specific value sets of the EQ-5D, the reported mapping functions can be more widely applied by users from other countries. However, it should be noted that the generalizability of the indirect response mapping functions depends on whether PD patients from other countries will have a similar response pattern to patients in Spain. Further external validation is warranted.

Some limitations should also be considered to accurately interpret the results. Firstly, with respect to the sample, although the study included 280 patients, all analyses were conducted with only the 228 patients who completed all the questionnaires. Also, the patients with the highest degree of severity possible (Stage V) were excluded because of lack of autonomy to complete the questionnaire set by themselves. Furthermore, the convenience sample does not guarantee representativeness for all patients with Parkinson's disease. To ratify the results reported here and to determine other psychometric properties, such as longitudinal response, more research with a larger sample size would be desirable. Secondly, the use of an external dataset

for validation is strongly recommended. Thirdly, although EQ-5D-5L and EQ-5D-3L are very similar with regard to the health state classification system (i.e., they use five identical dimensions), the distributions are very different. The selected predictors (items) from the PDQ-8 in each optimal mapping algorithm differ slightly in Table 4. It is not the aim of this study to investigate further why such differences exist between EQ-5D-3L and EQ-5D-5L value sets; however, it is an interesting topic that deserves further investigation.

Despite the growing popularity of mapping algorithms, there are doubts about whether it should be the first-choice method. McCabe et al. [35] support the idea of expressing the value of healthcare interventions using a consistent tool, but according to them, these methods are not as reliable as expected when the aim is to measure preferences about health. Likewise, in the same year, another study [36] concluded that mapping should mostly be viewed as a second-best solution. In response to these observations, other probabilistic methods have recently been published to predict the utilities of EQ-5D from PDQ-8 and PDQ-39, using multidimensional Bayesian network classifiers [37, 38].



**Table 5** Indirect response mapping equations from PDQ-8 scores to each EQ-5D dimension (order logit estimates),  $N = 172$ 

	EQ-5D-3L dimensions				EQ-5D-5L dimensions			
	MO	SC	UA	AD	MO	SC	UA	AD
PDQ1 (mobility)	0.622117 [0.195]**		0.686542 [0.203]**		0.95393 [0.166]**	0.439332 [0.161]**	0.572511 [0.156]**	
PDQ2 (activities of daily living)	0.496243 [0.195]*	1.825439 [0.230]**	1.366004 [0.245]**	0.31239 [0.189]	0.547628 [0.148]**	1.187733 [0.172]**	0.764347 [0.159]**	0.325193 [0.134]*
PDQ3 (emotional well-being)			0.552333 [0.179]**	2.058504 [0.285]**				1.316509 [0.174]**
PDQ6 (communication)	0.392542 [0.173]*			0.432654 [0.183]*			0.300542 [0.139]*	0.449215 [0.142]**
PDQ7 (bodily discomfort)				0.384708 [0.179]*	0.434697 [0.217]*			
PDQ8 (stigma)				0.728402 [0.181]**				0.542318 [0.151]**
Male				- 0.983751 [0.378]**	- 1.139082 [0.449]*			- 1.265199 [0.317]**
Age						0.037944 [0.016]*		
/cut1	1.098648 [0.292]**	2.134081 [0.312]**	1.918255 [0.355]**	- 0.154559 [0.416]	0.049458 [0.231]	3.468294 [1.133]**	0.314927 [0.250]	- 0.746302 [0.375]*
/cut2	7.306325 [0.986]**	7.308964 [0.841]**	8.477382 [1.062]**	4.268743 [0.600]**	2.090251 [0.279]**	5.839697 [1.171]**	2.455812 [0.308]**	1.04959 [0.379]**
/cut3					4.44236 [0.438]**	7.741082 [1.278]**	4.616362 [0.465]**	3.212781 [0.463]**
/cut4					7.611413 [0.881]**	9.565966 [1.420]**	6.80653 [0.694]**	4.884142 [0.585]**

EQ-5D dimensions: MO mobility, SC self-care, UA usual activities, PD pain/discomfort, AD anxiety/depression, PDQ Parkinson's Disease Questionnaire./cut# estimated cutpoints  
 \*\* $p < 0.01$ , \* $p < 0.05$ . Standard errors in parentheses

## 5 Conclusions

The algorithm for each MAU instrument allows the calculation of health state utility scores with good precision in a PD sample. These algorithms can be used for CUA in order to assess the cost effectiveness of interventions in patients with PD when only the disease-specific PDQ instrument was used. In addition, the indirect mapping functions reported in this study for two EQ-5D questionnaires will further facilitate the calculation of EQ-5D utility scores using other country-specific tariffs.

**Acknowledgements** JMAP, FISM, MAGG, and JCA conceived and designed the study. JMCF, MAGG, DCM, BPC, and JCA participated in the acquisition of the data. MAGG and JCA oversaw all stages of data collection and entry. GC, JMAP, and FISM contributed to interpretation of data and performed the statistical analyses. MAGG drafted the manuscript. JMCF and DCM provided critical reviews. GC and JMCF gave feedback on the manuscript and revised it critically for important intellectual content. DCM and MAGG reviewed the content of the manuscript. BPC and GC reviewed it in order to improve the grammar. All authors read and approved the first version sent and subsequent versions.

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

### Compliance with Ethical Standards

**Financial support** None declared.

**Conflicts of interest** GC has no conflicts of interest. MAGG has no conflicts of interest. DCM has no conflicts of interest. BPC has no conflicts of interest. JCA has no conflicts of interest. JMCF has no conflicts of interest. JMAP has no conflicts of interest. FISM has no conflicts of interest.

## References

- Gibb WRG, Lees AJ. The significance of the lewy body in the diagnosis of idiopathic Parkinson's Disease. *Neuropathol Appl Neurobiol.* 1989;15(1):27–44.
- Lees AJ, Hardy J, Revesz T. Parkinson's disease. *Lancet.* 2009;373(9680):2055–66.
- Dowding CH, Shenton CL, Salek SS. A review of the health-related quality of life and economic impact of Parkinson's disease. *Drugs Aging.* 2006;23(9):693–721.
- Martinez-Martin P, Rodriguez-Blazquez C, Paz S, Forjaz MJ, Frades-Payo B, Cubo E, et al. Parkinson symptoms and health related quality of life as predictors of costs: a longitudinal observational study with linear mixed model analysis. *PLoS One.* 2015;10(12):e0145310.
- Tosh JC, Longworth LJ, George E. Utility values in National Institute for Health and Clinical Excellence (NICE) Technology Appraisals. *Value Health.* 2011;14(1):102–9.
- Earnshaw J, Lewis G. NICE guide to the methods of technology appraisal: pharmaceutical industry perspective. *PharmacoEconomics.* 2008;26(9):725–7.
- Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes.* Oxford: Oxford University Press; 2015.
- Richardson J, McKie J, Bariola E. Multi attribute utility instruments and their use. In: Culyer AJ, editor. *Online encyclopedia of health economics.* San Diego: Elsevier Science; 2014. p. 341–57.
- Heniford BT, Walters AL, Lincourt AE, Novitsky YW, Hope WW, Kercher KW. Comparison of generic versus specific quality-of-life scales for mesh hernia repairs. *J Am Coll Surg.* 2008;206(4):638–44.
- Maly M, Vondra V. Generic versus disease-specific instruments in quality-of-life assessment of chronic obstructive pulmonary disease. *Methods Inf Med.* 2006;45(2):211–5.
- Rowen D, Brazier J, Ara R, Zouraq IA. The role of condition-specific preference-based measures in health technology assessment. *PharmacoEconomics.* 2017;35(Suppl. 1):33–41.
- Teckle P, McTaggart-Cowan H, Van der Hoek K, Chia S, Melosky B, Gelmon K, et al. Mapping the FACT-G cancer-specific quality of life instrument to the EQ-5D and SF-6D. *Health Qual Life Outcomes.* 2013;11:203.
- Muangpaisan W, Mathews A, Hori H, Seidel D. A systematic review of the worldwide prevalence and incidence of Parkinson's disease. *J Med Assoc Thai.* 2011;94(6):749.
- Badia X, Roset M, Montserrat S, Herdman M, Segura A. The Spanish version of EuroQol: a description and its applications. *European Quality of Life scale.* *Med Clin.* 1999;112(Suppl 1):79–85.
- Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20(10):1727–36.
- Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Med Decis Making.* 2001;21(1):7–16.
- Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med Care.* 2017;55(7):e51–8.
- Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med.* 2001;33(5):328–36.
- García-Gordillo M, del Pozo-Cruz B, Adsuar JC, Sanchez-Martinez FI, Abellan-Perpignan JM. Validation and comparison of 15-D and EQ-5D-5L instruments in a Spanish Parkinson's disease population sample. *Qual Life Res.* 2014;23(4):1315–26.
- Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The PDQ-8: development and validation of a short-form Parkinson's disease questionnaire. *Psychol Health.* 1997;12(6):805–14.
- Chen G, Stevens K, Rowen D, Ratcliffe J. From KIDSCREEN-10 to CHU9D: creating a unique mapping algorithm for application in economic evaluation. *Health Qual Life Outcomes.* 2014;12:134.
- Chen G, Khan MA, Iezzi A, Ratcliffe J, Richardson J. Mapping between 6 multiattribute utility instruments. *Med Decis Making.* 2016;36(2):160–75.
- Fox J. *Applied Regression Analysis and Generalized Linear Models (Second Edition).* Thousand Oaks: SAGE Publications; 2008.
- Freese J, Long JS. *Regression models for categorical dependent variables using Stata.* Austin: Stata; 2006.
- Cheung Y, Tan L, Lau P, Au W, Luo N. Mapping the eight-item Parkinson's Disease Questionnaire (PDQ-8) to the EQ-5D utility index. *Qual Life Res.* 2008;17(9):1173–81.
- McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods.* 1996;1(1):30–46.
- Dams J, Klotsche J, Bornschein B, Reese JP, Balzer-Geldsetzer M, Winter Y, et al. Mapping the EQ-5D index by UPDRS and PDQ-8 in patients with Parkinson's disease. *Health Qual Life Outcomes.* 2013;11:35.

28. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal, April 2013. London: UK National Health Service, National Institute for Health and Care Excellence; 2013.
29. Young MK, Ng SK, Mellick G, Scuffham PA. Mapping of the PDQ-39 to EQ-5D scores in patients with Parkinson's disease. *Qual Life Res.* 2013;22(5):1065–72.
30. Kent S, Gray A, Schlackow I, Jenkinson C, McIntosh E. Mapping from the Parkinson's Disease Questionnaire PDQ-39 to the Generic EuroQol EQ-5D-3L: the value of mixture models. *Med Decis Making.* 2015;35(7):902–11.
31. Tan LC, Luo N, Nazri M, Li SC, Thumboo J. Validity and reliability of the PDQ-39 and the PDQ-8 in English-speaking Parkinson's disease patients in Singapore. *Parkinsonism Relat Disord.* 2004;10(8):493–9.
32. Metta V, Logishetty K, Martinez-Martin P, Gage HM, Schartau P, Kaluarachchi T, et al. The possible clinical predictors of fatigue in Parkinson's disease: a study of 135 patients as part of international nonmotor scale validation project. *Parkinson's Disease.* 2011;2011.
33. 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.
34. Ara R, Rowen D, Mukuria C. The use of mapping to estimate health state utility values. *PharmacoEconomics.* 2017;35(Suppl 1):57–66.
35. McCabe C, Edlin R, Meads D, Brown C, Kharroubi S. Constructing indirect utility models: some observations on the principles and practice of mapping to obtain health state utilities. *PharmacoEconomics.* 2013;31(8):635–41.
36. 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.
37. Borchani H, Bielza C, Martinez-Martin P, Larrañaga P. Predicting the EQ-5D from the Parkinson's disease questionnaire PDQ-8 using multi-dimensional Bayesian network classifiers. *Biomed Eng Appl Basis Commun.* 2014;26(01):1450015.
38. Borchani H, Bielza C, Marti P, Larrañaga P. Markov blanket-based approach for learning multi-dimensional Bayesian network classifiers: an application to predict the European Quality of Life-5 Dimensions (EQ-5D) from the 39-item Parkinson's Disease Questionnaire (PDQ-39). *J Biomed Inf.* 2012;45(6):1175–84.