Nutrición Hospitalaria

Nutr Hosp. 2015;32(6):2808-2821
ISSN 0212-1611 • CODEN NUHOEQ
S.V.R. 318

Original/*Ancianos*

Validation and comparison of EQ-5D-3L and SF-6D instruments in a Spanish Parkinson´s disease population sample

Miguel Ángel García-Gordillo[1,2], B del Pozo-Cruz[3], JC Adsuar[1,4], JM Cordero-Ferrera[2], JM Abellán-Perpiñán[1] y FI Sánchez-Martínez[1]

[1]Department of Applied Economics, Faculty of Economics and Business, University of Murcia, Murcia, Spain. [2]Department of Economics, Faculty of Economics and Business, University of Extremadura, Badajoz, Spain. [3]Department of Sport and Exercise Science, Faculty of Science, The University of Auckland, Auckland, New Zealand. [4]Faculty of Sport Science, University of Extremadura, Cáceres, Spain.

Abstract

Introduction: **generic, preference-based Health-Related Quality of Life instruments are receiving growing attention in health-care decision-making process. In spite of this, to our knowledge, EQ-5D and SF-6D have never been compared in a Parkinson´s disease population sample.**

Objective: **the aim of this paper was to assess the psychometric properties of both instruments in a Spanish PD population sample.**

Methods: **a total sample of 133 patients were interviewed using EQ-5D-3L and SF-6D. The validity, level of agreement and sensitivity of both instruments were computed and then compared. The Spanish tariff has been used in both instruments.**

Results: **utilities of EQ-5D-3L and SF-6D have shown a strong correlation (r>0.50 and p<0.001) with the summary score of the PDQ-8 and the EQ-VAS score. Significant differences were observed in the stages III-IV of the Hoehn & Yahr stage. SF-6D had 51% higher efficiency than EQ-5D at detecting differences in symptoms severity.**

Discussion: **both EQ-5D-3L and SF-6D seem to be adequate generic Health-Related Quality of Life measures in terms of validity and sensitivity.**

Conclusion: **EQ-5D-3L presents greater ceiling and floor effects than the SF-6D instrument in this sample. Besides, the instrument SF-6D was better at detecting changes in symptoms severity compared with EQ-5D-3L.**

(*Nutr Hosp*. 2015;32:2808-2821)

DOI:10.3305/nh.2015.32.6.9765

Key words: *Psychometric properties. Health. Quality of Life. EQ-5D. SF-6D.*

Correspondence: Miguel Ángel García-Gordillo .
Department of Applied Economics,
Faculty of Economics and Business, University of Murcia,
Campus de Espinardo s/n, 30100, Murcia, Spain.
E-mail: miguelgarciagordillo@gmail.com

Recibido: 14-VIII-2015.
Aceptado: 6-IX-2015.

VALIDACIÓN Y COMPARACIÓN DE LOS INSTRUMENTOS EQ-5D-3L Y SF-6D EN UNA MUESTRA DE POBLACIÓN ESPAÑOLA CON ENFERMEDAD DE PARKINSON

Resumen

Introducción: **el uso de cuestionarios de calidad de vida basados en preferencias poblacionales están recibiendo cada vez más atención en el proceso de toma de decisiones en el ámbito sanitario. Sin embargo, a nuestro entender, EQ-5D y SF-6D nunca han sido comparados en una muestra de población con la enfermedad de Parkinson.**

Objetivo: **el objetivo de este trabajo fue evaluar las propiedades psicométricas de ambos instrumentos en una muestra de población española con enfermos de Parkinson.**

Métodos: **un total de 133 pacientes fueron entrevistados utilizando EQ-5D y SF-6D. La validez, el grado de acuerdo y la sensibilidad de ambos instrumentos fueron calculados para su posterior comparación. Las preferencias de la población española fueron utilizadas en ambos instrumentos.**

Resultados: **las utilidades de EQ-5D y SF-6D han mostrado una fuerte correlación (r> 0,50 y p <0,001) con la puntuación resumen del PDQ-8 y la puntuación del EQ-VAS. Hubo diferencias significativas en los estadios III-IV de la etapa de Hoehn y Yahr. SF-6D mostró mayor eficiencia (51%) que EQ-5D en la detección de diferencias en la gravedad de los síntomas.**

Discusión: **tanto EQ-5D como SF-6D parecen ser cuestionarios adecuados en términos de validez y sensibilidad.**

Conclusión: **en esta muestra EQ-5D presenta mayor efecto techo y suelo que el instrumento SF-6D. Además, el instrumento SF-6D fue mejor en la detección de cambios en la gravedad de los síntomas en comparación con el EQ-5D.**

(*Nutr Hosp*. 2015;32:2808-2821)

DOI:10.3305/nh.2015.32.6.9765

Palabras clave: *Propiedades psicométricas. Salud. Calidad de vida. EQ-5D. SF-6D.*

Introduction

Parkinson's disease (PD) is a progressive neurological disorder which is extremely common worldwide and imposes an important social and economic burden in the older population[1,2]. It affects more than 4 million people around the world and it usually appears between the ages of 50 and 65, being slightly more frequent in men than in women. This disease is characterized by motor and non-motor symptoms[3,4]. The most common motors symptoms are bradykinesia, rigidity, tremor and loss of balance[5]. Both motors and non-motor symptoms have a negative effect upon the health-related quality of life (HRQoL)[6]. The etiology of PD is still unknown[7] and although the best available treatment at the moment is levodopa (L-dopa)[8,9], in the long term it affects quality of life, due to the involuntary movements or dyskinesia that L-dopa causes[10,11]. Some surgical processes have proved to be alternative clinically effective options[12,13,14]. Notwithstanding, the main limitation is that studies use different criteria to identify and assess costs in PD.[15,16]. The most important objective is to reduce the progression by disease as much as possible[17]. Patients suffer an important deterioration in their health condition[18] and it is crucial to have appropriate instruments which are capable of measuring the patients' HRQoL, since it can provide relevant information for the decision-making process regarding medical care[7]. According to a systematic revision by Marinus et al.[19] not all specific-instruments are suitable to measure HRQoL in patients with PD. If HRQoL questionnaires are to be used in cost-effectiveness studies, they should show great reliability, validity and sensitivity[20].

Specific questionnaires need to be developed and subsequently validated because generic questionnaires do not properly reflect frequent and important aspects of the disease. In PD there are several validated questionnaires: the 39 and the 8 item PD questionnaires (PDQ-39/PDQ-8)[21,22], being the most widely used HRQoL instrument. The PDQ-8, derived from PDQ-39, is also considered as a valid and reliable instrument to measure HRQoL in PD patients[23] Although these profile-based HRQoL instruments for PD can detect important clinical changes and distinct health states, they cannot be used to make comparisons of PD patients' HRQoL with that of patients suffering from other diseases in terms of cost-effectiveness, since they are not preference-based instruments[24].

Generic, preference-based HRQoL instruments (i.e. EQ-5D, SF-6D, HUI, 15-D, etc.) can provide utilities, which are needed to calculate quality adjusted life years (QALYs). QALYs are commonly used in economic evaluation of health care programmes being the outcome measure in cost-utility analysis (CUA), a variant of cost-effectiveness analysis in which QALYs are the result units[25,26,27,28]. One of the preference-based instruments, the EuroQol-five dimensions (EQ-5D), was originally designed with three levels in each of its five dimensions (EQ-5D-3L). Recently, a new version including five levels by dimension (EQ-5D-5L) has been developed[29]. The SF-6D is another preference-based questionnaire, which is derived from the SF-36 questionnaire[30-32]. Recently, a new utility scoring algorithm has been designed in Spain using the lottery equivalent (LE) valuation technique[33].

Some studies have been conducted to test the validity of EQ-5D-3L in PD patients[34,35]. However, to our best knowledge, there is no previous study exploring psychometric proprieties of SF-6D and EQ-5D in a worldwide PD population sample.

Objective

The main aim of this study has been to test and compare the validity, sensitivity and relative efficiency (RE) of the SF-6D and EQ-5D-3L instruments in a PD population sample, which might be helpful to decide which questionnaire to use according to the conditions to be explained.

Methods

Study Design and Patient Recruitment

In this paper we present a pilot cross-sectional study with a sample of PD patients. All the participants belonged to one of fifteen local Spanish PD associations a total of 44 of the national territory. These fifteen associations were located in 13 out of 17 different regions of Spain. Patients in the study were over 18 year of age and had been diagnosed with PD and not other similar disorder. Exclusion criteria included patients classified in the V Hoehn & Yahr stage (H&YS) since they would not have been able to self-complete the questionnaires. A total of 190 sets of questionnaires were mailed to the associations involved in the study between May, 1st and July, 15th of 2012. A total of 157 patients gave their written informed consent prior to participating in the study. Finally, 133 participants (aged 64 ± 10 years, range 34 - 86 years) responded completely all questionnaires under supervision of a qualified member of the center staff in own center. These questionnaires included several generic-HRQoL instruments (EQ-5D-3L/VAS, SF-36v2, 15-D and EQ-5D-5L) and two specific-HRQoL measures for the PD (PDQ-39 and PDQ-8). A study including all variables in the same paper would reduce the number of participants that would respond completely all questionnaires and also, the conclusions for this paper might not be understood by the reader. For that, only EQ-5D-3L plus the EQ Visual Analog Scale (VAS), the SF-36v2 questionnaire, the PDQ-8 and H&YS were included in this study. Socio-demographic questions were also included, as well as some relevant questions about clinical and therapeutic aspects of patients. There is a previous

paper with different results and conclusions from other analyses using the same sample[36]. 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.

Instruments

EuroQol 5 dimensions-3 levels (EQ-5D-3L)

The EQ-5D-3L is a multi-attribute instrument that has been largely validated across the literature in HRQoL assessment[37]. It was originally developed in 1990[38] and it includes two parts: the EQ-5D descriptive system and the EQ-VAS. The descriptive system contains five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three levels: no problems, some problems, severe problems[39]. Therefore, a total of 243 distinct health states, resulting from the combination of dimensions and levels, can be obtained.

The EQ-VAS[39], is a vertical scale whose endpoints are labelled 'Best imaginable health state' (value 100) and 'Worst imaginable health state' (value 0). Although EQ-VAS responses can be used as a quantitative measure of HRQoL, it is more frequent to use the EQ-5D health states, defined by the descriptive system, and convert them into a single summary index by applying an algorithm that attaches values (or weights) to each level in each dimension. Value sets or 'tariffs' have been derived to EQ-5D-3L instrument in several countries, using the "Time Trade Off" (TTO) valuation technique[40,41,42,43]. We used the Spanish tariff developed by Badia[44], whose values range from -0.59 for the worst possible health state to 1.0 for perfect health, with 0 on the scale representing the state of being dead.

Short form health survey questionnaire 6 dimensions (SF-6D)

This instrument provides a preference-based single index measure of health from the SF-36[30]. SF-6D is made of six dimensions: physical functioning, role limitation, social functioning, pain, mental health and vitality. Each dimension can have 4, 5 or 6 levels and thus 18,000 possible health states can be obtained[30]. SF-6D utility scoring algorithms have been derived for several countries[45, 46], most of them using a Standard Gamble (SG) valuation technique. In our study we used the SF-6D value set derived by Abellán et al.[33] from a representative sample of the Spanish general population of Murcia, using a Lottery equivalent (LE) method, with utilities ranging from -0.357 to 1 (full health). Several studies have validated the use of SF-6D as preference-based of HRQoL instrument[47,48].

8 item - PD questionnaire (PDQ-8)

The PDQ-8 is a short version derived from the 39-item PDQ (PDQ-39). It is a profile-based HRQoL, which is also commonly applied in PD patients[23]. The questions are grouped into 8 dimensions: mobility, activities of daily living, emotional well-being, stigma, social support, cognition, communication and bodily discomfort[49]. Five possible answers are associated to each dimension: never, occasionally, sometimes, often and always/cannot do at all. Each item or dimension has its score and the summary index of all items is standardized on a scale with a range of 0-100, where the higher index represents worse HRQoL[35]. The Spanish adaptation of PDQ-8 has been used in this study[50]. The responsiveness of the PDQ-8 has been tested as well[51].

Hoehn & Yahr scale (H&YS)

The original H&YS was developed in 1967. Although it is not a complex scale, many authors have used it to describe disease progression. In this scale five (from 1 to 5) broad categories of motor function in PD are defined[52]. Subsequently, in a modified version, the number of stages was amplified, including stages 1.5 and 2.5. Even so, for this study has taken into account the first version since it is the most known scale. There is correlation between progression in H&YS and studies of dopaminergic loss, and also high correlations between H&YS and some standardized scales of motor disorders, disability and quality of life[53].

Data analysis

Statistical analyses were performed using Statistical Package for Social Sciences, version 18.0 (SPSS Inc., Chicago, IL, USA). The significance level was set at $p < 0.05$ in all cases. The data did not follow a normal distribution so we used Spearman`s correlation coefficient.

Descriptive statistics of PDQ-8, SF-6D, EQ-5D-3L and EQ-VAS: Descriptive statistics were computed to characterize the sample and the distribution of PDQ-8, SF-6D and EQ-5D-3L and EQ-VAS. The mean, standard deviation (SD), median, inter-quartile range (IQR), and range were computed for continuous variables. The number and proportion in the sample was shown for categorical variables. Patients in the I or II (from now on I-II) stages of H&YS were compared with those in the III or IV (from now on III-IV) stages using Mann–Whitney U or chi-square tests.

Construct Validation: Convergent validity of the SF-6D and EQ-5D-3L was assessed by examining their association with PDQ-8 and EQ-VAS at domain and scale level. Validity coefficients were computed as Spearman's rank correlation coefficients (r), with $r > 0.5$ considered as a strong correlation, 0.3 to 0.5 as a moderate correlation and 0.2 to 0.3 as a weak correlation[54].

To further extend testing validity, a "known-group" scheme was used to survey the discriminative validity of the SF-36 and EQ-5D-3L based on its ability to discriminate patients with different levels of PD severity and self-reported health status groups, alongside with other variables such as social economic status, duration of PD, ongoing therapies and the presence of other medical conditions other than PD. Mann-Whitney U tests were used to detect statistically significant effects of the dichotomous variables on utility scores. The levels of PD severity were defined based on H&YS as follows: mild to moderate if H&YS result was equal to I-II, and severe if H&YS result was equal to III-IV. The EQ-VAS score was used to classify individuals into health status groups, covering the range from very poor to very good health, a technique employed in a quite similar study[55]. Each subject was included in one of six groups according to VAS score: 0–49, 50–59, 60–69, 70–79, 80–89, and 90–100.

Level of Agreement between SF-6D and EQ-5D-3L: The intra-class correlation coefficient (ICC) and Bland-Altman plot were computed to test the agreement between the two instruments. A value greater than 0.7 in the ICC suggests a strong agreement[56]. In Bland–Altman plot, the average of the 2 measurements was plotted on the x-axis, and the difference between the two measurements on the y-axis, where SF-36 was the subtrahend. Zero difference implies total agreement, hence the deviation from 0 indicates the degree of (dis)agreement of each subject on the plot[57]. Additionally, the SF-6D and EQ-5D-3L were compared across the sample as well as for subgroups based on socio-economic and clinical characteristics, by performing paired comparisons with Wilcoxon's signed rank test and Spearman's rank correlation for the association of them.

Efficiency and Sensitivity of SF-6D and EQ-5D-3L: The RE statistic was used to test the efficiency of the SF-6D and EQ-5D-3L and to detect clinically relevant differences between PD patients. RE is defined as the ratio of the square of the t-statistic of the comparator instrument (assumed to be the SF-6D utility score for the purposes of this study) over the square of the t-statistic of the reference instrument (assumed to be the EQ-5D-3L utility score for the purposes of this study)[56]. A coefficient greater than 1 suggests that SF-6D is more efficient than EQ-5D-3L at detecting clinically relevant differences with the given sample size, while a coefficient less than 1 denotes a lower efficiency of the comparator instrument (SF-6D in our case). The sensitivity of the SF-6D and EQ-5D-3L instruments were compared and tested using receiver operating characteristic (ROC) curves[58]. The utility measure that generates the largest area under the ROC curve is regarded as the most sensitive at detecting differences in the external indicator. A measure with perfect discrimination would generate an area under the curve (AUC) score of 1.0, whilst a measure with no discriminatory power would generate an AUC score of 0.5. Self-reported health status (PDQ-8) and severity of the symptoms (H&YS) were used as external indicators. For the purposes of RE and AUC analysis, different "cut-off" points for the self-reported health status were selected for PDQ-8: 5.8 and 7.4 based on the literature[59], and 21.87 based on the PDQ-8 median; and a "cut-off" point differentiating between patients in the I-II stage and those in the III-IV stage was used for the severity of the symptoms.

Results

Descriptive Statistics of SF-6D and EQ-5D-3L: Table I shows the clinical and socio-demographic characteristics of all participants in this study, as well as the distribution of PDQ-8 scores and EQ-5D-3L and SF-6D utility values. The sample was stratified by severity of the symptoms. We did not detect statistically significant differences between the H&YS groups for socio-demographic and clinical variables except for the number of years since clinical diagnosis, which was smaller in the lower severity group according to H&YS.

SF-6D and EQ-5D-3L utilities resulted to be higher in the lower severity group ($p<0.05$).

We also found greater scores for the EQ-VAS in the lower severity group, whereas PDQ-8 scores were lower for the same group ($p<0.05$).

Table II shows the distribution of EQ-5D-3L and SF-6D results within each dimension and it allows to detect the existence of ceiling effects (i.e., most of respondents declaring no problems in a certain dimension) or floor effects (i.e. large numbers of patients at the bottom level of certain dimensions). The only dimension of EQ-5D-3L with an apparent ceiling effect was self-care (SC), whereas no floor effect emerged in any dimension. Regarding EQ-5D utilities, only 1 patient had a 0 (the worst score); whereas 18 patients had an index equal to 1 (the best score). On the other hand, only rol limitations (RL) dimension presented a ceiling effect in the SF-6D, whereas utilities of SF-6D suggest a best distribution without the existence of either a ceiling or floor effect.

Construct Validity: It can be observed on Table III a strong correlation between EQ-5D-3L and SF-6D utilities and the summary score of the PDQ-8 (-0.721; $p<0.01$ and -0.705; $p<0.01$ respectively) and the EQ-VAS (0.677; $p<0.01$ and 0.535; $p<0.01$ respectively). These correlation coefficients were larger than the ones between EQ-5D-3L and SF-6D dimensions, and PDQ-8 dimensions, on one hand, and between EQ-5-3L and SF-6D dimensions, and EQ-VAS scores, in the other hand. Moreover, the correlation EQ-5D-3L utilities and EQ-VAS scores was stronger than the correlation between SF-6D utilities and EQ-VAS scores.

Table I
Characteristics of participants and distribution of PDQ-8, EQ-5D-3L and SF-6D utility scores

Variables	*Total (n = 133)*	*Hoehn & Yahr (stages I-II) (n = 49)*	*Hoehn & Yahr (stages III-IV) (n = 84)*	*P-value*
Age (years)				
Median (IQR)	65.00 (13.00)	65.00 (14.00)	66.00 (12.00)	0.134[a]
Mean (SD)	64.33 (9.74)	62.24 (10.10)	65.55 (9.37)	
Range	34 to 86	34 to 78	38 to 86	
Gender				
Male	95 (71.4)	35 (71.4)	60 (71.4)	0.582[b]
Female	38 (28.6)	14 (28.6)	24 (28.6)	
Level of Studies				
Primary studies	65 (48.9)	18 (36.7)	47 (56.0)	0.065[b]
Secondary studies	31 (23.3)	16 (32.7)	15 (17.9)	
University studies	37 (27.8)	15 (30.6)	22 (26.2)	
Occupational status				
Self-employee	3 (2.3)	1 (2.0)	2 (2.4)	0.417[b]
Government employee	5 (3.8)	3 (6.1)	2 (2.4)	
Employee	4 (3.0)	3 (6.1)	1 (1.2)	
Housewife	11 (8.3)	4 (8.2)	7 (8.3)	
Retired	110 (82.7)	38 (77.6)	72 (85.7)	
Household size[#]				
Median (IQR)	2.00 (1.00)	2.00 (1.00)	2.00 (1.00)	0.064[a]
Mean (SD)	2.42 (1.13)	2.67 (1.21)	2.28 (0.06)	
Range	1.00 to 6.00	1.00 to 6.00	1.00 to 5.00	
Household income, €[#]				
Median (IQR)	1,700.00 (1,492.00)	1,900.00 (1,750.00)	1,600.00 (2,294.00)	0.961[a]
Mean (SD)	2,087.89 (1,369.13)	2,097.00 (1,424.00)	2,052.00 (1,330.00)	
Range	400.00 to 5,000.00	400.00 to 5,000.00	1,008.00 to 4,000.00	
Other medical conditions*				
Median (IQR)	0.00 (1.00)	0.00 (1.00)	0.00 (1.00)	0.680[a]
Mean (SD)	0.54 (0.71)	0.47 (0.58)	0.58 (0.77)	
Range	0.00 to 3.00	0.00 to 2.00	0.00 to 3.00	
Ongoing therapies†				
Median (IQR)	3.00 (2.00)	3.00 (1.00)	3.00 (2.75)	0.811[a]
Mean (SD)	2.44 (1.38)	2.51 (1.26)	2.40 (1.45)	
Range	0.00 to 6.00	0.00 to 6.00	0.00 to 6.00	
Years since clinical diagnosis[#]				
Median (IQR)	6.00 (10.00)	3.00 (4.00)	8.00 (10.50)	<0.001[a]
Mean (SD)	7.70 (6.44)	4.39 (4.05)	9.74 (6.80)	
Range	0.50 to 32.00	0.50 to 21.00	0.50 to 32.00	
PDQ-8				
Median (IQR)	21.87 (26.56)	18.75 (14.06)	29.68 (33.59)	<0.001[a]

Table I (cont.)
Characteristics of participants and distribution of PDQ-8, EQ-5D-3L and SF-6D utility scores

Variables	*Total (n = 133)*	*Hoehn & Yahr (stages I-II) (n = 49)*	*Hoehn & Yahr (stages III-IV) (n = 84)*	*P-value*
Range	0.00 to 87.00	0.00 to 56.00	3.13 to 87.50	
EQ-5D-3L utility				
Median (IQR)	0.72 (0.40)	0.82 (0.30)	0.65 (0.31)	0.001[a]
Mean (SD)	0.64 (0.31)	0.76 (0.21)	0.57 (0.34)	
Range	-0.65 to 1.00	0.05 to 1.00	-0.65 to 1.00	
EQ-VAS				
Median (IQR)	60.00 (22.50)	70.00 (30.00)	50.00 (20.00)	<0.001[a]
Mean (SD)	57.63 (19.67)	66.57 (16.60)	52.42 (19.53)	
Range	10 to 100	30 to 95	10 to 100	
SF-6D utility				
Median (IQR)	0.60 (0.39)	0.69 (0.24)	0.51 (0.49)	<0.001[a]
Mean (SD)	0.53 (0.29)	0.67 (0.19)	0.45 (0.31)	
Range	-0.23 to 0.98	0.16 to 0.93	-0.22 to 0.98	

Values are presented as n (%), unless otherwise indicated; Hoehn & Yahr: Hoehn & Yahr scale; *Number of self-reported medical conditions Parkinson´s Disease apart; †: Number of self-reported ongoing therapies for the control of the Parkinson´s Disease; EQ-5D-3L utility, Utility index from the European Quality of Life Questionnaire five dimensions three levels; VAS: Visual Analogical Scale from EQ-5D-3L; SF-6D utility, Utility index from the Short Form 6 dimensions; PDQ-8, Parkinson disease questionnaire 8 items.
[#]: Lost values (Household size n = 4; Household income n= 114; Years since clinical diagnosis n = 7)
[a]: *p*-value from Mann–Whitney U test
[b]: *p*-value from chi-square test

Table II
*Distribution of EQ-5D-3L and SF-6D * results within each domain (n=133)*

EQ-5D-3L (%)					
Level	*MO*	*SC*	*UA*	*PD*	*AD*
1	39.1	**55.6**	40.6	30.1	45.9
2	**60.2**	39.8	**53.4**	**61.7**	**52.6**
3	0.8	4.5	6.0	8.3	1.5
At celling (11111); n, (%)			18 (13.5)		
At floor (33333); n, (%)			1 (0.8)		

SF-6D (%)						
Level	*PF*	*RL*	*SF*	*P*	*MH*	*V*
1	3.0	**45.9**	29.3	21.1	12.0	5.3
2	15.8	9.0	20.3	13.5	17.3	24.1
3	26.3	9.8	**33.8**	**30.1**	**45,9**	**45.9**
4	3.0	35.3	9.0	15.8	23.3	17.3
5	**36.8**	/	7.5	12.8	1.5	7.5
6	15.0	/	/	6.8	/	/
At celling (11111); n, (%)			0 (0.0)			
At floor (645655); n, (%)			0 (0.0)			

*EQ-5D-3L dimensions: MO, mobility; SC, self-care; UA, usual activities PD, pain/discomfort; AD, anxiety/depression; SF-6D dimensions: PF, physical functioning RL, role limitation; SF, social functioning; P, pain; MH, mental health; V, vitality. Level in mold is in bold.
EQ-5D-3L, European Quality of Life Questionnaire five dimensions three levels; SF-6D, Short Form 6D.

Table III
Spearman`s rank correlation coefficient between EQ-5D-3L, EQ-VAS or SF-6D and PDQ-8 (n=133)

	PDQ-8									
	Mobility	*Activities of daily living*	*Emotional well-being*	*Stigma*	*Social support*	*Cognitions*	*Communication*	*Bodily discomfort*	*Summary Score*	*EQ-VAS*
EQ-5D										
Utility	-.564**	-.628**	-.548**	-.352**	-.376**	-.469**	-.340**	-.469**	-.721**	.677**
MO	.472**	.480**	.324**	.195*	.276**	.366**	.229**	.274**	.506**	-.557**
SC	.550**	.648**	.429**	.294**	.339**	.343**	.338**	.333**	.614**	-.551**
UA	.538**	.599**	.408**	.296**	.320**	.385**	.327**	.321**	.599**	-.573**
PD	.286**	.323**	.378**	.209*	.262*	.389**	.128	.524**	.500**	-.510**
AD	.272**	.341**	.687**	.412**	.317**	.338**	.293**	.381**	.585**	-.393**
SF-6D										
Utility	-.450**	-.612**	-.577**	-.309**	-.451**	-.532**	-.328**	-.435**	-.711**	.535**
PF	.237**	.142	.007	.118	.008	.110	.171*	.034	.162	-.060
RL	.143	.097	.044	.122	-.011	.108	.100	.015	.103	-.007
SF	.060	.006	-.032	.078	.056	.163	.084	.029	.046	-.016
P	.183*	.223*	.050	.215*	.115	.100	.120	.055	.184*	-.137
MH	.089	.174*	-.040	.023	-.008	.134	-.037	.059	.080	-.159
V	.132	.107	.013	.111	.131	.179*	.180*	.057	.156	-.255**

*P <0.05 (two-tailed); **P <0.001 (two-tailed).
Hypothesized moderate-to-strong correlations were shaded.
EQ-5D dimensions: MO, mobility; SC, self-care; UA, usual activities; PD, pain/discomfort; AD, anxiety/depression.
SF-6D dimensions: PF, physical functioning; RL, role limitation; SF, social functioning; P, pain; MH, mental health; V, vitality.
PDQ-8: Parkinson´s Disease questionnaire 8 items.
EQ-5D-3L: European Quality of Life Questionnaire five dimensions three levels;
SF-6D, Short Form 6dimensions.
EQ-VAS, EuroQol Visual Analog Scale.

Univariate analyses for SF-6D and EQ-5D-3L (Table IV) showed that utility scores gradually decrease with increasing PDQ-8 score and H&YS. By contrast, SF-6D and EQ-5D-3L utilities increase with increasing EQ-VAS scores. The same increasing pattern by ranges (six subgroups) was observed for EQ-VAS scores in absolute terms (Table V).

There were significant differences between EQ-5D-3L and SF-6D utilities in all variables except in those patients that did not receive any therapy. Most of the variables-based subgroups show statistically significant differences for utilities derived from both instruments: it was the case in gender, level of studies, EQ-VAS score, PDQ-8 score and H&YS. Differences according to age and years since clinical diagnosis were only statistically significant for EQ-5D-3L utilities.

Level of Agreement between SF-6D and EQ-5D-3L: As can be seen in table IV, EQ-5D-3L utility scores were greater than SF-6D utilities and statistically significant differences were found in all variables except in those patients that did not receive any therapy. SF-6D and EQ-5D-3L scores had a strong correlation for the sample as a whole (n=133), with high Spearman's correlation coefficient and ICC (r = 0.735; ICC 0.85; p<0.001).

The levels of correlation and ICC for the rest of variables-based subgroups were similar.

Bland–Altman analysis indicated that the 95% limits of agreement between EQ-5D-3L and SF-6D ranged from -0.279 to 0.58 (0.15 (0.21) bias (SD)) and over 95% points lie within limits (Fig. 1). A systematic discrepancy in the utility difference of EQ-5D-3L and SF-6D scores was observed, with higher SF-6D at lower mean utilities, and lower SF-6D at higher mean utility scores.

Efficiency and Sensitivity of SF-6D and EQ-5D-3L: RE statistic calculations showed that both EQ-5D-3L and SF-6D utilities score are similar to detect differences between patients with optimal HRQoL when the selected "cut-off" point was 5.8 PDQ-8 summary score points. Meanwhile the efficiency of SF-6D was lower when compared with the EQ-5D-3L when the selected "cut-off" point was fixed at 7.4 PDQ-8 summary score points. Instead, when the selected "cut-off" point was 21.87 PDQ-8 median, SF-6D showed better RE (38%) than EQ-5D-3L. Moreover, AUC

Table IV
Univariate analyses and comparison of SF-6D and EQ-5D-3L Spain utility scores for all patients and for several socio-demographic or health characteristics-based subgroups

Variables-based subgroups	*n (%)*	*Mean (SD)*		*Median (IQR)*		*P-value* [a]	*ICC (95% IC)*	*Spearman´s rank correlation coefficient*
		$SF\text{-}6D_{utility}$	$EQ\text{-}5D\text{-}3L_{utility}$	$SF\text{-}6D_{utility}$	$EQ\text{-}5D\text{-}3L_{utility}$			
All patients	133	0.53 (0.29)	0.64 (0.31)	0.60 (0.39)	0.72 (0.34)	<0.001	0.85 (0.79 to 0.89)	0.735**
Age (years)								
34-65	67 (50.40)	0.50 (0.27)	0.60 (0.27)	0.54 (0.40)	0.60 (0.31)	0.001	0.80 (0.67 to 0.87)	0.736**
66-90	66 (49.60)	0.55 (0.31)	0.68 (0.34)	0.66 (0.41)	0.80 (0.32) [b]	<0.001	0.88 (0.80 to 0.92)	0.684**
Gender								
Male	95 (71.40)	0.58 (0.27)	0.69 (0.28)	0.65 (0.39) [b]	0.78 (0.37)[b]	<0.001	0.80 (0.70 to 0.87)	0.638**
Female	38 (28.60)	0.40 (0.31)	0.52 (0.35)	0.39 (0.46)	0.57 (0.71)	0.003	0.89 (0.79 to 0.94)	0.879**
Level of Studies								
Non-university studies	96 (72.20)	0.49 (0.29)	0.60 (0.33)	0.54 (0.43)	0.70 (0.31)	<0.001	0.83 (0.75 to 0.89)	0.728**
University studies	37 (27.80)	0.63 (0.27)	0.74 (0.23)	0.69 (0.33)[b]	0.80 (0.30)[b]	<0.001	0.86 (0.72 to 0.92)	0.687**
Occupational status								
Employee	23 (17.30)	0.51 (0.30)	0.58 (0.28)	0.58 (0.33)	0.60 (0.24)	0.06	0.88 (0.72 to 0.95)	0.628**
Retired	110 (82.70)	0.53 (0.29)	0.65 (0.31)	0.60 (0.41)	0.78 (0.37)	<0.001	0.84 (0.77 to 0.89)	0.722**
Other medical conditions[∞]								
Yes	60 (45.10)	0.50 (0.27)	0.62 (0.30)	0.54 (0.39)	0.71 (0.34)	0.001	0.76 (0.61 to 0.86)	0.635**
No	73 (54.90)	0.55 (0.31)	0.65 (0.32)	0.66 (0.39)	0.75 (0.34)	<0.001	0.89 (0.83 to 0.93)	0.793**
Ongoing therapies†								
Yes	120 (90.20)	0.54 (0.28)	0.65 (0.30)	0.62 (0.39)	0.73 (0.36)	<0.001	0.85 (0.78 to 0.89)	0.728**
No	13 (9.80)	0.39 (0.33)	0.49 (0.32)	0.36 (0.50)	0.51 (0.61)	0.133	0.82 (0.41 to 0.94)	0.794**

Table IV (cont.)
Univariate analyses and comparison of SF-6D and EQ-5D-3L Spain utility scores for all patients and for several socio-demographic or health characteristics-based subgroups

Variables-based subgroups	*n (%)*	*Mean (SD)*		*Median (IQR)*		*P-value* [a]	*ICC (95% IC)*	*Spearman´s rank correlation coefficient*
		SF-6D$_{utility}$	*EQ-5D-3L$_{utility}$*	*SF-6D$_{utility}$*	*EQ-5D-3L$_{utility}$*			
Years since clinical diagnosis[#]								
≤10	91 (68.40)	0.56 (0.28)	0.68 (0.26)	0.62 (0.36)	0.75 (0.37)[b]	<0.001	0.83 (0.75 to 0.89)	0.712**
>10	35 (26.30)	0.45 (0.32)	0.51 (0.39)	0.53 (0.54)	0.51 (0.72)	0.08	0.86 (0.72 to 0.93)	0.789**
EQ-VAS								
≤65	86 (64.70)	0.43 (0.29)	0.52 (0.31)	0.46 (0.45)	0.57 (0.24)	<0.001	0.83 (0.74 to 0.89)	0.780**
>65	47 (35.30)	0.71 (0.19)	0.85 (0.18)	0.75 (0.20)[b]	0.88 (0.19)[b]	<0.001	0.61 (0.31 to 0.78)	0.251
PDQ-8 $_{\text{summary score}}$								
≤5.8	8 (6.01)	0.85 (0.08)	0.95 (0.08)	0.86 (0.12)[b]	1.00 (0.09)[b]	0.011	0.76 (-0.19 to 0.95)	0.671
>5.8	125 (93.99)	0.51 (0.29)	0.62 (0.31)	0.56 (0.39)	0.71 (0.31)	<0.001	0.83 (0.76 to 0.88)	0.713**
≤7.4	15 (11.27)	0.79 (0.14)	0.92 (0.08)	0.84 (0.09)[b]	1.00 (0.18)[b]	0.001	0.16 (-1.49 to 0.71)	0.421
>7.4	118 (88.73)	0.49 (0.29)	0.60 (0.31)	0.55 (0.38)	0.65 (0.31)	<0.001	0.83 (0.76 to 0.88)	0.730**
≤21.87	59 (44.36)	0.71 (0.17)	0.81 (0.19)	0.74 (0.21)[b]	0.84 (0.21)[b]	<0.001	0.643 (0.40 to 0.78)	0.437**
>21.87	74 (55.63)	0.38 (0.29)	0.50 (0.32)	0.38 (0.44)	0.51 (0.20)	<0.001	0.815 (0.70 to 0.88)	0.756**
Hoehn & Yahr								
Stages I-II	49 (36.84)	0.67 (0.19)	0.76 (0.21)	0.69 (0.24)[b]	0.82 (0.30)[b]	0.003	0.68 (0.44 to 0.82)	0.557**
Stages III-IV	84 (63.16)	0.45 (0.31)	0.57 (0.34)	0. 51(0.49)	0.64 (0.31)	<0.001	0.85 (0.78 to 0.90)	0.802**

Values are presented as n (%), unless otherwise indicated; [∞] self-reported medical conditions Parkinson´s Disease apart; †: self-reported ongoing therapies for the control of the Parkinson´s Disease; EQ-5D-3L$_{utility}$: Spanish utility index from the European Quality of Life Questionnaire 5 dimensions three levels; EQ-VAS: EuroQol Visual Analogical Scale score L; SF-6D$_{utility}$: Spanish utility index from the Short Form 6 Dimensions; PDQ-8: Parkinson´s Disease questionnaire 8 items score; Hoehn & Yahr: Hoehn & Yahr scale; ICC: intraclass correlation coefficient; IQR: interquartile range.
[#]: Lost values: Years since clinical diagnosis n = 7 (5.3%);
[a]: Paired comparisons of SF-6D and EQ-5D-3L utilities scores were made with Wilcoxon's signed-rank test.
*p<0.05 (two-tailed); **p<0.001 (two-tailed).
[b]: Independent comparisons of SF-6D and EQ-5D-3L utilities scores were made with Mann–Whitney U test.

Table V
Comparison of the SF-6D and EQ-5D-3L utilities across the EQ-VAS-based health groups

Health status groups			*Utility comparisons*			
EQ-VAS Range	*n (%)*	*Age Mean (SD)*	*SF-6D$_{utility}$ Mean (SD)*	*EQ-5D-3L$_{utility}$ Mean (SD)*	*Difference†*	*Effect size[a]*
0-100	133 (100)	65.00 (13.00)	0.53 (0.29)	0.64 (0.31)	0.11**	0.51
0-49	33 (24.81)	65.58 (8.35)	0.27 (0.30)	0.32 (0.34)	0.05	0.22
50-59	32 (24.06)	65.19 (9.90)	0.56 (0.25)	0.68 (0.18)	0.12**	0.77
60-69	24 (18.04)	61.13 (9.87)	0.49 (0.21)	0.65 (0.25)	0.16**	0.85
70-79	15 (11.27)	61.60 (12.12)	0.68 (0.29)	0.75 (0.26)	0.7	0.51
80-89	19 (14.28)	66.74 (9.35)	0.71 (0.11)	0.86 (0.11)	0.15**	1.99
90-100	10 (7.51)	64.70 (9.56)	0.78 (0.16)	0.93 (0.07)	0.15*	1.78
Explained variance[b]			.295	.406		

EQ-VAS: EuroQol Visual Analogue Scale score. EQ-5D-3L$_{utility:}$ Spanish utility index from the European Quality of Life Questionnaire 5 dimensions 3 levels. SF-6D$_{utility:}$ Spanish utility index from SF-6 dimensions.
*P<0.05; **P<0.001
†Paired comparisons of SF-6D and EQ-5D-3L utility scores were made with Wilcoxon's signed-rank test.
[a]Negative effect sizes indicate a lower EQ-5D-3L mean score compared to the respective SF-6D mean score
[b]Expressed as R^2 and corresponds to the % of variance in the SF-6D$_{utility}$ and the EQ-5D-3L$_{utility}$ explained by the EQ-5D-3L VAS-based health status groups

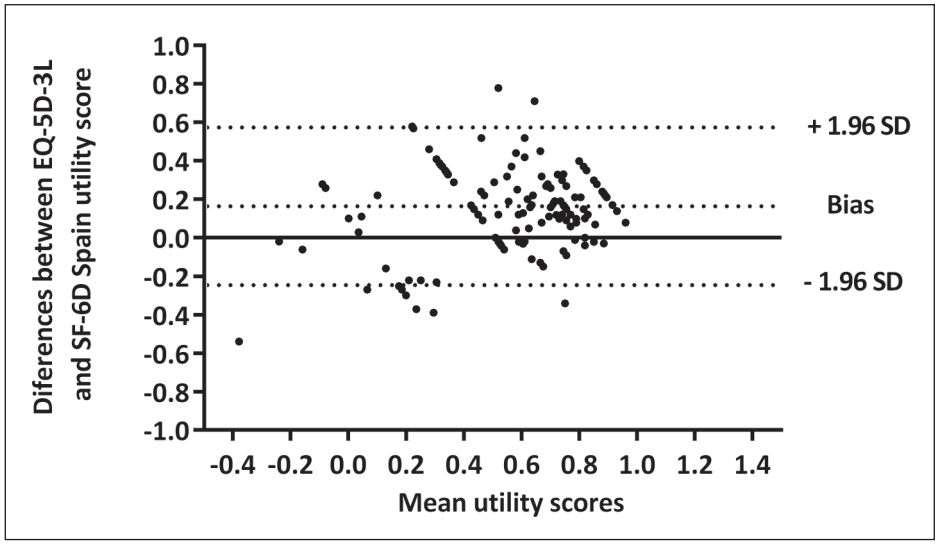

Fig. 1.—Bland-Altman plot for all 133 individuals who completed both the EQ-5D and SF-6D.

scores above 0.5 confirm the ability of the two instruments to detect clinical differences on HRQoL in PD patients. Conversely, the SF-6D was more efficient at detecting differences between patients at mild and moderate to strong severity symptoms as measured by H&YS, ratified when the greater AUC of SF-6D was observed.

Discussion

To our best knowledge, this is the first study in the world testing and comparing the validity and performance of SF-6D and EQ-5D-3L HRQoL measures in a PD population sample. Previous studies have verified the validity of EQ-5D-3L instrument in this disease elsewhere[35,34,51], but no one has verified the validity of SF-6D instrument in PD patients. The present study provides relevant information on the psychometric properties of SF-6D. This information will be relevant for future applications on the correct preference-based HRQL measurement and economic evaluations related to PD.

According to our results, both instruments have shown validity and sensitivity as instruments to assessing HRQoL in patients with PD varying in socio-demographic, clinical characteristics and EQ-VAS-based health status. In consequence, they are feasible and

Table VI
Efficiency of EQ-5D-3L and SF-6D to detect clinically relevant difference in quality of life and in the severity of the symptoms

				t-Test			*ROC curve*	
Measure	*PDQ-8*	*n*	*Mean (SD)*	*t-statistic*	*p-value*	*RE†*	*AUC*	*95%CI*
$EQ\text{-}5D\text{-}3L_{utility}$	≤ 5.8	8	0.95 (0.08)	8.37	<0.001	1.00	0.894*	(0.802 to 0.986)
	>5.8	125	0.62 (0.31)					
$SF\text{-}6D_{utility}$	≤ 5.8	8	0.85 (0.08)	8.83	<0.001	1.11	0.898*	(0.829 to 0.967)
	>5.8	125	0.51 (0.29)					
$EQ\text{-}5D\text{-}3L_{utility}$	≤ 7.4	15	0.92 (0.08)	8.55	<0.001	1.00	0.860*	(0.785 to 0.936)
	>7.4	118	0.60 (0.31)					
$SF\text{-}6D_{utility}$	≤ 7.4	15	0.79 (0.14)	6.54	<0.001	0.58	0.842*	(0.744 to 0.939)
	>7.4	118	0.49 (0.29)					
$EQ\text{-}5D\text{-}3L_{utility}$	≤21.87	59	0.81 (0.19)	6.92	<0.001	1.00	0.835*	(0.765 to 0.905)
	>21.87	74	0.50 (0.32)					
$SF\text{-}6D_{utility}$	≤21.87	59	0.71 (0.17)	8.15	<0.001	1.38	0.842*	(0.775 to 0.909)
	>21.87	74	0.38 (0.29)					
Measure	*Hoehn & Yahr*	*n*	*Mean (SD)*	*t-statistic*	*p-value*	*RE†*	*AUC*	*95%CI*
$EQ\text{-}5D\text{-}3L_{utility}$	Stages I-II	49	0.76 (0.21)	3.993	<0.001	1.00	0.679*	(0.588 to 0.771)
	Stages III-IV	84	0.57 (0.34)					
$SF\text{-}6D_{utility}$	Stages I-II	49	0.67 (0.19)	4.916	<0.001	1.51	0.713*	(0.625 to 0.800)
	Stages III-IV	84	0.45 (0.31)					

$EQ\text{-}5D\text{-}3L_{utility:}$ Spanish utility index from the European Quality of Life Questionnaire 5 dimensions 3 levels.
$SF\text{-}6D_{utility:}$ Spanish utility index from SF-6 dimensions.
Hoehn & Yahr: Hoehn & Yahr scale;
AUC: area under ROC curve; CI: confidence interval; RE: relative efficiency; ROC: receiver operating characteristic.
*$p < 0.001$ indicates that AUC statistically significantly greater than 0.5.
†Reference is $EQ\text{-}5D\text{-}3L_{utility}$.

acceptable instruments to obtain the utility score in this population. Even though these 2 instruments show similar performance, a difference at individual level was observed

The convergent validity of SF-6D and EQ-5D-3L was demonstrated through their moderate to strong correlations with PDQ-8. It is clear that EQ-5D-3L is a shorter instrument than SF-36, therefore it is easier and faster to complete by patients; however, SF-6D covers other areas and therefore generates further information. In our case, utility scores from EQ-5D-3L and SF-6D have showed a similar and strong correlation with the summary score of the PDQ-8 (Table III). Nevertheless, the limited evidence in the scientific literature suggests that different measures are not always interchangeable for their use in CUA[36,60,61].

Although EQ-5D-3L may not measure similar constructs than SF-6D, both scores were strongly correlated with summary score of PDQ-8 and their ICC showed a strong agreement in almost all tested cases. Both the EQ-5D-3L and the SF-6D showed a decrease in utilities scores with increasing of PDQ-8 summary scores and H&YS. Conversely, the utilities from both instruments increase with a higher classification according to EQ-VAS. SF-6D and EQ-5D-3L utilities were found to be higher in the lower severity group ($p<0.05$). Even so, in the higher severity group, the EQ-5D-3L utilities score were significantly higher than the SF-6D (Tables I and IV). Regarding ceiling and floor effects, a 13.5% of participants had a score of 1 and only 0.8% had a 0 score in EQ-5D-3L. One dimension presented a ceiling effect in SF-6D responses. Comparing EQ-5D-3L and SF-6D utilities across the EQ-VAS-based health groups in this Spanish PD population sample, it has been observed that EQ-5D-3L produces higher utilities than SF-6D. According to the results shown in table V, differences between the two instruments were less likely to be accounted for in PD patients with poorer health condition –attending to VAS scores–. These results could have potential implications in CUA so that QALY gains valuations might differ depending on utility at baseline.

Both instruments were valid to discriminate patients with different self-reported health status and severity of symptoms. Depending on the "cut-off" (5.8 or 7.4), the efficiency of SF-6D to detect clinically relevant differences in quality of life may be lower or higher than EQ-5D-3L. The use of the "cut-off" is valid if it is applied to absolute scores but, its practice is more efficient in scores derived in longitudinal studies[59] as the "known-groups" are not so unbalanced in size. Adding the median of PDQ-8 as "cut-off" in our paper plus 5.8 and 7.4 "cut-off" it was possible to double check whether the two sub-groups are meaningfully different and also the statistic power was increased. Meanwhile, SF-6D seems to be significantly better in terms of efficiency at detecting clinical changes in PD severity of symptoms.

The choice of an instrument may influence the results measured; therefore, it would be interesting to do comparisons of cost-effectiveness ratios using the same HRQoL instrument and scoring algorithm in interventions, therapies or programs. There exists a study which has showed how the incremental cost-effectiveness ratio might lead to different recommendations if the QALY measurement was based on EQ-5D or SF-6D data[62]. We have shown that SF-6D may be a valid instrument in economic evaluation of interventions targeting PD patients. It was efficient in detecting clinically relevant differences in quality of life and severity of symptoms and, it may be less prone to suffer from ceiling effects than EQ-5D-3L.

This study has several limitations that need to be understood to achieve a logical interpretation of the results. One of them is related to the cross-sectional design. Although sensitive measures are generally considered to be reliable[56] it would be appropriate to make a longitudinal study in patients with PD, since PD is a chronic disease. On the other hand, the clinical conditions were self-reported, and there is the possibility that the reliance on such data may result in biased estimates of the prevalence of some conditions[63]. Furthermore, the use of EQ-VAS in a specific condition could fail to capture underlying disease severity; but in this study it has shown to be valid as a discriminator of overall perceived health, as well as in specific health conditions[39]. Two other limitations associated with the characteristics and the size of the sample need to be acknowledged. Although the sample was collected from 15 different local PD associations, to some extent representative of the Spanish context, the voluntariness of the participation in the study could have introduced a selection bias. The relative small sample size does not allow separating H&YS in the four levels measured; hence this could have potentially produced systematic bias resulting from the possible differences of patients' experience. It would be interesting to conduct new studies with larger sample size which could ratify the results in this study, as well as to assess other psychometric properties, such as longitudinal response and reliability.

Conclusions

Both EQ-5D-3L and SF-6D instruments seem to be adequate generic HRQoL measures in terms of validity and sensibility with regard to indicators of health status in this Spanish PD population sample. The EQ-5D-3L descriptive system is simpler than SF-6Dbut presents higher ceiling effects between its utilities. Besides, the SF-6D was better in detecting clinical changes in severity of the symptoms of this disease with respect to EQ-5D-3L and has better efficiency and greater sensitivity to detect clinical HRQoL changes in patients with PD when the selected "cut-off" point was fixed at 21.87 PDQ-8 median. Since methods and instruments used to measure HRQoL are of great importance in economic evaluation, further research is needed to confirm our results as well as to assess other properties, such as reliability and longitudinal response.

Acknowledgements

Our gratitude extends to the following Parkinson associations: Albacete, Aragón, Astorga, Asturias, Bahía de Cádiz, Córdoba, Elche, Extremadura, Gran Canaria, Granada, Lorca, Málaga, Mallorca, Sevilla and Valencia. José María Abellán-Perpiñán and Fernando Ignacio Sánchez- Martínez gratefully acknowledge the financial support from Ministerio de Economía y Competitividad grant ECO2010-22041-C02-02 and from Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia grant 15357/PHCS/10.

References

1. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. *Lancet Neurol*. 2006;5(6):525-35.
2. Lees AJ, Hardy J, Revesz T. Parkinson's disease. *Lancet*. 2009;373(9680):2055-66.
3. Haapaniemi TH, Sotaniemi KA, Sintonen H, Taimela E. The generic 15D instrument is valid and feasible for measuring health related quality of life in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2004;75(7):976-83.
4. Samii A, Nutt JG, Ransom BR. Parkinson's disease. *Lancet*. 2004;363(9423):1783-93.
5. Klockgether T. Parkinson's disease: clinical aspects. *Cell Tissue Res*. 2004;318(1):115-20.
6. Yousefi B, Tadibi V, Khoei AF, Montazeri A. Exercise therapy, quality of life, and activities of daily living in patients with Parkinson disease: a small scale quasi-randomised trial. *Trials*. 2009;10:67.
7. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA*. 1995;273(1):59-65.
8. Voelker R. Parkinson disease guidelines aid diagnosis, management. *JAMA*. 2006;295(18):2126-8.
9. Schapira AH. Treatment options in the modern management of Parkinson disease. *Arch Neurol*. 2007;64(8):1083-8.
10. Pearce RK, Heikkila M, Linden IB, Jenner P. L-dopa induces dyskinesia in normal monkeys: behavioural and pharmacokinetic observations. *Psychopharmacology* (Berl). 2001;156(4):402-9.

11. Bezard E, Brotchie JM, Gross CE. Pathophysiology of levodopa-induced dyskinesia: potential for new therapies. *Nat Rev Neurosci*. 2001;2(8):577-88.
12. Siegfried J, Lippitz B. Bilateral chronic electrostimulation of ventroposterolateral pallidum: a new therapeutic approach for alleviating all parkinsonian symptoms. *Neurosurgery*. 1994;35(6):1126-9; discussion 9-30.
13. Limousin P, Pollak P, Benazzouz A, Hoffmann D, Le Bas JF, Broussolle E, et al. Effect of parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. *Lancet*. 1995;345(8942):91-5.
14. Krack P, Pollak P, Limousin P, Hoffmann D, Xie J, Benazzouz A, et al. Subthalamic nucleus or internal pallidal stimulation in young onset Parkinson's disease. *Brain*. 1998;121 (Pt 3):451-7.
15. Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jonsson B. The economic cost of brain disorders in Europe. *Eur J Neurol*. 2012;19(1):155-62.
16. Dodel R. The Economic Burden of Parkinson's Disease. European Neurological Review. 2008;3(2 suppl):11-4.
17. Marchesini G, Bianchi G, Amodio P, Salerno F, Merli M, Panella C, et al. Factors associated with poor health-related quality of life of patients with cirrhosis. *Gastroenterology*. 2001;120(1):170-8.
18. Karlsen KH, Tandberg E, Arsland D, Larsen JP. Health related quality of life in Parkinson's disease: a prospective longitudinal study. J Neurol Neurosurg Psychiatry. 2000;69(5):584-9.
19. Marinus J, Ramaker C, van Hilten JJ, Stiggelbout AM. Health related quality of life in Parkinson's disease: a systematic review of disease specific instruments. *J Neurol Neurosurg Psychiatry*. 2002;72(2):241-8.
20. Fitzpatrick R, Jenkinson C, Peto V, Hyman N, Greenhall R. Desirable properties for instruments assessing quality of life: evidence from the PDQ-39. *J Neurol Neurosurg Psychiatry*. 1997;62(1):104.
21. Peto V, Jenkinson C, Fitzpatrick R, Greenhall R. The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. *Quality of Life Research*. 1995;4(3):241-8.
22. Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The PDQ-8: development and validation of a short-form Parkinson's disease questionnaire. *Psychology and Health*. 1997;12(6):805-14.
23. 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.
24. Maynard A, Bloor K. Our certain fate: rationing in health care: OHE; 1998.
25. Younossi ZM, Guyatt G. Quality-of-life assessments and chronic liver disease. *Am J Gastroenterol*. 1998;93(7):1037-41.
26. Zárate V. Evaluaciones económicas en salud: Conceptos básicos y clasificación. Revista médica de Chile. 2010;138:93-7.
27. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care*. 1989;5(4):559-75.
28. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(16):1339-41.
29. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708-15.
30. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21(2):271-92.
31. Kharroubi SA, Brazier JE, Roberts J, O'Hagan A. Modelling SF-6D health state preference data using a nonparametric Bayesian method. *J Health Econ*. 2007;26(3):597-612.
32. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ*. 2004;13(9):873-84.
33. Abellan Perpinan JM, Sanchez Martinez FI, Martinez Perez JE, Mendez I. Lowering the 'floor' of the SF-6D scoring algorithm using a lottery equivalent method. *Health Econ*. 2012;21(11):1271-85.
34. Schrag A, Selai C, Jahanshahi M, Quinn NP. The EQ-5D--a generic quality of life measure-is a useful instrument to measure quality of life in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2000;69(1):67-73.
35. Luo N, Low S, Lau PN, Au WL, Tan LC. Is EQ-5D a valid quality of life instrument in patients with Parkinson's disease? A study in Singapore. *Ann Acad Med Singapore*. 2009;38(6):521-8.
36. Garcia-Gordillo M, del Pozo-Cruz B, Adsuar JC, Sanchez-Martinez FI, Abellan-Perpinan JM. Validation and comparison of 15-D and EQ-5D-5L instruments in a Spanish Parkinson's disease population sample. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2014;23(4):1315-26.
37. Rasanen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryynanen OP, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. *Int J Technol Assess Health Care*. 2006;22(2):235-41.
38. EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. 1990.
39. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37(1):53-72.
40. Froberg DG, Kane RL. Methodology for measuring health-state preferences--II: Scaling methods. *J Clin Epidemiol*. 1989;42(5):459-71.
41. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey: Centre for Health Economics; 1995.
42. Williams A. The measurement and valuation of health: a chronicle: University of York York; 1995.
43. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35(11):1095-108.
44. 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* (Barc). 1999;112 Suppl 1:79-85.
45. Lam CL, Brazier J, McGhee SM. Valuation of the SF-6D Health States Is Feasible, Acceptable, Reliable, and Valid in a Chinese Population. *Value Health*. 2008;11(2):295-303.
46. Ferreira LN, Ferreira PL, Pereira LN, Brazier J, Rowen D. A Portuguese value set for the SF-6D. *Value Health*. 2010;13(5):624-30.
47. Alonso J, Prieto L, Anto JM. [The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results]. *Med Clin* (Barc). 1995;104(20):771-6.
48. Rebollo P, Moris J, Ortega T, Valdes C, Ortega F. [Estimating utility values for health status using the Spanish version of the SF-36. Validity of the SF-6D index vs EQ-5D]. *Med Clin* (Barc). 2007;128(14):536-7.
49. 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.
50. Martinez-Martin P, Benito-Leon J, Alonso F, Catalan MJ, Pondal M, Zamarbide I, et al. Quality of life of caregivers in Parkinson's disease. *Qual Life Res*. 2005;14(2):463-72.
51. Luo N, Ng WY, Lau PN, Au WL, Tan LC. Responsiveness of the EQ-5D and 8-item Parkinson's Disease Questionnaire (PDQ-8) in a 4-year follow-up study. *Qual Life Res*. 2010;19(4):565-9.
52. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology*. 1967;17(5):427-42.
53. Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord*. 2004;19(9):1020-8.
54. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd. edit. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

55. Kontodimopoulos N, Pappa E, Papadopoulos AA, Tountas Y, Niakas D. Comparing SF-6D and EQ-5D utilities across groups differing in health status. *Qual Life Res*. 2009;18(1):87-97.
56. Fayers PM. Quality of Life: the Assessment, Analysis and Interpretation of Patient-Reported Outcomes. John Wiley & Sons,. 2007.
57. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-10.
58. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol*. 1995;48(11):1369-78.
59. Luo N, Tan LC, Zhao Y, Lau PN, Au WL, Li SC. Determination of the longitudinal validity and minimally important difference of the 8-item Parkinson's Disease Questionnaire (PDQ-8). *Mov Disord*. 2009;24(2):183-7.
60. O'Brien BJ, Spath M, Blackhouse G, Severens JL, Dorian P, Brazier J. A view from the bridge: agreement between the SF-6D utility algorithm and the Health Utilities Index. *Health economics*. 2003;12(11):975-81.
61. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. *Med Care*. 2003;41(7):791-801.
62. Sach TH, Barton GR, Jenkinson C, Doherty M, Avery AJ, Muir KR. Comparing cost-utility estimates: does the choice of EQ-5D or SF-6D matter? *Med Care*. 2009;47(8):889-94.
63. Lubetkin EI, Jia H, Franks P, Gold MR. Relationship among sociodemographic factors, clinical conditions, and health-related quality of life: examining the EQ-5D in the U.S. general population. *Qual Life Res*. 2005;14(10):2187-96.