

Mapping CushingQoL Scores onto SF-6D Utility Values in Patients with Cushing's Syndrome

Montse Roset · Xavier Badia · Anna Forsythe · Susan M. Webb

Published online: 3 April 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract

Objectives To construct a prediction model of preference-adjusted health status (SF-6D) for Cushing's syndrome using a disease-specific health-related quality of life (HRQOL) measure (CushingQoL).

Methods Data were obtained from the original multicenter, multinational study to validate the CushingQoL questionnaire. HRQOL was measured using the CushingQoL and the SF-36 questionnaires. SF-6D scores were calculated from responses on the SF-36. Sociodemographic and clinical data were also collected. Various predictive models were tested and the final one was selected on the basis of four criteria: explanatory power, consistency of estimated coefficients, normality of prediction errors, and parsimony.

Results For the mapping analysis, data were available from 116 of the 125 patients included in the original validation study. Mean (SD) age was 45.3 (13.1) years and the sample was predominantly (83 %) female. Patients had a mean (SD) CushingQoL score of 52.9 (21.9), whereas the SF-6D (derived from SF-36) was skewed towards better health with

a mean of 0.71 (median 0.74) on a scale of -0.704 to 1. Of the various models tested, a model which included the intercept (0.61), CushingQoL overall score, level one in CushingQoL item 2 (always have pain preventing me from leading a normal life), and level one in CushingQoL item 10 (my illness always affects my everyday activities) best met the four criteria for model selection. The model had an adjusted R^2 of 0.60 and a root mean square error of 0.084.

Conclusions Although the mapping function finally selected appears to be able to accurately map CushingQoL scores onto SF-6D outcomes at the group level, further testing is required to validate the model in independent patient samples.

On behalf of the CushingQoL Development Group, the members of which are given in the “Appendix”.

M. Roset (✉) · X. Badia
Health Economics and Outcomes Research, IMS Health,
C/Dr. Ferran 25-27, 2º, 08034 Barcelona, Spain
e-mail: mrosset@es.imshealth.com

A. Forsythe
Global Health Economics and Market Access, Novartis
Oncology, New York, USA

S. M. Webb
Endocrinology/Medicine Departments, Hospital Sant Pau,
IIB-Sant Pau, Centro de Investigación Biomédica en Red de
Enfermedades Raras (CIBERER Unit 747), ISCIII; Universitat
Autònoma de Barcelona (UAB), Barcelona, Spain

Key Points for Decision Makers

- Index values produced by preference-based measures are essential for cost–utility analysis, where they are used in combination with the time in a given health state to generate quality-adjusted life years (QALYs).
- A mapping approach allows one to carry out cost–utility measurements even when a preference-based instrument is not used in the initial evaluation of an intervention.
- This study allows one to derive a simple prediction model to map scores on the disease-specific CushingQoL, for use in patients with Cushing's syndrome, to the SF-6D.
- Although the mapping function finally selected appears to be able to accurately map CushingQoL scores onto SF-6D outcomes at the group level, further testing is required to validate the model in independent patient samples.

1 Introduction

Endogenous Cushing's syndrome (CS) results from chronic exposure to excess glucocorticoids (GC) produced by the adrenal cortex. It is caused by excess adrenocorticotrophic hormone (ACTH) production (80–85 %), usually by a pituitary corticotroph adenoma [Cushing's disease (CD)], less frequently by an extrapituitary tumor (ectopic ACTH syndrome), or very rarely by a tumor secreting corticotropin-releasing hormone (CRH) (ectopic CRH syndrome). CS can also be ACTH-independent (15–20 %) when it results from excess secretion of cortisol by unilateral adrenocortical tumors, either benign or malignant, or by bilateral adrenal hyperplasia or dysplasia [1]. The condition is more prevalent in females, who are at three times greater risk of having the condition than males. The incidence of CS ranges from 0.7 to 2.4 per million population per year [2]. New data, however, suggest that CS is more common than previously thought. In screening studies of obese patients with type 2 diabetes, reported prevalence of CS is between 2 and 5 % [3]. Moreover, CS may be present in adrenal incidentalomas [2].

The condition leads to many symptoms and disorders which can impact negatively on the patient's quality of life, including central obesity, gonadal dysfunction, hirsutism, delayed wound healing, muscle weakness, hypertension, hyperglycemia, osteoporosis, and depression, among others [1]. Studies have shown that even patients who have been cured of the disease score lower in terms of general well-being, anxiety and depression, and overall quality of life than healthy controls [4]. Successful treatment of CD often ameliorates clinical symptoms and leads to an improvement in health-related quality of life (HRQOL). Published studies have reported significant improvements in patients' postoperative physical and mental functioning [5, 6].

For that reason, HRQOL instruments are useful for assessing the burden of disease in patients with the syndrome as well as for evaluating the outcomes of specific interventions. HRQOL instruments are usually self-report measures and can generally be classified as either generic (suitable for use over a wide range of conditions or illnesses, as well as in the general population) or disease-specific (instruments for use in a given condition or illness) [7]. A further category of HRQOL instruments are those designed to collect preferences (or utilities) from patients or other groups, including the general population [8]. Examples include the SF-6D [9], the EQ-5D [10], and the Health Utilities Index [11].

Index values produced by preference-based measures are essential for cost–utility analysis, where they are used in combination with the time in a given health state to generate quality-adjusted life years (QALYs) [12]. In cases where a preference-based measure of this type has not been

included in the initial data collection, but a disease-specific measure has, it is sometimes possible to create a preference function which allows scores on the disease-specific measure to be 'mapped' to index values on the preference-based instrument. If mapping is successful, this approach can allow cost–utility measurements to be carried out even when a preference-based instrument was not used in the initial evaluation of an intervention. A recent review indicated that, although still a relatively new field of research, this type of mapping exercise has been carried out in a range of disease areas and using different preference-based measures [13]. None of the studies reviewed had been carried out in patients with CS, however.

The aim of the present study was to construct a prediction model of preference-adjusted health status (SF-6D) for non-malignant CS using a disease-specific HRQOL measure (CushingQoL).

2 Methods

2.1 Study Sample and Data Collection

Data used in the present analysis were collected in the original validation study of the CushingQoL [14]. That study was performed in 125 patients aged 18 years or above with histologically determined CS of pituitary or adrenal origin, or whose hypercortisolism disappeared after adrenal or pituitary surgery. Data were collected by 14 investigators in 5 European countries (Spain, France, Germany, the Netherlands, and Italy) over a 2-month period from August to October 2006. Data were collected at a single visit from the medical records and through self-report on the relevant HRQOL questionnaires.

2.2 Instruments and Other Variables

HRQOL was measured using the CushingQoL [14] and the Short Form-36 (SF-36) (licensed by Quality Metrics) [15] questionnaires.

2.2.1 CushingQoL

The CushingQoL is a disease-specific questionnaire designed to assess HRQOL in CS. It is a self-reported instrument consisting of 12 questions which cover the areas of trouble sleeping, wound healing/bruising, irritability/mood swings/anger, self-confidence, physical changes, ability to participate in activities, interactions with friends and family, memory issues, and future health concerns. Content for the questionnaire was derived from interviews with ten patients with the condition [16]. Patients respond on unipolar rating scales with five response categories

(‘Always’, ‘Often’, ‘Sometimes’, ‘Rarely’, and ‘Never’, or ‘Very much’, ‘Quite a bit’, ‘Somewhat’, ‘Very little’, and ‘Not at all’). Responses are scored on a scale of 1–5, where ‘1’ corresponds to ‘Always’ or ‘Very much’ and ‘5’ to ‘Never’ or ‘Not at all’. The overall score is calculated by summing responses on all items and ranges from 12 (worst HRQOL) to 60 (best HRQOL). To facilitate the interpretation of scores, they can be standardized on a scale from 0 (worst HRQOL) to 100 (best HRQOL).

2.2.2 SF-36

The SF-36 is a 36-item instrument designed to measure health status in a range of populations, including the general population. It measures eight dimensions of health: physical functioning (10 items), role limitations because of physical problems (4 items), bodily pain (2 items), general health (5 items), energy/vitality (4 items), social functioning (2 items), role limitations caused by emotional problems (3 items), and mental health/emotional well-being (5 items). Each dimension is scored from 0 “poor health” to 100 “optimal health” [15]. As well as scores for the individual dimensions, the SF-36 generates two summary measures, the Physical Component Summary (PCS) scale and the Mental Component Summary (MCS) scale [17, 18].

2.2.3 SF-6D

The SF-6D is a preference-based measure derived from the SF-36 [9]. It covers the six domains of physical functioning, role limitation, social functioning, pain, mental health, and vitality. Each domain has 4–6 levels of severity and an SF-6D health state is defined when a respondent selects one level of severity in each domain. The SF-6D can define a total of 18,000 health states. Preference-based weights for the SF-6D were obtained from a representative sample of the UK population using the standard gamble technique [19]. SF-6D preference weights can be generated in any study in which the SF-36 or SF-12 instruments are used.

2.2.4 Other Variables

Sociodemographic data (age, gender, level of studies, and current employment status) and the following clinical variables were collected: weight, height, blood pressure, date of diagnosis of CS and cause (pituitary or adrenal adenoma), history and persistence or not of adrenal insufficiency and hypercortisolism, surgery undergone for the disease (type, date, route, and results of histology), and history, dose, and date of pituitary radiotherapy.

2.3 Model Development and Selection

Before a predictive model could be developed, it was necessary to transform SF-36 scores to SF-6D scores. This was done using the items from the SF-36 corresponding to each dimension on the SF-6D [9]. In order to calculate the utility weights for this study, we used model 10, as described by Brazier et al. [18]; coefficients in this model run from 0.291 to 1, with 1 representing full health and 0 representing death.

Regression analysis was used to analyze the relationship between the SF-6D utility score and the CushingQoL score. In all models, the dependent variable was the SF-6D utility score. Models were additive generalized linear models incorporating main effects. The simplest reference model was one which included only the SF-6D utility scores and CushingQoL scores. Clinical and sociodemographic variables that showed a statistical relationship with CushingQoL scores (presence of depression and hospitalizations during the previous year) were also included in the model. Individual CushingQoL items and categorizations of CushingQoL scores (level 1 or 2 vs. other responses) were then also included in the model in order to improve goodness of fit. Finally, transformations (logarithm or square root) were applied to the CushingQoL and interactions and/or quadratic terms were tested as predictors.

Clinical and sociodemographic variables were initially tested for potential inclusion in the models by determining whether they showed a statistically significant association with SF-6D utility scores. Categorical variables were analyzed using the analysis of variance and continuous variables using the Pearson’s correlation coefficient. Variables tested were age, gender, level of education, age at diagnosis, time since diagnosis, body mass index (BMI), type of CS (pituitary-dependent or adrenal adenoma), adrenal insufficiency development, previous surgery for CS, presence of any concomitant disease, hypertension, diabetes mellitus, osteoporosis, osteopenia, depression, hormone levels, and hospitalizations related to CS or its complications during the previous year.

With regard to the CushingQoL itself, as well as the overall score, we also tested each item individually in the models, by including them as discrete dummy variables (always vs. other response options). Items included were those significant at the 0.01 level in bivariate analysis. We also tested the following categorizations of CushingQoL scores by including them as dummy variables: presence of ‘1’ in any of the items answered; presence of ‘5’ in any of the items answered; overall score at most 20, between 21 and 40, between 41 and 60, between 61 and 80, and greater than 80.

Analyses were performed using SAS[®] (PROC REG and PROC GLM) and four criteria were used to select the final model, namely the model's explanatory power (assessed using adjusted R^2), the consistency of the estimated coefficients (sign and parameter estimation), normality of prediction errors, and simplicity. The normality of the prediction errors was assessed using mean error (ME), mean absolute error (MAE), root mean squared error (RMSE), and a percentage error under 5, 10, and 15 % of the overall scale of independent variable. The model's simplicity was evaluated by determining whether predictors were readily available and use of the minimum number of predictor variables in the model. The criterion of simplicity was important in order to optimize model usability. It has also been pointed out that in general in this type of modeling exercise, simple additive models performed almost as well as more complex models with greater complexity providing little extra advantage [13].

3 Results

A total of 125 patients were included in the original CushingQoL validation study. Table 1 shows the sample characteristics. The patients were included from five countries in roughly equal proportions. Mean [standard deviation (SD)] age was 45.3 (13.1) years and a large majority (83 %) of the sample were female. The vast majority of the sample had pituitary-dependent CS (85 %), with only 18 patients (15 %) having cortisol-secreting adrenal adenoma. Most of the sample (84 %) had also received prior surgery and most (80 %) had at least one co-morbid condition. The most frequent co-morbidity was osteopenia or osteoporosis (34 % of overall sample). For the mapping analysis presented here, we were able to use data from 116 of the original 125 patients. The remaining patients had to be excluded because of missing scores on the SF-36 or CushingQoL.

Table 2 shows the score distributions on the CushingQoL and SF-6D. Whilst the CushingQoL showed a mean score (52.9) at approximately the mid-point on the scale, the SF-6D showed a noticeably skewed distribution towards better health with a mean of 0.71 on a scale of -0.704 to 1, and minimum and maximum scores of 0.37 and 0.95.

Figure 1 shows the relative distribution of CushingQoL and SF-6D values in the basic, two-variable regression model. The correlation between the two scores, as shown in the model which only included the SF-6D as an independent variable, was 0.68.

Table 3 shows the results of the most promising models tested alongside the basic reference model (model 1). In bivariate analyses, only depression and hospitalization over

the previous year showed a statistically significant correlation with SF-6D utility values (p values of 0.002 and 0.01, respectively). Hypertension and presence of any co-morbidity almost achieved statistical significance with p values of 0.05 and 0.06, respectively, but were not included for further testing, as the threshold for inclusion was a p value of 0.01 or under. When depression and hospitalization over the previous year were included in different regression models, only depression improved model fit and was retained in the different models tested. We also observed that seven of the dummy variables related to CushingQoL scores showed a statistically significant relationship with SF-6D (Table 3). In terms of R^2 and adjusted R^2 , the best-performing model (model 2) of those tested incorporated five of the seven dummy variables related to CushingQoL scores as well as depression.

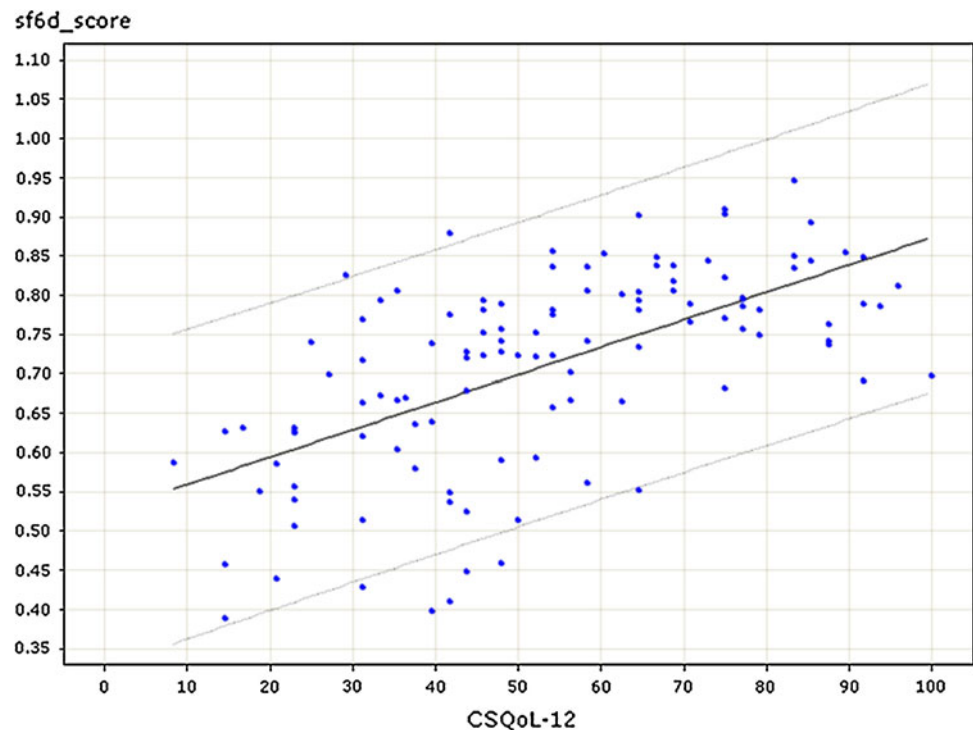
Table 1 Sociodemographic and clinical characteristics of the original validation study sample

Characteristic	<i>N</i> = 125
Country, <i>n</i>	
Spain	29
France	26
Germany	24
Netherlands	21
Italy	25
Mean (SD) age in years	45.3 (13.1)
Sex, female, <i>n</i> (%)	104 (83 %)
Education, secondary or university studies, <i>n</i> (%)	99 (79 %)
BMI, mean (SD)	27.8 (6.0)
Time since diagnosis in months, mean (SD)	60.8 (67.9)
Clinical type, <i>n</i> (%)	
Pituitary dependent	107 (85 %)
Cortisol-secreting adrenal adenoma	18 (15 %)
Hypercortisolemic at time of visit, <i>n</i> (%)	39 (31 %)
Receiving pharmacological treatment for CS, <i>n</i> (%)	28 (22 %)
Prior surgery, <i>n</i> (%)	105 (84 %)
Adrenal insufficient at time of visit, <i>n</i> (%)	47 (38 %)
Concomitant diseases, <i>n</i> (%)	
Any	100 (80 %)
Osteopenia/osteoporosis	42 (34 %)
Hypertension	40 (32 %)
Depression	28 (22 %)
Diabetes mellitus	20 (16 %)
CushingQoL score	
Mean (SD)	52.92 (21.92)
Median (IQR)	39.60 (29.20)
SF-6D utility values	
Mean (SD)	0.708 (0.132)
Median (IQR)	0.736 (0.161)

CS Cushing's syndrome, SD standard deviation, IQR interquartile range

Table 2 Score distributions on the CushingQoL and SF-6D

Values	Mean	Theoretical range	Standard deviation	Minimum	Median	Maximum	Valid <i>N</i>
CushingQoL	52.9	0 to 100	21.9	8.3	51.0	100.00	124
SF-6D	0.71	-0.704 to 1	0.13	0.37	0.74	0.95	116

Fig. 1 Correlation between CushingQoL scores and SF-6D utility values in the simplest reference model

However, the model finally selected (model 4) achieved very similar goodness of fit statistics with only three variables, and therefore more fully met the requirement for parsimony. This model was also selected because it included only independent variables derived from CushingQoL, and no additional variables were required. This model also considerably improved on the fit obtained with the simplest reference model with an adjusted R^2 of 0.60 compared to 0.46 for the basic model. Very similar results were obtained with models which incorporated interaction and quadratic terms and these were therefore not selected as they were considered to unnecessarily complicate the model.

Table 4 shows the results of the analysis of residuals in the selected model. Only minimal differences were observed between observed and estimated mean and median values. Error terms were obviously larger for maximum and minimum values because of the much smaller number of patients scoring at the extremes. Prediction errors showed a normal distribution, according to the Shapiro–Wilks test ($p = 0.216$). Estimated utility scores showed a minimum error of 10 %, compared with

observed utility values, in 32.8 % of the study patients. In comparison to observed values, the figures for maximum and minimum scores as well as for the first and third quartiles indicate some compression in estimated values at the extremes. The correspondence between predicted and observed values is also shown in Fig. 2. Figure 3 shows the distribution of observed utility values and predicted values according to self-perceived health status using the final model.

4 Discussion

In the present study, we were able to derive a simple prediction model to map scores on the disease-specific CushingQoL, for use in patients with CS, to the SF-6D. The model finally selected was a parsimonious model which achieved acceptable goodness of fit with only three variables (the CushingQoL overall score, a level 1 score in CushingQoL item 2, or a level 1 score in CushingQoL item 10). A model which included only variables derived from the CushingQoL questionnaire was selected in order to

Table 3 Comparison of results obtained with the most promising models and the reference model (model 1)

Parameter	Model 1	Model 2	Model 3	Model 4
Intercept	0.48766	0.81962	0.80097	0.60682
CushingQoL score	0.00411			0.00259
CushingQoL 21–40		−0.10069		
CushingQoL 41–60		−0.08230		
Level 1 or 2 in Q-1			−0.05084	
Level 1 in Q-2		−0.11696	−0.13194	−0.09419 ^a
Level 1 in Q-9		−0.10205	−0.09192	
Level 1 in Q-10		−0.08872	−0.09476	−0.08354 ^b
Level 2 in Q-12			−0.05366	
Depression		−0.04239		
Model fit statistics				
R^2	0.4685	0.6553	0.5944	0.6074
Adjusted R^2	0.4639	0.6363	0.5760	0.5969
RMSE	0.0969	0.0798	0.0861	0.0840

Q question or item from CushingQoL questionnaire, RMSE root mean square error

^a I always have pain preventing me from leading a normal life

^b My illness always affects my everyday activities

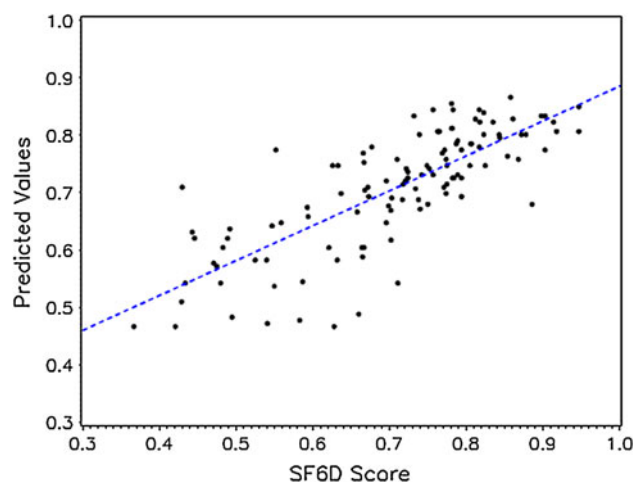
reduce the number of data required to obtain the corresponding utility value. This model best met the four predefined criteria for model selection, namely explanatory power, consistency of estimated coefficients, normality of prediction errors, and simplicity. The final model had an adjusted R^2 of 0.60 and an RMSE of 0.084.

Although this type of mapping is a relatively new field of research, in a recent review Brazier et al. identified 28 published mapping studies carried out to date [13]. A number of those mapped from generic HRQOL measures, particularly the SF-36 and SF-12, to a preference-based measure, but a substantial number mapped from disease-specific measures, including instruments used in asthma, rheumatoid arthritis, osteoarthritis, overactive bladder, irritable bowel syndrome, intermittent claudication, dental, dyspepsia, obesity, cancer, and heart disease.

Table 4 Analysis of residuals in the final model

	Observed SF-6D score	Predicted SF-6D score	Error term	Absolute error term	Individual % absolute error
Mean	0.71	0.71	−0.0005	0.0645	10.16
SD	0.13	0.10	0.0829	0.0518	10.12
Minimum	0.37	0.47	−0.2799	0.0002	0.03
Q1	0.63	0.65	−0.0531	0.0232	3.22
Median	0.74	0.73	0.0100	0.0577	7.76
Q3	0.79	0.79	0.0588	0.0901	12.81
Maximum	0.95	0.87	0.2059	0.2799	65.10
Valid N	116	116	116	116	116

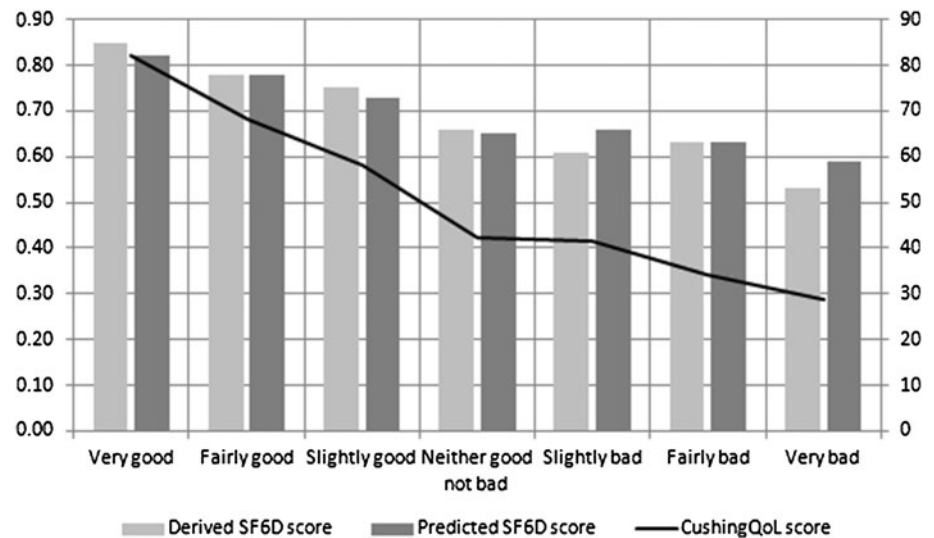
Q1 quartile 1, Q3 quartile 3, SD standard deviation

**Fig. 2** Observed SF-6D utility values and predicted values using the final model

However, this is the first study we are aware of to report mapping from a disease-specific measure for CS to a preference-based measure (SF-6D). This type of mapping exercise is useful in deriving a prediction model to generate preference-based index values which can then be used to, for example, populate an economic model when a preference-based measure was not included in the original trial or study. This might be the case, for example, if it was felt that including the SF-36 would overburden patients in a trial.

The results of testing model performance were generally satisfactory. The adjusted R^2 of 0.60 for the final model would be within the upper range of the various condition-specific to generic mapping exercises reported by Brazier et al., which indicated that only a relatively small proportion of this type of model achieved an adjusted R^2 over 0.60 [13]. The RMSE was 0.084, which again is at the lower end of the spectrum of values reported by Brazier et al. in their review of similar mapping exercises. This represented a percentage error of 10.2 % of the overall scale on the dependent variable.

Fig. 3 Observed utility values and predicted values by self-perceived health status using the final model



The results also met our criterion of consistency of estimated coefficients, in that the presence of a positive response at level 1 on CushingQoL items 2 or 10 also led to a more negative preference-value on the SF-6D. A level 1 response on these items indicates that respondents “always have pain preventing [them] from leading a normal life” and that their “illness always affects [their] everyday activities”. They are therefore clearly items with a substantial impact on respondents’ quality of life, a point which is highlighted further by the fact that they are the only individual CushingQoL items included in the final regression model. Despite their severity, a total of 19/116 (16.4 %) respondents reported level 1 on item 2 and 31/116 (26.7 %) reported level 1 on item 10 of the CushingQoL. Twelve of 116 (10.3 %) reported level 1 on both of these items. Prediction errors also showed a normal distribution, as indicated by the results of the Shapiro–Wilks test, and we consider the model to meet the criterion of parsimony. We found that adding interaction or quadratic terms did little to improve model performance, which is in line with Brazier et al.’s comment in their review of mapping studies that only “quite modest or negligible improvements were achieved from increasing model complexity” [13]. Finally, it should be noted that the prediction model derived here is suitable for use only at aggregate or group level, and should not be used to predict individual scores on the SF-6D. The prediction model and the corresponding error are assessed at aggregate level, not for individual scores.

One area of interest in developing this type of mapping exercise is how variables for the model are selected. Although the most frequent approach seems to be that of testing a wide range of possible variables for the model and including those which, in bivariate testing, show some association with the dependent variable, there may also be an argument for taking a more deterministic, theory-based

approach. However, our experience of this approach when developing models for other utility-based instruments, as well as reports in the literature [20], suggest that it does not lead to better-fitting models. In this study we therefore used the more standard approach. One concern with this approach is that, by testing a wide range of possible variables, it might magnify chance effects of finding significant predictors in the data. However, in the present case, the relationship of each CushingQoL item (considering four response options and merging options 1 and 2) with utility values was analyzed and those items with a significance level less than 0.01 were included in the regression model. The CushingQoL questionnaire only includes 12 questions and the individual significance level selected was 0.01. Given that the possibility of identifying significant variables is directly related to the number of tested variables, if we analyze the relationship with 24 variables obtained from individual items then the increase in the likelihood of obtaining significant variables exists, but is very small (just in 0.25 %).

4.1 Study Limitations

The study had a number of limitations. One was that we did not perform a cross-validation test in another sample, or in half of the original sample. Although this approach has the advantage of testing the prediction model in a sample other than the one used to actually develop the model, it was not practical here because the sample size was insufficient. This was due to the difficulty of recruiting patients with a clearly rare condition. Sample size calculations also indicated that for one independent variable in the final model, we would require a minimum of 63 patients, and for three independent variables, a minimum sample of 105 patients [21]. On the other hand, Brazier et al. also found in their

review of mapping studies that, in general, “out-of-sample tests found little reduction, if any, in the performance of the models” [13]. This is not always the case, though; for example, Wu et al. found that when mapping other specific HRQOL questionnaires (FACT-P and EORTC QLQ-C30) to EQ-5D scores their model predicted only 58.2 % of the observed EQ-5D variation in the cross-validation sample compared to 73.2 % in the development sample [21]. Until further testing in a different sample can be carried out, however, the model presented here should be considered provisional; it is nevertheless a useful first step in providing an algorithm for mapping from CushingQoL to SF-6D.

The small sample size in the present study also meant that we could not test the model in different patient subgroups. For example, only 18 patients (15 %) had cortisol-secreting adrenal adenoma, whereas the vast majority had pituitary-dependent CD. It would be useful to be able to test the performance of the model in other relevant subgroups. Finally, the distribution of scores was skewed on the SF-6D. In general, scores were skewed towards better health; thus, although there is less discriminative capacity, the fact that utility values are compressed within a relatively small range will likely mean that model fit and levels of residual error are improved.

In terms of mean age, percentage of female patients, and etiologic groups, the profile of CS patients included in the study was quite similar to patients included in the European Registry on Cushing’s syndrome (ERCUSYN) [22]. The percentage of patients with biochemically cured disease comprised nearly two-thirds of the included subjects, which may explain the high utility values observed here. The mapping function obtained should be tested in a sample of patients with more active disease in order to assess its goodness of fit in that patient profile.

5 Conclusions

Although the mapping function finally selected appears to be able to accurately map CushingQoL scores onto SF-6D outcomes at the group level, further testing is required to validate the model in independent patient samples. It would also be of interest to test whether the model performs equally well in samples with a different mix of patient types or in specific patient subgroups.

Acknowledgements Montse Roset and Xavier Badia have no conflict of interest and were involved in the conception and planning of the work, statistical analysis, interpretation of the data, and the preparation of the manuscript. Anna Forsythe is a Novartis Pharmaceuticals employee. Anna Forsythe was involved in the conception and planning of the work, interpretation of the data, and the critical revision of the manuscript. Dr. Susan M. Webb received a fee for scientific and clinical assessment in the project and was involved in

conception and planning of the work, interpretation of the data, and the preparation of the manuscript. Other members of the CushingQoL Development Group have no conflict of interest and were involved in interpretation of the data and the clinical review of the manuscript. All the authors approved the final submitted version of the manuscript. Montse Roset will act as overall guarantor. This study was supported by an unrestricted grant from Novartis Oncology.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited. The exclusive right to any commercial use of the article is with Springer.

Appendix

CushingQoL Development Group

MJ Barahona: Department of Endocrinology, Hospital Mútua de Terrassa, Pl Dr Robert 5, 08221 Terrassa, Barcelona, Spain; CJ Strasburger: Division of Clinical Endocrinology, Department of Medicine for Endocrinology, Diabetes and Nutritional Medicine Charité –Universitätsmedizin, Campus Mitte Charitéplatz 1, 10117 Berlin, Germany; A Tabarin: Department of Endocrinology, University of Bordeaux, 2 CHU Haut Leveque, 33604 Pessac, France; MO van Aken: Department of Internal Medicine, Haga Hospital, Leyweg 275, 2545 CH the Hague, the Netherlands; GK Stalla: Department of Endocrinology, Max Planck Institute of Psychiatry, Kraepelinstr. 10, 80804 Munich, Germany.

References

1. Arnaldi G, Angeli A, Atkinson AB, et al. Diagnosis and complications of Cushing’s syndrome: a consensus statement. *J Clin Endocrinol Metab.* 2003;88:5593–602.
2. Newell-Price J, Bertagna X, Grossman AB, et al. Cushing’s syndrome. *Lancet.* 2006;367:1605–17.
3. Tabarin A, Perez P. Pros and cons of screening for occult Cushing syndrome. *Nat Rev Endocrinol.* 2011;7:445–55.
4. van Aken MO, Pereira AM, Biermasz NR, et al. Quality of life in patients after long-term biochemical cure of Cushing’s disease. *J Clin Endocrinol Metab.* 2005;90:3279–86.
5. Lindsay JR, Nansel T, Baid S, et al. Long-term impaired quality of life in Cushing’s syndrome despite initial improvement after surgical remission. *J Clin Endocrinol Metab.* 2006;91:447–53.
6. Thompson SK, Hayman AV, Ludlam WH, Deveney CW, Loriaux DL, Sheppard BC. Improved quality of life after bilateral laparoscopic adrenalectomy for Cushing’s disease: a 10-year experience. *Ann Surg.* 2007;245:790–4.
7. Guyatt GH. A taxonomy of health status instruments. *J Rheumatol.* 1995;22:1188–90.
8. Coons SJ, Rao S, Keininger DL, et al. A comparative review of generic quality-of-life instruments. *Pharmacoeconomics.* 2000;17:13–35.

9. Brazier J, Usherwood T, Harper R, et al. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol.* 1998;51:1115–28.
10. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med.* 2001;33:337–43.
11. Feeny D, Furlong W, Boyle M, et al. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics.* 1995;7:490–502.
12. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care.* 1989;5:559–75.
13. Brazier JE, Yang Y, Tsuchiya A, et al. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ.* 2010;11:215–25.
14. Webb SM, Badia X, Barahona MJ, et al. Evaluation of health-related quality of life in patients with Cushing's syndrome with a new questionnaire. *Eur J Endocrinol.* 2008;158:623–30.
15. Ware JE Jr, Sherbourne CD. The MOS 36-item short form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30:473–83.
16. Badia X, Barahona MJ, Glusman J, et al. Strategy for developing a specific Cushing's syndrome QoL questionnaire. In: 12th ISOQOL conference on patient reported outcomes in clinical practice. Budapest (Hungary), 2007 June 24–26 (Abstract).
17. Ware JE, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Institute; 1994.
18. Brazier J, Roberts J, Deverill M. The estimation of a preference based single index measure for health from the SF-36. *J Health Econ.* 2002;21:271–92.
19. Cheung YB, Tan LC, Lau PN, Au WL, Luo N. Mapping the eight-item Parkinson's Disease Questionnaire (PDQ-8) to the EQ-5D utility index. *Qual Life Res.* 2008;17:1173–81.
20. Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. *Stat Med.* 1998;17:1623–34.
21. Wu EQ, Mulani P, Farrell MH, et al. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. *Value Health.* 2007;10:408–14.
22. Valassi E, Santos A, Yaneva M, et al. The European Registry on Cushing's syndrome: 2-year experience. Baseline demographic and clinical characteristics. *Eur J Endocrinol.* 2011;165:383–92.