VALUE IN HEALTH 18 (2015) 299–307

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

# Mapping the Clinical Chronic Obstructive Pulmonary Disease Questionnaire onto Generic Preference-Based EQ-5D Values

*Melinde R.S. Boland[1,*], Job F.M. van Boven[2], Janwillem W.H. Kocks[3], Thys van der Molen[3], Lucas M. Goossens[1], Niels H. Chavannes[4], Maureen P.M.H. Rutten-van Mölken[1]*

[1]*Institute for Medical Technology Assessment, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands;* [2]*Unit of PharmacoEpidemiology & PharmaEconomics, Department of Pharmacy, University of Groningen, Groningen, The Netherlands;* [3]*Department of Primary Care, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands;* [4]*Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands*

ABSTRACT

**Objectives:** To develop a model to predict EuroQol five-dimensional questionnaire (EQ-5D) values from clinical chronic obstructive pulmonary disease (COPD) questionnaire (CCQ) scores. **Methods:** We used data from three clinical trials (the Randomized Clinical Trial on Effectiveness of Integrated COPD Management in Primary Care [RECODE], the Assessment Of Going Home Under Early Assisted Discharge [GO-AHEAD], and the Health Status Guided COPD Care [MARCH]). Data were randomly split into an estimation sample and a validation sample. The conceptual similarity between patient-reported CCQ and preference-based EQ-5D scores was assessed using correlation and principal-component analysis. Different types of models were estimated with increasing complexity. We selected the final models on the basis of mean absolute error and root mean square error when comparing predicted and observed values from the same population (internal validity) and from different trial populations (external validity). We also developed models for different country-specific EQ-5D value sets. **Results:** The principal-component analysis showed that the CCQ domains functional state and mental state are associated with four dimensions of the EQ-5D. The EQ-5D dimension pain/discomfort formed a separate construct on which no CCQ item loaded. The mean observed EQ-5D values were not significantly different from the mean predicted EQ-5D values in internal validation samples but did significantly differ in external validation samples. The models underestimated EQ-5D values in milder health states and overestimated them in more severe health states. The predictive ability of the models was similar across different EQ-5D value sets. **Conclusions:** The models can predict mean EQ-5D values that are similar to observed mean values in a similar population. The overestimating/underestimating of the low/high EQ-5D values, however, limits its use in Markov models. Therefore, mapping should be used cautiously.

***Keywords:*** CCQ, EQ-5D, mapping, utility.



## Introduction

Studies assessing the effectiveness of new chronic obstructive pulmonary disease (COPD) treatments commonly use disease-specific health-related quality-of-life (HRQOL) instruments such as the St. George's Respiratory Questionnaire (SGRQ) [1] and the clinical COPD questionnaire (CCQ) [2]. Preference-based HRQOL, however, measures such as the EuroQol five-dimensional questionnaire (EQ-5D) [3] are often not included in these studies. Nevertheless, data from such studies frequently form the basis of post hoc cost-utility analyses that aim to estimate the incremental costs per quality-adjusted life-year gained. To enable the calculation of quality-adjusted life-years in the absence of directly collected EQ-5D data, the SGRQ or CCQ scores from these studies need to be converted to utilities. This can be done with a model or algorithm that maps disease-specific HRQOL data to EQ-5D data. This solution has also been recommended by the National Institute for Health and Care Excellence [4] and is seen as a better approach than judgments of experts [5]. Before July 2013, 90 studies estimated 121 mapping models and reported algorithms to allow other researchers to use them to predict EQ-5D values [6].

A model to map SGRQ scores onto EQ-5D values has been published [7]. A model to predict EQ-5D values from CCQ data, however, has not been developed. It is relevant to do so because the CCQ is increasingly used, not only because it is brief and takes little time to complete [8] but also because the 2013 revision of the Global Initiative for Chronic Obstructive Lung Disease guidelines recommends the CCQ as one of the options to define

* *Address correspondence to:* Melinde R.S. Boland, Institute for Medical Technology Assessment, Institute of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands.

E-mail: boland@bmg.eur.nl.

1098-3015$36.00 – see front matter 

http://dx.doi.org/10.1016/j.jval.2014.11.006

symptom level, that is, one of the three components of the COPD classification [9]. This study aimed to develop a model to predict EQ-5D values from CCQ scores obtained in clinical trials.

## Methods

### Instruments

#### CCQ

The CCQ is an instrument to measure HRQOL in patients with COPD on three domains (symptoms, functional state, and mental state). The symptoms and functional state domains contain four items each, and the mental state domain contains two items. Patients have to respond to each item on a seven-point scale, resulting in more than 282 million possible health states. Response options on CCQ items 1 to 6 are as follows: never/hardly ever/a few times/several times/many times/a great many times/almost all the time. Response options on CCQ items 7 to 10 are as follows: not limited at all/very slightly limited/slightly limited/moderately limited/very limited/extremely limited/totally limited or unable to do. The total CCQ score and the scores of the domains are calculated by adding up the item scores and dividing this sum by the number of items, where 0 is the best and 6 is the worst score [2].

#### EQ-5D

The EQ-5D is a generic HRQOL questionnaire. It consists of five dimensions to describe the current health states of patients: mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Each dimension is measured with one item, and respondents have to respond to each item on a three-point scale (no, some, or extreme problems). This results in 243 potential health states. EQ-5D values are calculated by combining the responses on each dimension with off-the-shelf preference-based weights (i.e., a value set). In the base case, we used the Dutch EQ-5D value set to calculate the EQ-5D values of the patients with COPD. These EQ-5D values can range from −0.329 to 1, where 0 indicates a health state equivalent to dead and 1 indicates full health [3,10]. In the sensitivity analyses, we used UK and US EQ-5D value sets.

### Setting and Participants

To include a broad range of patients with different severity levels of COPD, we combined the data from three trials: the Randomized Clinical Trial on Effectiveness of Integrated COPD Management in Primary Care (RECODE), a 2-year, cluster-randomized controlled trial with 1086 patients recruited from general practice [11]; the Assessment Of Going Home Under Early Assisted Discharge (GO-AHEAD) trial, a 3-month, multicenter, randomized trial with 166 patients hospitalized for a COPD exacerbation [12]; and the Health Status Guided COPD Care (MARCH) trial, a 6-month, randomized controlled trial with 53 patients recruited from general practice [13]. Patients completed both the CCQ and the EQ-5D at two (MARCH), three (GO-AHEAD), and six (RECODE) time points.

### Conceptual Similarity

Mapping would be able to appropriately predict EQ-5D values from the CCQ only if there are no major conceptual differences between the CCQ and the EQ-5D [14]. Therefore, we first investigated the conceptual (dis)similarities between the EQ-5D dimensions and the CCQ items using Spearman rank correlations and principal-component analysis (PCA). PCA explores which questions included in the two instruments are related to each other and generate information on the same underlying construct. The CCQ may include items related to domains that are not included in the EQ-5D and hence will not be reflected in changes in the patient's EQ-5D value, and vice versa. In the explorative PCA, all constructs with an eigenvalue of more than 1 were selected [15], whereas in the confirmatory PCA the number of constructs was set to five to investigate whether these five constructs would mirror the five dimensions of the EQ-5D. Eigenvalues of a construct represent the relative share of variance accounted for by the construct. The individual items have meaningful loadings on a construct if their absolute value exceeds 0.40. After extracting the initial constructs, varimax rotation was used to improve the differentiation and the interpretability of the results [16].

### Model Development

The data set was randomly split into a sample that was used to estimate the model (50%) and a sample that was used to validate the estimated model (50%). Thereafter, different types of models were estimated with increasing complexity. We started with simple models that predicted the EQ-5D value from the total CCQ score and the CCQ domain scores. We also investigated models that included individual items of the CCQ, either as categorical variables or as dummy variables for the seven possible response options of each CCQ item. We further tested whether the model improved when including polynomial terms and patient characteristics (age, Charlson comorbidity index [17], and sex) and when the seven-point response scale of the CCQ was collapsed into a three-point scale by combining the response options 1 and 2, 3 and 4, and 5, 6, and 7.

These models were estimated using ordinary least squares (OLS) with backward selection procedures. Variables with a $P$ value of more than 0.10 were removed. To correct for multiple testing, Bonferroni corrections were applied to the models with dummy variables for the seven possible response options of each CCQ item.

Because OLS regression ignores censoring in the EQ-5D values, we investigated whether Tobit and generalized linear models performed better than OLS models [18]. The distribution and link function of these models were selected after comparing the goodness of fit of models with different specifications of the distribution and link functions (i.e., normal, inverse Gaussian, gamma, Poisson distributions in combination with log and identity link). Models that had the lowest Akaike's information criterion and Bayesian information criterion were selected [19].

To analyze the effect of using repeated measurements of the same patients as independent observations, we estimated a model on the basis of 1) the baseline data of the RECODE trial and 2) data from all measurements in the RECODE trial. In addition, we estimated models on the basis of 3) data from the GO-AHEAD trial only (because the patients in that trial are more severely ill) and 4) data from the combined RECODE, GO-AHEAD, and MARCH trials.

### Model Validation

We validated the models that were developed for each data set (i.e., data sets 1–4 as mentioned above). First, the models were used to predict EQ-5D values of patients in the validation sample (internal validation). In addition, the models that were developed using data from one trial were used to predict EQ-5D values in the other trials (external validation). Predicted EQ-5D values were compared with observed EQ-5D values. The mean absolute error (MAE) was reported for the overall range of observed EQ-5D values and for observed EQ-5D values in a specific range; that is, EQ-5D $< 0.25$, $0.25 \leq$ EQ-5D $< 0.50$, $0.50 \leq$ EQ-5D $< 0.75$, $0.75 \leq$ EQ-5D $\leq 1$. We also calculated the root mean square error (RMSE), which attaches greater weight to larger errors. A scatter plot of observed and predicted values in the validation sample was provided, and we tested whether the mean predicted EQ-5D value significantly differed from the mean observed EQ-5D value. We selected the final models on the basis of the MAE and the RMSE when comparing the predicted and observed EQ-5D values

from the same trial population (internal validity) and from different trial populations (external validity). This was done for models in which EQ-5D values were based on the Dutch value set and, in sensitivity analyses, for models in which EQ-5D values were based on the US value set and the UK value set.

Finally, we assessed the precision of the predicted mean EQ-5D value by applying a bootstrap procedure: 1) we randomly sampled patients, with replacement, to create a group of size 10, 25, 50, 100, 250, and 500; 2) we computed the mean predicted EQ-5D value and the mean predicted error for each of the six group sizes; 3) we repeated steps 1) and 2) 1000 times to generate a distribution of the group predicted error for each of the six group sizes.

## Results

### *Descriptive Statistics*

The patient characteristics are presented in Table 1. Observations with missing data on the total EQ-5D and/or CCQ score were excluded, resulting in a study population containing 5751 observations. The patients were mainly elderly (68 years), included slightly more men (55%), and had moderate airflow obstruction. The mean CCQ score was 1.66 ± 1.04, and the mean EQ-5D value was 0.75 ± 0.25, whereas the mean EQ-5D value for a representative sample of the Dutch population has been estimated at 0.89 [20]. The distribution of EQ-5D values was left-skewed: 25% of the observations were at 1 (full health), 25% were between 0.8 and 1, 25% were between 0.7 and 0.8, and 25% were between −0.3 and 0.7. The CCQ scores also showed skewness toward the severe end of the scale: 1% of the CCQ scores were at 0 (full health), 50% of the observations were between 0 and 1.5, 25% were between 1.5 and 2.3, and 25% were between 2.3 and 6. Decreased HRQOL was mainly due to the CCQ items 2, 5, 6, and 7 (Table 1 gives a description of these items), and more than 50% of the patients reported some or extreme problems in the EQ-5D dimensions mobility and pain/discomfort. The Charlson comorbidity index, a weighted sum score of the comorbid conditions of a patient, was lower (i.e., better) in patients in the GO-AHEAD trial than in patients in the MARCH trial [21]. Despite this, the rank order from best to worst mean HRQOL in the trials was MARCH, RECODE, GO-AHEAD.

**Table 1 – Patient characteristics at baseline and average EQ-5D and CCQ scores of all the time points in the different databases.**

| Characteristic | Total data | RECODE | GO-AHEAD | MARCH |
|---|---|---|---|---|
| Patients, n | 1303 | 1084 | 166 | 53 |
| Observations, n | 5751 | 5268 | 382 | 101 |
| Age (y) | 68 ± 11 | 68 ± 11 | 68 ± 11 | 64 ± 11 |
| Men, % | 55.0 | 54.0 | 61.0 | 58.5 |
| $FEV_1$, % predicted | 65.7 | 67.8 | 48.3 | 75.4 |
| $FEV_1$ ≥50% predicted | 76.2 | 77.7 | 41.3 | 94.3 |
| $FEV_1$ <50% predicted | 23.8 | 22.3 | 58.8 | 5.7 |
| Charlson comorbidity index | 2.27 ± 1.29 | 2.34 ± 1.26 | 1.75 ± 1.09 | 2.56 ± 1.83 |
| EQ-5D | 0.75 ± 0.25 | 0.75 ± 0.25 | 0.67 ± 0.25 | 0.88 ± 0.13 |
| Mobility (no/some/extreme problems), % | 48.7/50.8/0.5 | 49.7/49.9/0.5 | 30.69/67.5/1.8 | 67.3/32.7/0 |
| Self-care (no/some/extreme problems), % | 79.0/19.1/1.9 | 80.9/17.5/1.6 | 48.2/45.0/6.8 | 96.0/4.0/0 |
| Usual activities (no/some/extreme problems), % | 58.3/37.5/4.2 | 60.7/35.8/3.5 | 23.6/61.3/15.2 | 65.3/34.7/0 |
| Pain/discomfort (no/some/extreme problems), % | 47.5/41.4/11.1 | 46.3/41.9/11.8 | 57.9/37.7/4.5 | 71.3/27.7/1.0 |
| Depression/anxiety (no/some/extreme problems), % | 77.0/20.6/2.4 | 78.1/19.6/ 2.3 | 60.7/34.6/4.7 | 82.2/17.8/0 |
| CCQ* | 1.66 ± 1.04 | 1.61 ± 1.00 | 2.51 ± 1.17 | 1.10 ± 0.78 |
| Symptoms | 2.13 ± 1.18 | 2.12 ± 1.18 | 2.44 ± 1.16 | 1.68 ± 1.06 |
| Functional state | 1.69 ± 1.37 | 1.60 ± 1.30 | 3.07 ± 1.58 | 0.92 ± 0.94 |
| Mental state | 0.65 ± 1.04 | 0.59 ± 0.99 | 1.52 ± 1.39 | 0.29 ± 0.60 |
| CCQ-1 Short of breath at rest | 1.18 ± 1.31 | 1.16 ± 1.30 | 1.68 ± 1.37 | 0.64 ± 0.84 |
| CCQ-2 Short of breath doing physical activities | 2.76 ± 1.73 | 2.72 ± 1.71 | 3.54 ± 1.70 | 2.15 ± 1.62 |
| CCQ-3 Concerned about getting a cold or your breathing getting worse | 0.59 ± 1.11 | 0.53 ± 1.05 | 1.52 ± 1.55 | 0.18 ± 0.46 |
| CCQ-4 Depressed (down) because of your breathing problems | 0.71 ± 1.19 | 0.66 ± 1.14 | 1.51 ± 1.56 | 0.41 ± 0.82 |
| CCQ-5 Cough | 2.47 ± 1.58 | 2.47 ± 1.59 | 2.54 ± 1.45 | 2.14 ± 1.72 |
| CCQ-6 Produce phlegm | 2.11 ± 1.76 | 2.12 ± 1.76 | 2.00 ± 1.64 | 1.78 ± 1.78 |
| CCQ-7 Strenuous physical activities | 2.82 ± 1.85 | 2.75 ± 1.82 | 4.04 ± 1.89 | 1.92 ± 1.57 |
| CCQ-8 Moderate physical activities | 1.87 ± 1.63 | 1.78 ± 1.57 | 3.22 ± 1.84 | 1.02 ± 1.30 |
| CCQ-9 Daily activities at home | 1.16 ± 1.48 | 1.05 ± 1.36 | 2.92 ± 1.94 | 0.37 ± 0.86 |
| CCQ-10 Social activities | 0.91 ± 1.32 | 0.83 ± 1.24 | 2.14 ± 1.72 | 0.39 ± 0.91 |

*Note*. Values are means ± SD unless indicated otherwise.

CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; $FEV_1$, forced expiratory volume in 1 s, a pulmonary function measurement to indicate the severity of airflow obstruction; GO-AHEAD, Assessment Of Going Home under Early Assisted Discharge; MARCH, Health Status Guided COPD Care; RECODE, Randomized Clinical Trial on Effectiveness of Integrated COPD Management in Primary Care.

* Response options on CCQ items 1 to 6 are as follows: never (0)/hardly ever (1)/a few times (2)/several times (3)/many times (4)/a great many times (5)/almost all the time (6). Response options on CCQ items 7 to 10 are as follows: not limited at all (0)/very slightly limited (1)/slightly limited (2)/moderately limited (3)/very limited (4)/extremely limited (5)/totally limited or unable to do (6).

**Table 2 – Mapping algorithm (i.e., regression coefficients) of the recommended models.**

| Model A (Dutch value set) | | | | Model B (Dutch value set) | | | |
|---|---|---|---|---|---|---|---|
| Variable | B | SE | P | Variable | β | SE | P |
| Intercept | .24995 | .06295 | 0.0001 | Intercept | .28244 | .06506 | <0.0001 |
| Response 0 or 1 on CCQ-1 | .13716 | .06105 | 0.0259 | Response 0 or 1 on CCQ-1 | .16570 | .06267 | 0.0089 |
| Response 2 or 3 on CCQ-1 | .13146 | .05734 | 0.0230 | Response 2 or 3 on CCQ-1 | .14444 | .05742 | 0.0128 |
| Response 0 or 1 on CCQ-4 | .14506 | .05049 | 0.0046 | Response 0 or 1 on CCQ-4 | .13021 | .05083 | 0.0113 |
| Response 2 or 3 on CCQ-4 | .09468 | .05243 | 0.0726 | Response 2 or 3 on CCQ-4 | .08218 | .05255 | 0.1196 |
| Response 0 or 1 on CCQ-8 | .14845 | .05814 | 0.0115 | Response 0 or 1 on CCQ-8 | .12698 | .05897 | 0.0327 |
| Response 2 or 3 on CCQ-8 | .06910 | .03915 | 0.0793 | Response 2 or 3 on CCQ-8 | .05938 | .03927 | 0.1323 |
| Response 0 or 1 on CCQ-9 | .18815 | .04875 | 0.0002 | Response 0 or 1 on CCQ-9 | .18981 | .04844 | 0.0001 |
| Response 2 or 3 on CCQ-9 | .14164 | .03762 | 0.0002 | Response 2 or 3 on CCQ-9 | .14809 | .03755 | 0.0001 |
| Male | .06658 | .03028 | 0.0292 | Male | .08202 | .03127 | 0.0095 |
| | | | | Charlson comorbidity index | −.02427 | .01340 | 0.0717 |
| **Model C (US value set)** | | | | **Model D (UK value set)** | | | |
| Variable | β | SE | P | Variable | β | SE | P |
| Intercept | .94682 | .03204 | <0.0001 | Intercept | .95497 | .04719 | <0.0001 |
| CCQ symptoms$^2$ | −.00683 | .00213 | 0.0016 | CCQ symptoms$^2$ | −.01041 | .00314 | 0.0011 |
| CCQ functional | −.06249 | .00970 | <0.0001 | CCQ functional | −.09133 | .01428 | <0.0001 |
| CCQ mental$^2$ | −.01529 | .00684 | 0.0265 | CCQ mental$^2$ | −.02242 | .01007 | 0.0272 |
| CCQ mental$^3$ | .00293 | .00150 | 0.0526 | CCQ mental$^3$ | .00449 | .00221 | 0.0437 |
| Male | .05102 | .02276 | 0.0262 | Male | .07866 | .03352 | 0.0200 |

*Note*. CCQ symptoms$^2$ = CCQ symptoms × CCQ symptoms; CCQ mental$^2$ = CCQ mental × CCQ mental; CCQ mental$^3$ = CCQ mental × CCQ mental × CCQ mental.
CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease.

**Table 3 – Internal validity of the recommended models based on the GO-AHEAD database using different EQ-5D value sets.**

| EQ-5D value set | Dutch model A | | Dutch model B | | US model C | | UK model D | |
|---|---|---|---|---|---|---|---|---|
| | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted |
| Mean EQ-5D | .6625 | .6843 | .6625 | .6901 | .6993 | .7134 | .6002 | .6161 |
| MAE | .1520 | | .1471 | | .1211 | | .1766 | |
| RMSE | .1983 | | .1963 | | .1552 | | .2289 | |
| Mean if observed EQ-5D < 0.25 | .1562 | .5181 | .1562 | .5238 | .2190 | .5184 | .0912 | .4224 |
| Mean if 0.25 ≤ observed EQ-5D < 0.5 | .3513 | .5618 | .3513 | .5567 | .3849 | .5964 | .3101 | .5135 |
| Mean if 0.50 ≤ observed EQ-5D < 0.75 | .6624 | .6530 | .6624 | .6623 | .6364 | .6439 | .6530 | .5939 |
| Mean if 0.75 ≤ observed EQ-5D ≤ 1 | .8566 | .7753 | .8566 | .7821 | .8425 | .7914 | .8841 | .7875 |
| MAE if observed EQ-5D < 0.25 | .3618 | | .3676 | | .2994 | | .3311 | |
| MAE if 0.25 ≤ observed EQ-5D < 0.5 | .2113 | | .2068 | | .2157 | | .2283 | |
| MAE if 0.50 ≤ observed EQ-5D < 0.75 | .1254 | | .1215 | | .0968 | | .1387 | |
| MAE if 0.75 ≤ observed EQ-5D ≤ 1 | .1105 | | .1027 | | .0993 | | .1429 | |

EQ-5D, EuroQol five-dimensional questionnaire; GO-AHEAD, Assessment Of Going Home under Early Assisted Discharge; MAE, mean absolute error; RMSE, root mean square error.

### Correlation between CCQ and EQ-5D

The correlation matrix between the two instruments is shown in Appendix 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.11.006. The correlation between total CCQ scores and EQ-5D values was moderate (−0.514). The correlations were negative because a better score is indicated by a higher value on the EQ-5D and a lower value on the CCQ. Correlations between EQ-5D dimensions and the CCQ total score was mostly lower. Weak correlation (<0.3) was found between the total CCQ score and the EQ-5D dimension pain/discomfort. The highest correlation (0.583) was found between the total CCQ score and the EQ-5D dimension usual activity. The lowest correlations (<0.2) were those between any of the EQ-5D dimensions and CCQ items 5 and 6 related to cough and phlegm production.

The exploratory PCA showed that four constructs had an eigenvalue of more than 1 and explained 69% of the total variance (see Appendix 2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.11.006). CCQ items 1, 2, 7, 8, 9, and 10, which all related to being active and the impact of being active on breathlessness, and the EQ-5D dimensions mobility, self-care, and usual activities all loaded onto the same construct. CCQ items 3 and 4 related to being concerned and depressed loaded on the same construct as the EQ-5D dimension depression/anxiety. CCQ items 5 and 6 related to cough and phlegm production formed a distinct construct, unrelated to any of the EQ-5D items. Likewise, the EQ-5D dimensions pain/discomfort and mobility formed a construct on which none of the CCQ items loaded. Hence, the EQ-5D dimension mobility had a loading of more than 0.4 on two constructs. The two constructs containing both CCQ and EQ-5D items explained a total variance of 48%.

The confirmatory PCA in which the number of components was fixed at five is presented in Appendix 3 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.11.006. Construct 1 of the exploratory PCA was split into two constructs: one construct with CCQ items 1, 2, 7, 8, 9, and 10 and one construct with CCQ items 7, 8, 9, and 10 and EQ-5D dimensions mobility, self-care, and usual activities.

### Models with the Dutch EQ-5D Value Set

The simplest equation for predicting EQ-5D values using the Dutch value set is the 10-parameter model A in Table 2. This model, which was based on the GO-AHEAD data set, had the lowest MAE and RMSE when comparing predicted and observed EQ-5D values from the same population and from different trial populations, especially when observed EQ-5D values were below 0.5 (see Appendix 4 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.11.006). Therefore, this model was chosen as the best model. Model A was able to predict a mean EQ-5D value that was not significantly different from the observed mean EQ-5D value in the validation data set (Table 3). In the external validation, the model was able to predict a mean EQ-5D value that did not significantly differ from the mean observed EQ-5D value in the MARCH data set but the model was unable to predict an accurate mean EQ-5D value in the RECODE data set (Table 4). Model A is a classical OLS regression model because that model outperformed generalized linear models with different distributions and link functions. Furthermore, using a Tobit model instead of an OLS model increased the MAE (0.152 vs. 0.161) and the RMSE (0.198 vs. 0.214).

Scatter plots of predicted and observed values for model A are shown in Figure 1. The scatter plots reveal that the mapping did not produce accurate predictions of the EQ-5D value on an individual level, especially not for the more severe health states (observed EQ-5D values <0.5).

**Table 4 – External validity of the recommended models based on the GO-AHEAD database using different EQ-5D value sets.**

| Data set | EQ-5D value set | Dutch model A | | Dutch model B | | US model C | | UK model D | |
|---|---|---|---|---|---|---|---|---|---|
| | | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted |
| RECODE | Mean EQ-5D | .7553 | .8088* | .7570 | .7907* | .7838 | .8259* | .7085 | .7788* |
| | MAE | .1509 | | .1484 | | .1224 | | .1787 | |
| | RMSE | .2235 | | .2161 | | .1720 | | .2592 | |
| MARCH | Mean EQ-5D | .8830 | .8639 | .8807 | .8389* | .8904 | .8883 | .8703 | .8627 |
| | MAE | .0933 | | .0979 | | .0749 | | .1010 | |
| | RMSE | .1107 | | .1179 | | .0908 | | .1276 | |

EQ-5D, EuroQol five-dimensional questionnaire; GO-AHEAD, Assessment Of Going Home under Early Assisted Discharge; MAE, mean absolute error; MARCH, Health Status Guided COPD Care; RMSE, root mean square error; RECODE, Randomized Clinical Trial on Effectiveness of Integrated COPD Management in Primary Care.
* Significant ($P < 0.01$).

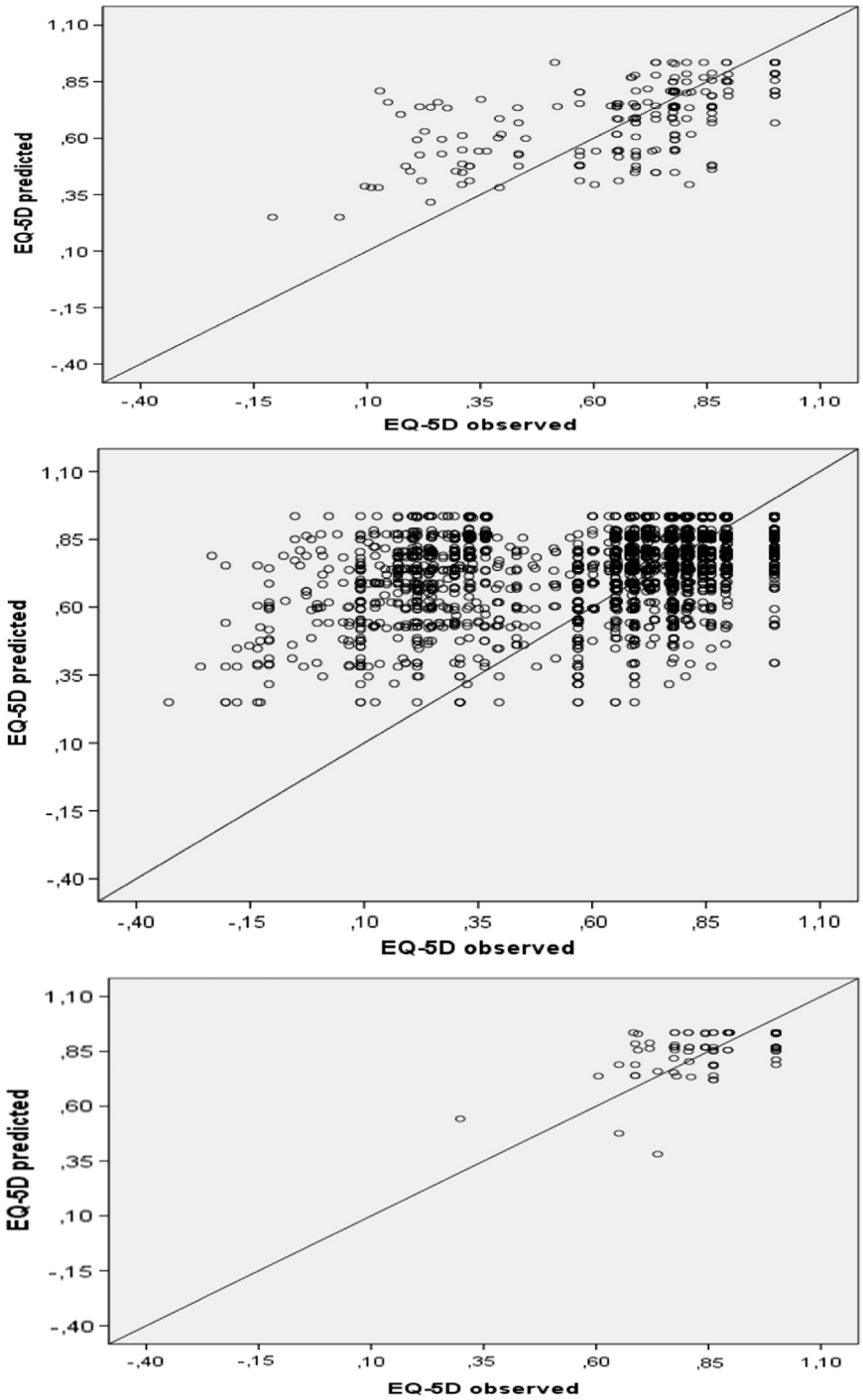


**Fig. 1 – Scatter plots of the observed and predicted EQ-5D in Model A based on the GO-AHEAD development data in (I) GO-AHEAD validation dataset (II) RECODE dataset (III) MARCH dataset. EQ-5D, EuroQol five-dimensionalquestionnaire; GO-AHEAD, Assessment Of Going Home under Early Assisted Discharge; GO-AHEAD, Assessment Of Going Home under Early Assisted Discharge; MARCH; RECODE, Randomized Clinical Trial on Effectiveness of Integrated COPD Management in Primary Care.**

Figure 2 shows the results of the bootstrap procedure based on model A. From these graphs, we can define the range of certainty of the group mean predicted error by group size: with a group size of 10, errors range from +0.3 to −0.2, and when the group size is 500, errors range from +0.1 to 0. The group mean prediction error decreased when the group mean predicted EQ-5D value increased.

We extended model A with the Charlson comorbidity index to build model B, which improved the model and reduced the MAE and RMSE slightly (Table 3). We recommend this model when the Charlson comorbidity index is available. The performance of the other estimated OLS models using the Dutch value set is shown in Appendix 5 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.11.006. We investigated the impact of

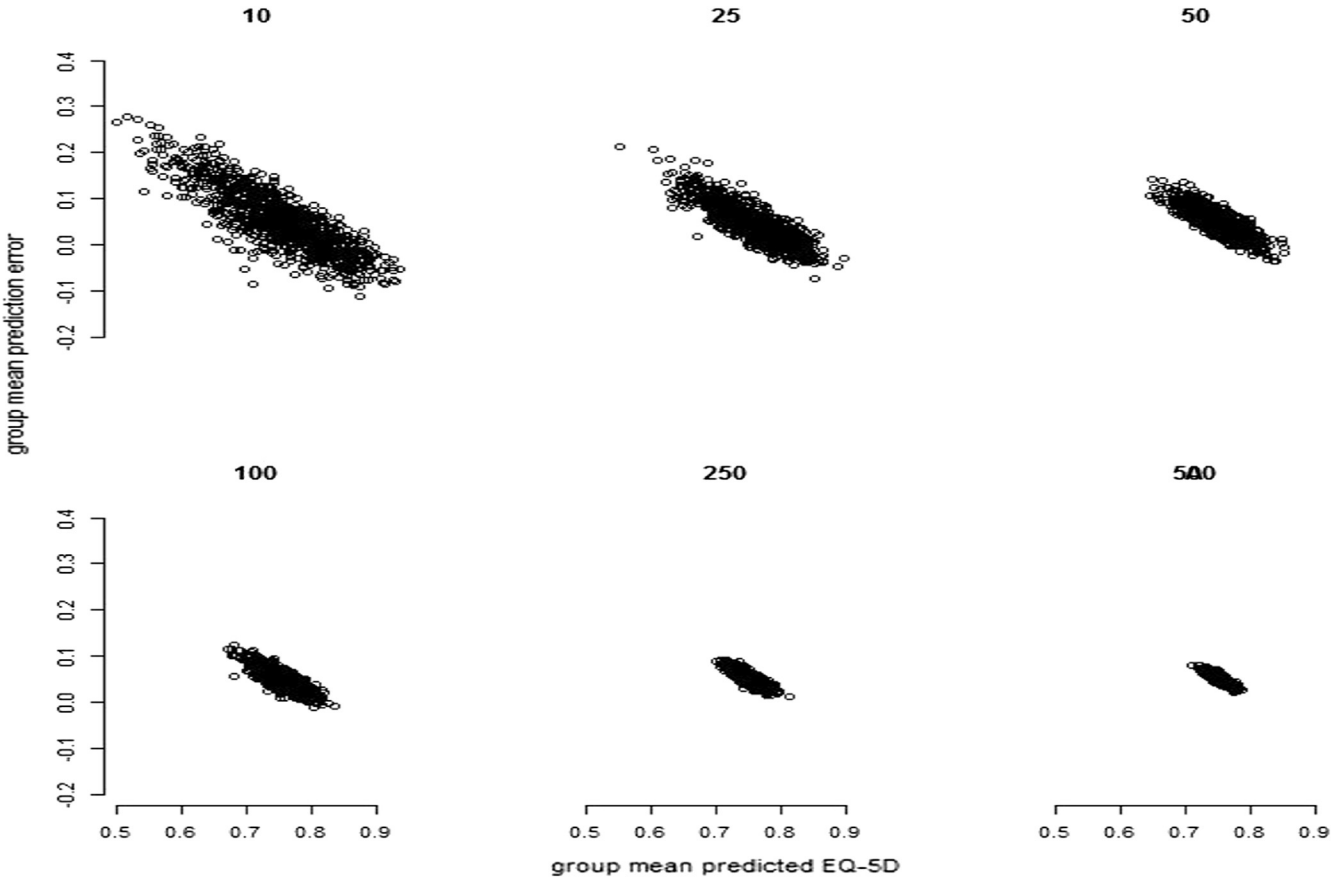


**Fig. 2 – Group mean prediction errors, by group size (10, 25, 50, 100, 250, 500) and group mean predicted EQ-5D value. EQ-5D, EuroQol five-dimensional questionnaire.**

treating repeated measurements from the RECODE data set as independent observations by comparing a model that was based on the entire data set with a model that was based on the baseline measurement only. We found that the model based on the entire data set had better predicted performance (i.e., lower MAE and RMSE) than did the model based on the data set including only the baseline measurements of RECODE. Merging the three different data sets only slightly decreased the MAE from 0.153 (RECODE) to 0.152 (full data set). The RMSE was the lowest (0.198) when the model was based on the GO-AHEAD trial.

### *Sensitivity Analyses with US and UK EQ-5D Value Sets*

When using US and UK value sets, the models based on the GO-AHEAD data set had the lowest RMSE and the lowest MAE for patients with an observed EQ-5D value below 0.5. Although the MAE values of the models based on GO-AHEAD data set were somewhat higher than those for the other models in case of milder impaired health states, only the GO-AHEAD models predicted a mean EQ-5D value that was not significantly different from the observed mean in the MARCH trial. Therefore, the recommended six-parameter model for the US value set (model C) and the recommended six-parameter model for the UK value set (model D) are based on the GO-AHEAD data set. The predicted performances of these models are given in Tables 3 and 4, and the algorithm can be found in Table 2. The predictive ability of the model based on the US value set (MAE 0.121; RMSE 0.155) was slightly better in comparison with those of the model based on the UK value set (MAE 0.177; RMSE 0.229) and the models based on the Dutch value set (MAE 0.147 and 0.152; RMSE 0.196 and 0.198). The scatter plots of models C and D look quite similar to the scatter plot of model A. These plots reveal that we were not able to produce accurate predictions on the individual level. The mapping models underestimated the EQ-5D scores for the mild health states, whereas they overestimated those for the more severe health states. Appendix 6 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2014.11.006 presents the results of the other estimated models for predicting EQ-5D values based on US and UK value sets.

## Discussion

This study aimed to develop a model that predicts EQ-5D values from CCQ scores, which can be used in the absence of directly collected EQ-5D data. The recommended models were estimated from the GO-AHEAD data set because it had the lowest MAE and RMSE, especially for EQ-5D values below 0.5. The main reason is that the GO-AHEAD data set had the widest variation in EQ-5D values and the highest proportion of severe health states. On a group level, the models predicted mean EQ-5D values that were similar to the mean observed EQ-5D values in the same population (internal validity) and the errors as percentage of the EQ-5D range were lower than the typically found percentage of error of the EQ-5D range up to 15% [22]. The predictive ability of the models, however, varied with the severity of HRQOL impairment. The mapping models underestimated the EQ-5D values for the mild health states, whereas they overestimated those for the more severe health states. This “misfit” is a general problem with mapping studies because of regression to the mean [22,23]. The overestimation and underestimation may cancel out when

predicting the overall mean EQ-5D value. The underestimation for patients with good health states, however, is less than the overestimation for patients with more severe health states [22]. This combination of findings may indicate reduced room for improvement in severe health states. In typical Markov models of COPD, patients with COPD are divided into different disease severity states with different EQ-5D values. When using the mapping models to predict EQ-5D values for these states, these predictions are likely to be biased and the overestimation of the low values and the underestimation of the high values would probably no longer cancel each other out.

An important cause of the problematic map of the CCQ scores onto EQ-5D values is the conceptual difference between the two instruments. The correlation between the CCQ and the EQ-5D was moderate (−0.514), and PCA suggests that there are differences in the underlying constructs of the CCQ and the EQ-5D. The EQ-5D dimension pain/discomfort formed a separate construct on which none of the CCQ items loaded. Because this dimension is an important driver of a reduced EQ-5D value in this population, this is an important limitation. Furthermore, CCQ symptom items 5 and 6 related to cough and phlegm production formed a distinct construct, unrelated to any of the EQ-5D items.

If we compare our mapping model with the mapping model of the SGRQ [7], another frequently used disease-specific HRQOL instrument in COPD, the overall predictive ability is comparable. The MAE and RMSE of the model that predicts EQ-5D values (US value set) from SGRQ scores were 0.124 and 0.172, respectively [7], and the MAE and RMSE of our model that predicts EQ-5D values (US value set) from CCQ scores were 0.121 and 0.155, respectively. Unfortunately, we cannot compare the misfit (i.e., underestimation and overestimation of the EQ-5D values) of the model with the SGRQ because a scatter plot of predicted and observed EQ-5D values is not presented in the SGRQ mapping study.

This is the first mapping study that developed models for EQ-5D value sets of different countries. Because the model coefficients differ between countries, future mapping studies are recommended to account for different value sets of the EQ-5D. Apart from the fact that we investigated a wide range of different models with increasing complexity for different EQ-5D value sets, the strength of our study lies in the large number of observations compared with previous mapping studies [6]. Using repeated EQ-5D measurements of the same patients as unique observations, as previous mapping studies also did, increased the predictive ability of the model. The bootstrap procedure showed to which extent the accuracy of the mapping increases when they are estimated in a larger sample size. These results are consistent with the study from Grootendorst et al. [24]. The fact that the model estimated from the GO-AHEAD trial performed best demonstrates, however, that not only the size of the data set matters but also the variation in EQ-5D values as we mentioned earlier. A wide range of disease severity is a premise for undertaking mapping [22].

A limitation of this study is that the numbers of patients with an EQ-5D value of less than 0.5 was relatively small and 25% of the observations were at full health. The final models were based on OLS regression because generalized linear models and Tobit models did not improve the prediction. Some studies suggest that the prediction performance of mapping models may be improved by using more complex models such as the censored least absolute deviation model [22]. Their use in previous mapping studies, however, was limited (80% of the mapping studies used OLS models [6]) and showed conflicting results on improvement in predictive ability [22,25,26]. Recently, latent class analysis showed some promising improvements in the predictive ability of mapping models [27]. For future studies, it would be worthwhile to investigate whether these new modeling approaches will actually lead to better predictability in patients with COPD.

## Conclusions

On a group level, the CCQ mapping models can predict mean EQ-5D values that are similar to observed mean values if the patient population is similar to the population in which the algorithm was developed. The overestimating of the low observed EQ-5D values and the underestimation of the high observed EQ-5D values, however, limits its use in cost-effectiveness Markov models in which utility estimates by disease severity state are required. Therefore, mapping algorithms should be used cautiously.

Source of financial support: Boehringer Ingelheim supported this study.

## Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2014.11.006 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

## REFERENCES

[1] Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation: the St. George's Respiratory Questionnaire. Am Rev Respir Dis 1992;145:1321–7.
[2] van der Molen T, Willemse BW, Schokker S, et al. Development, validity and responsiveness of the clinical COPD questionnaire. Health Qual Life Outcomes 2003;1:13.
[3] EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16:199–208.
[4] National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal. London: National Institute for Health and Care Excellence, 2008.
[5] Coast J. Reprocessing data to form QALYs. BMJ 1992;305:87–90.
[6] Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. Health Qual Life Outcomes 2013;11:151.
[7] Starkie HJ, Briggs AH, Chambers MG, Jones P. Predicting EQ-5D values using the SGRQ. Value Health 2011;14:354–60.
[8] Ringbaek T, Martinez G, Lange P. A comparison of the assessment of quality of life with CAT, CCQ, and SGRQ in COPD patients participating in pulmonary rehabilitation. COPD 2012;9:12–5.
[9] Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. Available from: http://www.goldcopd.org. [Accessed December 1, 2013].
[10] Lamers LM, McDonnell J, Stalmeier PF, et al. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ 2006;15:1121–32.
[11] Kruis AL, Boland MR, Schoonvelde CH, et al. RECODE: design and baseline results of a cluster randomized trial on cost-effectiveness of integrated COPD management in primary care. BMC Pulm Med 2013;13:17.
[12] Utens CM, Goossens LM, Smeenk FW, et al. Effectiveness and cost-effectiveness of early assisted discharge for chronic obstructive pulmonary disease exacerbations: the design of a randomised controlled trial. BMC Public Health 2010;10:618.
[13] Kocks J, de Jong C, Berger MY, et al. Putting health status guided COPD management to the test: protocol of the MARCH study. BMC Pulm Med 2013;13:41.
[14] Oppe M, Devlin N, Black N. Comparison of the underlying constructs of the EQ-5D and Oxford Hip Score: implications for mapping. Value Health 2011;14:884–91.
[15] Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 1960;20:141–51.
[16] Hatcher L. Principal component analysis. In: Hatcher L, ed. A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. Cary, NC: SAS Institute, Inc., 2003.
[17] Charlson ME, Pompei P, Ales KL, Mackenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
[18] Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. Qual Life Res 2000;9:901–10.

[19] Guisan A, Edward TC Jr, Hastie T. Generalized linear and generalized additive models in studies of species distributions: setting the scene 2002;157:89–100.
[20] Janssen B, Szende A. Population norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, eds. Self-Reported Population Health: An International Perspective based on EQ-5D. London: Springer, 2014.
[21] Charlson ME, Charlson RE, Peterson JC, et al. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. J Clin Epidemiol 2008;61:1234–40.
[22] Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ 2010;11: 215–25.
[23] Fayers PM, Hays RD. Should linking replace regression when mapping from profile-based measures to preference-based measures? Value Health 2014;17:261–5.
[24] Grootendorst P, Marshall D, Pericak D, et al. A model to estimate health utilities index mark 3 utility scores from WOMAC index scores in patients with osteoarthritis of the knee. J Rheumatol 2007;34:534–42.
[25] Austin PC. A comparison of methods for analyzing health-related quality-of-life measures. Value Health 2002;5:329–37.
[26] Pullenayegum EM, Tarride JE, Xie F, et al. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? Value Health 2010;13:487–94.
[27] Hernandez Alava M, Wailoo AJ, Ara R. Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. Value Health 2012;15:550–61.