Mapping the Modified Rankin Scale (mRS) measurement into the generic EuroQol EQ-5D health outcome

Melissa Ouellet, MSc[1]; Oliver Rivero-Arias, MSc[1]; Alastair Gray, PhD[1]; Jane Wolstenholme, PhD[1]; Peter M. Rothwell, MD, PhD, FRCP[2]; Ramon Luengo-Fernandez, MA, MSc[1]

[1]Health Economics Research Centre, Department of Public Health, University of Oxford, UK
[2]Stroke Prevention Research Unit, University Department of Clinical Neurology, Radcliffe Infirmary, Oxford, UK

Correspondence to Melissa Ouellet:

Health Economics Research Centre; Department of Public Health; University of Oxford; Old Road Campus; Oxford; OX3 7LF; Tel: +44 (0) 1865289266; Fax: +44 (0) 1865289271; Email: melissa.ouellet@dphpc.ox.ac.uk

Abstract

Background: Mapping disease-specific instruments or generic health outcomes into utility values is an expanding field of interest in health economics, and the existence of reliable mapping algorithms may be valuable in circumstances where information only from a disease-specific instrument has been collected. In the area of stroke, for example, the modified Rankin Scale (mRS) has been widely collected alongside clinical trials to evaluate treatment effectiveness, and translating mRS results into, for example, EQ-5D utility tariffs, may help in calculating quality adjusted life years gained, and be of potential value to decision-makers as well as professionals in the field.

Methods: The Oxford Vascular study (OXVASC), a large observational longitudinal study which collected mRS and EQ-5D measurements from stroke or transient ischaemic attack (TIA) patients followed over a 24-month period, was the main source of data on which to perform the mapping exercises in this paper. The analysis was performed on a sample of 1283 patients clustered at four different follow-up points. Ordinary least squares (OLS) regression was used to predict UK EQ-5D tariffs from mRS scores. An alternative method, using multinomial logistic regression to predict responses to each EQ-5D question, was also explored. The effects of including age and gender in the regressions were also examined. The validity of the models was evaluated according to the magnitude of their predicted-to-actual mean EQ-5D tariff difference, of their mean squared and mean absolute errors, and by the use of quantile-quantile (q-q) plots.

Results: A total of 2425 observations of the clustered data provided matched pairs of mRS and EQ-5D measurements which were used in the analysis. The OLS model with age and gender predicted tariffs closest to the mean with smallest MSE (0.0448) and MAE (0.1533) values. The mutinomial logistic model including the same baseline characteristics produced larger MSE (0.0736) and MAE (0.1925) but predicted better individual EQ-5D tariffs for a particular patient. These results were corroborated by the distribution of differences between actual and predicted tariffs, and by examining q-q plots of the individual models.

Conclusions: Two algorithms are introduced to predict from modified Rankin Scale data either average EQ-5D tariff values for a particular group with the OLS approach or individual EQ-5D response levels from the multinomial method. The choice of algorithm by its user(s) will be dependent upon their study aim. Individuals outside the UK may find it more useful to use the multinomial results, which can be used with different country-specific tariff valuations. Finally, these algorithms should not be viewed as alternatives to prospective collection of utility data.

Introduction

Stroke is second only to coronary heart disease as the most common cause of death and the leading cause of disability in adults in the UK (Rothwell et al. 2004). Researchers worldwide have explored the effectiveness of stroke-related treatments in clinical and epidemiological studies over the last two decades (Antithrombotic Trialists Collaboration 2002; CAPRIE Steering Committee 1996; SPARCL 2006; EAFT Study Group 1993; Rothwell et al. 2004). The development and availability of disease-specific instruments to measure clinical effectiveness in this area as a result has increased considerably (Pickard et al. 2005).

At the same time, health economists have enthusiastically advocated the use of a standard metric to measure the health benefits of healthcare interventions, in order to increase the comparability of cost-effectiveness estimates across disease areas. Policy makers in the UK nowadays thus increasingly use quality-adjusted life years (QALYs) as the main measure of health benefits when performing economic evaluations of a given medical intervention (Rawlins and Culyer 2004). Although the inclusion of instruments to derive QALYs is becoming more common in new research protocols, a large number of studies have not included or are not including the necessary information to calculate QALYs, and this is likely to be the case in at least some future epidemiological protocols.

The EuroQol EQ-5D is a self-administered questionnaire and currently one of the most commonly used instruments for obtaining utility values to derive QALYs for a particular intervention, for example, by attaching a population-derived valuation or tariff to each health state described by respondents. Although used alongside randomised and observational clinical trials, further evidence of EQ-5D utility values in stroke patients may be useful. From the clinical side, the modified Rankin Scale (mRS) is a well-known disease specific instrument which measure levels of disability post incident or treatment in patients after brain injuries or stroke (Haacke et al. 2006). It has become one of the standard outcomes in evaluating the effectiveness of stroke-related treatments in clinical trials (Wilson et al. 2002). Therefore, translating disease-specific stroke instruments such as the mRS into outcomes which would enable the calculation of QALYs is of potential interest to researchers and decision-makers in this field.

The use of regression modelling, or mapping, to translate health-related outcomes measurements into utility values is not a novel approach. The EQ-5D outcome has received particular attention and there are already available algorithms to obtain its utilities from standard generic health outcome instruments, such as the SF-36 and SF-12 (Franks et al. 2003; Franks et al. 2004; Gray et al. 2006; Sullivan and Ghushchyan 2006).

In the case of mapping the mRS to obtain EQ-5D data, the literature is very limited or nonexistent. Haacke et al. (2006) evaluated long term health-related quality of life (HRQoL) of stroke survivors. A univariate analysis was carried out to identify EQ-5D scores for two different groups of the mRS and a prognostic model was also constructed to identify independent predictors of the EQ-5D index. The mRS health outcome as a binary variable was found not to be a statistically significant predictor in their model when included alongside other variables such as the Barthel Index (which measures ability to perform activities of daily living). In another study, Pickard et al. (2005) investigated the responsiveness of generic HRQoL measures in stroke and used the mRS as an anchor-based measurement to determine whether a change had really occurred. They found strong correlation between changes in EQ-5D tariff and changes in the mRS scores. Therefore to our knowledge, no exercise has attempted to translate the modified Rankin Scale scores into EQ-5D response levels or "tariff" values.

This study aims to estimate the association between the mRS scores and the EQ-5D UK tariff values or individual question responses using data from a large observational study. The association between the EQ-5D tariff and the mRS scores will be explored using Ordinary Least Squares regression whereas the dimension responses from the EQ-5D questionnaires will be investigated using a multinomial logistic regression. A second study with information on mRS and EQ-5D will be used to validate the results from these models.

Methods

Data

The Oxford Vascular Study (OXVASC), a large longitudinal observational study, was used in this methodological research paper. OXVASC monitored the occurrence of stroke and transient ischaemic attacks (TIA) for individuals registered in nine general practices located in Oxfordshire, United Kingdom, who were recruited into the study between April 2002 and March 2007. The sample used for this analysis consisted of 1283 stroke and TIA patients followed up at 1, 6, 12 and 24 months. As part of their follow-up, which was conducted face-to-face by a research nurse, patients were assessed using the mRS and asked to complete the EQ-5D. Unfortunately, during the first 18 months of OXVASC, the EQ-5D was not given at the 6-month follow-up. Baseline demographics age and gender were also extracted and included in the final dataset. In addition, some clinical information, such as whether a stroke or TIA had occurred, was obtained in order to conduct validity checks in sub-groups.

Instruments

The modified Rankin Scale is one of the most commonly used health outcome tools for clinical evaluation of patients in the area of stroke (Haacke et al. 2006). The score is derived from an interview with a patient performed either by a nurse or a clinician, in a clinic or over the phone (Eriksson et al. 2007). The mRS assesses the level of independence of a patient post-stroke on both mental and physical levels (Sulter et al. 1999). It evaluates a patient's level of handicap by looking at activity limitations and lifestyle changes (Wolfe et al. 1991). The scale spans 6 grades from 0 to 5, where 0 stands for the patient exhibiting no symptoms at all, and 5, for severe disability. Individuals are commonly classified as "independent" if their score ranges from 0 to 2 inclusive, or as "severely disabled" if from 3 to 5 (Haacke et al. 2006) (in many studies 6 is used for "dead"). A detailed description of the mRS scale levels is given in Table 1.

The generic EuroQol EQ-5D is a multiattribute questionnaire for measuring preferences associated with a particular patient's health state. A profile with five dimensions (mobility, self-care, usual activities, pain, anxiety) each with three levels (no problem, some problems, extreme problems) can be described. EQ-5D health states are defined by five numbers as each of the levels are coded from 1 to 3, with 11111 being perfect health, and 33333, the worst possible health state. Therefore, the EQ-5D classification defines 243 possible health states (3^5) which can then be transformed into utility values, for example by using available "tariffs". In this study, the UK tariff derived from a large British sample population using time trade-off methods and regression analysis was used (Dolan et al. 1996). This tariff ranges from -0.594 to 1, where 1 is defined as "full health", 0 represents "dead" and negative values health states worse than dead.

Statistical Analysis

Descriptive statistics for both instruments were produced to understand the relation between them. A univariate analysis of EQ-5D tariffs by mRS level and comparison across the groups was done using standard parametric testing. To make use of the potential of having a longitudinal study available, cluster regression analysis was considered, where each respondent acted as a possible cluster, hence adjusting the standard errors. The interclass correlation coefficient (ICC) supported the evidence for this analysis. The ICC values are approximately equal to 0.77 for the mRS scores and 0.59 for EQ-tariffs at each follow-up point. Values over 0.60 are commonly considered good indicators of the inter-dependence of measurements.

Mapping the mRS to the EQ-5D tariff

Ordinary least squares (OLS) regression was fitted using the mRS, as a discrete ordinal independent dummy variable x_{it} for respondent *i* at follow-up time *t*, and the EQ-5D tariff, as a continuous dependent variable y_{it}, following the standard approach:

$$y_{it} = \beta x_{it} + \varepsilon_i$$

where β represents the coefficients associated with the dummies mRS including a constant term, and ε_i, the associated error term.

Mapping the mRS to the EQ-5D domains

As a second approach, a multinomial logistic (Mlogit) regression was explored to study a response mapping approach to the EQ-5D tariff. In this method, each of the EQ-5D domains which can take any of the three possible values is regressed to individual mRS scores using a multinomial logistic regression. Although, the response categories from the EQ-5D follow a natural ordering, an ordered logistic approach was not considered in case the proportional odds or parallel regression assumption was not valid.

The Mlogit predicts the probability of being in a given level *m*, and then assigns the level with the highest probability to each individual. A reference category needs to be selected to be able to calculate the probabilities for each level. For this analysis, the response 1 "no problem" was used as the baseline category, so that the general model format for a predicted probability of a given response level is as follows (Green 1997; Long 1997):

$$\Pr(y_{it} = m | x_{it}) = \frac{\exp(x_{it}\beta_m)}{1 + \sum_{j=2}^{J} \exp(x_{it}\beta_j)}$$

where y_{it} stands for the EQ-5D the response level, *j* = 1,2,..., *m*-1 for a given set of explanatory variables, and x_{it}, for the stroke-specific mRS dummy scores

Simply choosing the response level with the highest predicted probability of occurring will not give the expected value of the response. To correct for this, Monte Carlo simulation was used when predicting response level categories, using random numbers compared against the predicted probabilities.

Age and gender were included to account for any potential significant influence of these characteristics on the overall EQ-5D tariffs or response values in both models.

Internal validation

The validity of the individual regression models was assessed on the basis of the magnitude of the difference between the predicted and actual mean EQ-5D tariffs, using the mean squared error (MSE) and mean absolute error (MAE). The MSE is the square of the difference between the actual and predicted EQ-5D tariff, and the MAE is the absolute value of this same difference in utility values. The models were evaluated from observing the distribution of the difference between the actual and predicted tariff values. Quantile-Quantile plots were plotted to visualize how close the predicted and actual distributions were to each other for the multinomial logistic model, whereas quantile-quantile plots of the residuals of the OLS regression were plotted against ones of a normal distribution to check for the normality of errors in simple linear models. All of the analyses were carried out using STATA version 10 (StataCorp. 2007).

Results

Outcome descriptives

Baseline characteristics of the respondents of the OXVASC study are reported in Table 2. The age distribution of this study population indicates relatively old patients in this sample concentrated around the 75-85 age interval. The ratio of females to male is approximately 1:1 with a female proportion of 53.2% compared to 46.8% in the case of males.

Table 3 summarizes the distribution of deaths and of missing EQ-5D tariff and mRS-data across the different follow-up points at 1, 6, 12, and 24 months. Due to the longitudinal nature of the study and the ongoing collection of follow-up data, the proportion of missing values mainly increases over time.

The frequency distribution of the EQ-5D tariffs and of the mRS scores are reported in detail for the different follow-up points in Tables 4 and 5 respectively. The EQ-5D tariffs values seem to remain relatively stable over the different follow-up points. The highest proportion of respondents is consistently found in the mRS score 1 category over time. This is also reflected in the higher frequency of respondents classified as in the "independent" category of the mRS scoring system. Table 6 summarizes the distribution of the EQ-5D response levels by domain and follow-up. Most of the data points are concentrated towards the higher level category, in this case response level 1, consistently over time for each domain.

Table 7 shows the mean EQ-5D tariff corresponding to a given mRS score. The results have face validity in that the tariff values are inversely proportional to the level of mRS scores across different follow up points. The corresponding mean tariff values for a specific mRS score are also quite close to each other across different follow-up points.

The tabulation of corresponding mean EQ-5D tariffs for changes in mRS scores can be found in Table 8. Parametric testing was performed and indicated that there is no evidence to indicate that there is a significant difference in EQ-5D tariff change between either going up or down 1 or 2 units in the mRS. A patient increasing 1 level on the mRS scale would be expected to decrease 0.092 in tariff value, and decreasing 1 mRS score would result in a 0.130 increase in tariff. Going up or down 2 units on the mRS scale would yield a 0.275 decrease and 0.257 increase in utility values respectively. Due to a limited number of observations, such inferences could not be drawn for changes of 3, 4 or 5 units in mRS scores.

Regression results

Table 9 shows the results of the internal validity test for the different regression modelling approaches. It suggests that the OLS approach performed better than the Mlogit when predicting the actual mean EQ-5D tariff of 0.7135. Including the explanatory variables age and gender in both methods improved overall mean tariff prediction for the OLS and reduced the errors for both. The OLS with age and gender yielded the lowest MSE and MAE values of 0.0448 and 0.1533 respectively. The other OLS model containing only mRS scores predicted a similar tariff value of 0.7134, but with larger MSE and MAE errors of 0.0453 and of 0.1453. The Mlogit with age and gender yielded a mean tariff estimate of 0.7044 with higher MSE and MAE values of 0.0727 and 0.1900, but lower errors than its logistic counterpart. However, the Mlogit approach did produce a range of values closer to the observed data than the OLS range, as will be inferred when comparing frequency distributions.

The superiority of the age-gender OLS model is not as clear cut when visualizing the comparison between the regressions by plotting the frequency distribution of the difference between the actual and predicted EQ-5D tariffs for the two OLS and the two Mlogit models (Figure 1). The Mlogit models predict a higher proportion of quasi-exact tariff values (close to 0.1%) than the OLS models, yielding at its best (with the age-gender Mlogit) 14.9% exact values compared to 0.7% and 0% in the age-gender and mRS-only OLS models. The Mlogit models furthermore predict a higher proportion within a 5% range of actual tariff values: 22.2% from the age-gender Mlogit compared to 19.0% from the age-gender OLS. Visualizing the regression results in this way furthermore also illustrates clearly the gain resulting from including age and gender as variables, with a considerable increase in the proportion of

predicted values lying within a 5% range of the actual ones. In the case of the OLS model, this proportion grew from 17.1% to 19.0% and in the Mlogit case, from 21.7% to 22.2%.
In addition, Figure 2 shows that the Mlogit estimates accurately individual predictions or mostly with an error of one level up or down for all domains.

The values of the quantiles of the residuals of the linear model versus a normal distribution are reported in Figure 3. In both OLS models, it would seem that the distribution of the residuals can be said to be relatively normal, even though there is indication of some negative skewness, and consequently that the models are in line with the assumption of normally distributed residuals for linear models. When plotting the residuals of the actual versus predicted quantiles for the logistic models as seen in Figure 3, one is able to visualize the relatively small dispersion of the predictions around the sample observations.
Hence, two algorithms (with 2 subsets each, depending on whether one holds baseline characteristics on the sample) can be used to translate individual mRS measurements into specific EQ-5D tariffs through the use of the coefficients reported in Table 10.

Discussion

It is difficult to nominate a 'best-fit' model for predicting EQ-5D response values or tariffs from given mRS scores. The OLS approach predicts nearest to mean tariffs, and contains a lower model error compared to the Mlogit. However, when looking at the frequency distribution of the differences in observed and predicted values, the Mlogit regressions consistently predict a higher proportion of quasi-exact values, and also perform better in the proportion of estimated values within 5% of the actual value.

Given the small number of explanatory variables, the OLS models examined here are essentially an algebraic reiteration of the descriptive data presented in Table 8 showing the mean corresponding EQ-5D tariffs for given mRS scores. The model looks at the incremental change in tariff from one score to the other, and bases the coefficient parameters solely on this. The addition of age and gender improve the predictions, but not markedly. In addition, the OLS models predict the tariffs as continuous variables, so that exact utility values will rarely if ever be predicted.

The Mlogit regression models have different strengths and limitations. One of the inherent weaknesses of using the Mlogit approach is largely related to the format of the EQ-5D questionnaire, which only allows for three levels within a given domain. A wrong prediction inevitably results in a significant error in the resulting valuation. However, even though the Mlogit models seemed to have performed worse than the OLS in terms of mean prediction, the predicted seemed to perform better in terms of the proportion of individual predicted tariffs close to and on the actual values. One reason why the Mlogit models exhibit these results is that the tariffs are not treated as continuous variables as in the OLS, but derived from a specific set of predicted responses. If the responses are predicted correctly, the resulting tariff will be an exact match. The Mlogit is also advantageous in that the predicted response values can easily be used in different countries with country-specific tariffs derived from these response levels.

In this study, only three explanatory variables (i.e. mRS levels, age, and gender) were examined. Requiring additional explanatory variables may restrict the usefulness of an algorithm to data sets containing all such variables. There might also be some potential limitations in attempting to explain the variability of the EQ-5D instrument using the mRS as the only instrument. All predictions were made in-sample, as there were no data available at the time for external validation. It is therefore possible that out-of-sample predictions might indicate weakness in either of the models.

Finally, the instruments used in this study are quite precise, and any deviation from the exact wording used in the mRS and EQ-5D instruments could render the algorithms examined here invalid.

Conclusion

This study evaluated different approaches to predict EQ-5D tariffs or response values from the Modified Rankin Scale using regression modelling on a large longitudinal database. Preliminary descriptives show that mRS scores are inversely proportional to mean EQ-5D tariffs, indicating a consistent trend across different follow-up points, providing the basis on which to examine different models. Simple univariate (mRS) OLS, multivariate (mRS, age, and gender) OLS, and Mlogit regressions were run to predict mean EQ-5D tariffs. The OLS approach yielded the closest prediction to the actual mean EQ-5D tariffs, and the age-gender OLS model performed best in estimating actual mean values and with lowest mean square and mean absolute model errors. When analysing the frequency distribution of these predictions however, i.e. the predicted-to-actual differences, the multinomial models showed a higher degree of consistency in the proportion of individual estimated values close to the actual ones. Both these approaches have advantages, and the choice of algorithm by its user(s) will be dependent upon their study aim. Individuals outside the UK may find it more useful to use the multinomial results, which can be used with different country-specific tariff valuations. Finally, these algorithms should not be viewed as alternatives to prospective collection of utility data.

References

Antithrombotic Trialists' Collaboration (2002). "Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients." BMJ **324**: 71–86.

Brazier, J., Roberts, J. and Deverill, M. (2002). "The estimation of a preference-based measure of health from the SF-36." J Health Econ **21**(2): 271-92.

CAPRIE Steering Committee (1996). "A randomised, blinded trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE)." Lancet **348**: 1329-39.

Dolan, P., Gudex, C., Kind, P. and Williams, A. (1996). "The time trade-off method: results from a general population study." Health Econ. **5**(2): 141–54.

EAFT (European Atrial Fibrillation Trial) (1993). "Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke." Lancet **342**: 1255-62.

Eriksson, M., Appelros, P., Norrving, B., Terént, A. and Stegmayr, B. (2007). "Assessment of Functional Outcome in a National Quality Register for Acute Stroke: Can Simple Self-Reported Items Be Transformed Into the Modified Rankin Scale?" Stroke **38**(1384-1386).

Franks, P., Lubetkin, E. I., Gold, M. R. and Tancredi, D. J. (2003). "Mapping the SF-12 to preference-based instruments: convergent validity in a low-income, minority population." Med Care **41**(11): 1277-83.

Franks, P., Lubetkin, E. I., Gold, M. R., Tancredi, D. J. and Jia, H. (2004). "Mapping the SF-12 to the EuroQol EQ-5D Index in a National US Sample." Medical Decision Making **24**(3): 247-254.

Glick, H. A., Polsky, D., Willke, R. J. and Schulman, K. A. (1999). "A comparison of preference assessment instruments used in a clinical trial: responses to the visual analog scale from the EuroQol EQ-5D and the Health Utilities Index." Med. Decis. Mak. **19**(3): 265-75.

Gnanalingham, K. K., Apostolopoulos, V., Barazi, S. and O'Neill, K. (2006). "The impact of the international subarachnoid aneurysm trial (ISAT) on the management of aneurysmal subarachnoid haemorrhage in a neurosurgical unit in the UK." Clinical Neurology and Neurosurgery **108**: 117-23.

Gray, A., Rivero-Arias, O. and Clarke, P. (2006). "Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping." Medical Decision Making **26**(1): 18-29.

Greene, W. H. (1997). Econometric Analysis. London, Prentice Hall.

Haacke, C., Althaus, A., Spottke, A., Siebert, U., Back, T. and Dodel, R. (2006). "Long-Term Outcome After Stroke: Evaluating Health-Related Quality of Life Using Utility Measurements." Stroke **37**: 193-198.

Holtslag, H. R., Van Beeck, E. F., Lindeman, E. and Leenen, L. P. (2007). "Determinants of long-term functional consequences after major trauma." J Trauma **62**(4): 919-27.

International Subarachnoid Aneurysm Trial (ISAT) Collaborative Group (2002). "International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial." Lancet **360**: 1267-74.

Kind, P., Dolan, P., Gudex, C. and Williams, A. (1998). "Variations in population health status; results from a United Kingdom national questionnaire survey." BMJ **316**: 736-41.

Lanzino, G., Fraser, K., Kanaan, Y. and Wagenbach, A. (2006). "Treatment of ruptured intracranial aneurysms since International Subarachnoid Aneurysm Trial: practice utilizing clip ligation and coil embolization as individual or complementary therapies." J Neurosurg **104**: 344-349.

Laukontaus, S. J., Pettilä, V., Kantonen, I., Salo, J. A., Ohinmaa, A. and Lepäntalo, M. (2006). "Utility of surgery for ruptured abdominal aortic aneurysm." Ann Vasc Surg **20**(1): 42-8.

Long, J. S. (1997). Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks (CA), Sage.

Molyneux, A. J., Kerr, R., Yu, L.-M., Clarke, M., Sneade, M. and Yarnold, J. (2005). "International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised comparison of effects on survival, dependency, seizures, rebleeding, subgroups, and aneurysm occlusion." Lancet **366**: 809-817.

Oliver, A. J. (2003). "Putting the quality into quality-adjusted life years." J Public Health Med **25**(1): 8-12.

Pickard, A. S., Johnson, J. A. and Feeny, D. H. (2005). "Responsiveness of generic health-related quality of life measures in stroke." Quality of Life Research **14**(1): 207-219.

Rawlins, M. D. and Culyer, A. J. (2004). "National Institute for Clinical Excellence and its value judgments." BMJ **329**: 224-27.

Rothwell, P. M., Coull, A. J., Giles, M. F., Howard, S. C., Silver, L. E., Bull, L. M., et al. (2004). "Change in stroke incidence, mortality, case-fatality, severity, and risk factors in Oxfordshire, UK from 1981 to 2004 (Oxford Vascular Study)." Lancet **363**: 1925-33.

Rothwell, P. M., Eliasziw, M., Gutnikov, S. A., Warlow, C. P. and Barnett, H. J. M. (2004). "Endarterectomy for symptomatic carotid stenosis in relation to clinical subgroups and timing of surgery." Lancet **363**: 915-24.

Suarez, J., Tarr, R. and Selman, W. (2006). "Aneurysmal Subarachnoid Hemorrhage." N Engl J Med **354**: 387-96.

Sullivan, P. W. and Ghushchyan, V. (2006). "Mapping the EQ-5D Index from the SF-12: US General Population Preferences in a Nationally Representative Sample " Medical Decision Making **26**(4): 401-9

Sulter, G., Steen, C. and De Keyser, J. (1999). "Use of the Barthel index and modified Rankin scale in acute stroke trials." Stroke **30**(8): 1538-41.

The EuroQol Group (1990). "EuroQol - a new facility for the measurment of health-related quality of life." Health Policy **16**: 199-208.

The Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) Investigators (2006). "High-Dose Atorvastatin after Stroke or Transient Ischemic Attack." N Engl J Med **355**: 549-59.

Townend, W. (2001). "Relation between Glasgow outcome score extended (GOSE) and the EQ-5D health status questionnaire after head injury." Journal of Neurology, Neurosurgery & Psychiatry **70**(2): 267-8.

Van den Berg, R., Rinkel, G. J. E. and Vandertop, W. P. (2003). "Treatment of ruptured intracranial aneurysms: implications of the ISAT on clipping versus coiling." European Journal of Radiology **46**: 172-177.

Van Gijn, J., Kerr, R. and Rinkel, G. (2007). "Subarachnoid haemorrhage." Lancet **369**: 306-318.

Van Swieten, J. C., Koudstaal, P. J., Visser, M. C., Schouten, H. J. A. and Van Gijn, J. (1988). "Interobserver Agreement for the Assessment of Handicap in Stroke Patients." Stroke **19**: 604-07.

Williams, L. S., Weinberger, M., Harris, L. E., Clark, D. O. and Biller, J. (1999). "Development of a Stroke-Specific Quality of Life Scale." Stroke **30**: 1362-1369.

Wilson, J. T. L., Hareendran, A., Grant, M., Baird, T., Schulz, U. G. R., Muir, K. W., et al. (2002). "Improving the assessment of outcomes in stroke: Use of a structured interview to assign grades on the modified Rankin scale." Stroke **33**: 2243-46.

Wolfe, C. D., Taub, N. A., Woodrow, E. J. and Burney, P. G. (1991). "Assessment of scales of disability and handicap for stroke patients." Stroke **22**(10): 1242-4.

Tables and Figures

Table 1. Generic mRS scale used in the OXVASC study*

Modified Rankin Score	
0	No symptoms at all.
1	No significant disability despite symptoms; able to perform all usual duties and activities.
2	Slight disability; unable to perform all normal activities but able to look after own affairs without assistance.
3	Moderate disability requiring some help but able to walk without assistance.
4	Moderately severe disability; unable to walk without assistance, and unable to attend to own bodily needs without assistance.
5	Severe disability; bedridden, incontinent and requiring constant nursing care and attention.

*(Sulter et al. 1999)

Table 2 . Baseline characteristics recorded for the OXVASC study

		Proportion (%)	Mean (Std. Dev.)
Age	<45	3.0	39.0 (6.21)
	45-55	5.4	51.3 (2.61)
	>55-65	12.9	60.5 (2.73)
	>65-75	23.3	70.3 (2.90)
	>75-85	36.9	80.3 (2.83)
	>85	18.6	89.2 (3.12)
	Total	100.0	74.4 (12.58)
Gender*	Male	46.8	72.3 (12.16)
	Female	53.2	76.1 (12.68)

* The "Mean" value for the "Gender" group represents the mean age of the particular group being described

Table 3. Deaths and missing EQ-5D and mRS data from the OXVASC study

	Baseline*	1 month	6 months	12 months	24 months
Alive	1283	1127	1051	1011	964
Deaths		156	76	40	47
Cum. Deaths		156	232	272	319
EQ-5D tariff					
N	0	764	579	667	428
Missing		363	472	344	536
% missing		32.2	44.9	34.0	55.6
mRS					
N	0	853	821	703	444
Missing		274	230	308	520
% missing		24.3	21.9	30.5	53.9

*Note: Clinical outcomes in the form of mRS and EQ-5D data were not recorded at baseline

Table 4. Mean EQ-5D utility values

Time	N	Mean (Std. Dev.)
1 month	764	0.700 (0.308)
6 months	579	0.721 (0.289)
12 months	667	0.730 (0.266)
24 months	428	0.702 (0.280)
Total	2438	0.713 (0.288)

Table 5. Frequency distribution of the mRS scores

Modified Rankin Scale †	1 month	6 months	12 months	24 months	Total
0 No symptoms	144 (16.9)	153 (18.6)	103 (14.7)	61 (13.7)	461 (16.2)
1 No significant disability	282 (33.1)	278 (33.9)	240 (34.1)	144 (32.4)	944 (33.5)
2 Slight disability	165 (19.3)	150 (18.3)	161 (22.9)	114 (25.7)	590 (20.7)
3 Moderate disability	127 (14.0)	158 (19.2)	138 (19.6)	88 (19.8)	511 (18.3)
4 Moderately severe disability	90 (10.6)	56 (6.8)	46 (6.5)	27 (6.1)	219 (7.9)
5 Severe disability	45 (5.3)	26 (3.2)	15 (2.1)	10 (2.3)	96 (3.4)
0-2 'Independent'	591 (69.3)	581 (70.8)	504 (71.7)	319 (71.8)	1995 (70.7)
3-5 'Severely disabled'	262 (30.7)	240 (29.2)	199 (28.3)	125 (28.2)	826 (29.3)
Total	853	821	703	444	**2821**

†Modified Rankin Scale scores of 6 (=dead) were excluded as deaths are reported with the missing values in Table3.

Table 6. Frequency of EQ-5D response levels

Problem	Mobility	Self care	Activities	Pain	Anxiety
	1 month				
No	382 (49.5)	563 (73.0)	406 (52.6)	473 (61.7)	506 (66.0)
Some	350 (45.3)	173 (22.4)	258 (33.4)	262 (34.2)	229 (29.9)
Extreme	40 (5.2)	35 (4.5)	108 (14.0)	32 (4.2)	32 (4.2)
Total	772	771	772	767	767
	6 months				
No	305 (52.4)	445 (76.6)	335 (57.8)	343 (59.0)	404 (69.7)
Some	262 (45.0)	115 (19.8)	194 (33.5)	207 (35.6)	157 (27.1)
Extreme	15 (2.6)	21 (3.6)	51 (8.8)	31 (5.3)	19 (3.3)
Total	582	581	580	581	580
	12 months				
No	314 (46.9)	493 (73.7)	361 (54.0)	389 (58.2)	477 (71.5)
Some	349 (52.2)	159 (23.8)	253 (37.8)	252 (37.7)	175 (26.2)
Extreme	6 (0.9)	17 (2.5)	55 (8.2)	27 (4.0)	15 (2.3)
Total	669	669	669	668	667
	24 months				
No	191 (44.6)	312 (72.9)	224 (52.3)	240 (56.1)	290 (67.8)
Some	225 (52.6)	98 (22.9)	162 (37.9)	170 (39.7)	127 (29.7)
Extreme	12 (2.8)	18 (4.2)	42 (9.8)	18 (4.2)	11 (2.6)
Total	428	428	428	428	428
	Total				
No	1192 (48.6)	1813 (74.0)	1326 (54.1)	1445 (59.1)	1677 (68.7)
Some	1186 (48.4)	545 (22.3)	867 (35.4)	891 (36.5)	688 (28.2)
Extreme	73 (3.0)	91 (3.7)	256 (10.5)	108 (4.4)	77 (3.2)
Total	**2451**	**2449**	**2449**	**2444**	**2442**

Table 7. Mean EQ-5D utility values by mRS score

mRS†	N (%)	Mean (Std. Dev.)
	1 month	
0	132 (17.3)	0.909 (0.153)
1	269 (35.3)	0.819 (0.181)
2	157 (20.6)	0.694 (0.220)
3	116 (15.2)	0.589 (0.280)
4	70 (9.2)	0.252 (0.301)
5	18 (2.4)	-0.112 (0.195)
Total	762	
	6 months	
0	109 (19.0)	0.931 (0.117)
1	211 (36.8)	0.803 (0.194)
2	119 (20.8)	0.682 (0.220)
3	98 (17.1)	0.547 (0.308)
4	31 (5.4)	0.236 (0.313)
5	5 (0.9)	0.101 (0.519)
Total	573	
	12 months	
0	101 (15.2)	0.961 (0.118)
1	238 (35.8)	0.832 (0.178)
2	157 (23.6)	0.684 (0.190)
3	129 (19.4)	0.547 (0.275)
4	38 (5.7)	0.323 (0.277)
5	2 (0.3)	-0.069 (0.137)
Total	665	
	24 months	
0	61 (14.4)	0.959 (0.074)
1	143 (33.6)	0.812 (0.181)
2	111 (26.1)	0.656 (0.218)
3	82 (19.3)	0.545 (0.277)
4	24 (5.6)	0.248 (0.281)
5	4 (0.9)	0.02 (0.046)
Total	425	
0	403 (16.6)	0.936 (0.127)
1	861 (35.5)	0.817 (0.183)
2	544 (22.4)	0.681 (0.211)
3	425 (17.5)	0.558 (0.284)
4	163 (6.7)	0.265 (0.294)
5	29 (1.2)	-0.054 (0.264)
Total	**2425**	

†mRS scores of 6 (=dead) were excluded as they corresponded to an EQ-5D tariff of 0

Table 8. Mean EQ-5D utility change stratified by mRS score changes

Type of mRS change		N	Mean (Std. Dev.)	Type of mRS change		N	Mean (Std. Dev.)
1	4→5	2	-0.446 (0.279)	**-1**	5 →4	6	0.324 (0.367)
	3→4	21	-0.156 (0.339)		4→3	33	0.164 (0.355)
	2→3	28	-0.055 (0.356)		3→2	34	0.144 (0.249)
	1→2	79	-0.063 (0.266)		2→1	50	0.077 (0.263)
	0→1	49	-0.117 (0.225)		1→0	45	0.128 (0.180)
	Total	**179**	**-0.092 (0.283)**		**Total**	**168**	**0.130 (0.268)**
2	3→5	0		**-2**	5 →3	2	0.402 (0.240)
	2→4	3	-0.350 (0.133)		4→2	1	0.596 (0.297)
	1→3	9	-0.320 (0.286)		3→1	10	0.288 (0.287)
	0→2	15	-0.232 (0.229)		2→0	11	0.172 (0.237)
	Total	**27**	**-0.275 (0.239)**		**Total**	**24**	**0.257 (0.264)**
3	2→5	0	-0.106 (0.049)	**-3**	5 →2	0	
	1→4	0	-0.190 (0.020)		4→1	1	-0.497 (.)†
	0→3	1	-0.740 (0.387)		3→0	2	-0.084 (0.270)
	Total	**1**	**-0.740 (.)†**		**Total**	**3**	**-0.222 (0.306)**
4	1→5	1	-0.972 (.)†	**-4**	5 →1	0	
	0→4	1			4→0	0	
	Total	**2**	**-1.008 (0.050)**		**Total**	**0**	
5	0→5	0		**-5**	5 →0	0	
	Total	**0**			**Total**	**0**	

† There is no standard deviation value since there is only one data point

Table 9. Mlogit and OLS regressions in-sample predictions of actual mean EQ-5D utility values *

Regression Model (explanatory variables)	N	Mean	MSE	MAE	Std. Dev.	Min	Max
Actual mean EQ-5D tariff	2438	**0.7135**	n/a†	n/a	0.2875	-0.594	1
OLS (mRS scores, age gender)	2425	**0.7134**	0.0448	0.1533	0.1938	-0.091	0.988
OLS (mRS scores)	2425	**0.7134**	0.0453	0.1543	0.1926	-0.054	0.936
Mlogit (mRS scores)	2425	**0.7063**	0.0736	0.1925	0.2653	-0.358	1
Mlogit (mRS, age gender)	2425	**0.7044**	0.0727	0.1900	0.2683	-0.484	1

*Results are presented in order of closest to the actual mean tariff
† n/a stands for not applicable

Figure 1 . Frequency distribution of the difference between actual and predicted EQ-5D utility values: a1) OLS with mRS scores, a2) OLS with mRS scores, age, and gender, b1) Mlogit with mRS scores, and b2) Mlogit with mRS scores age, and gender models

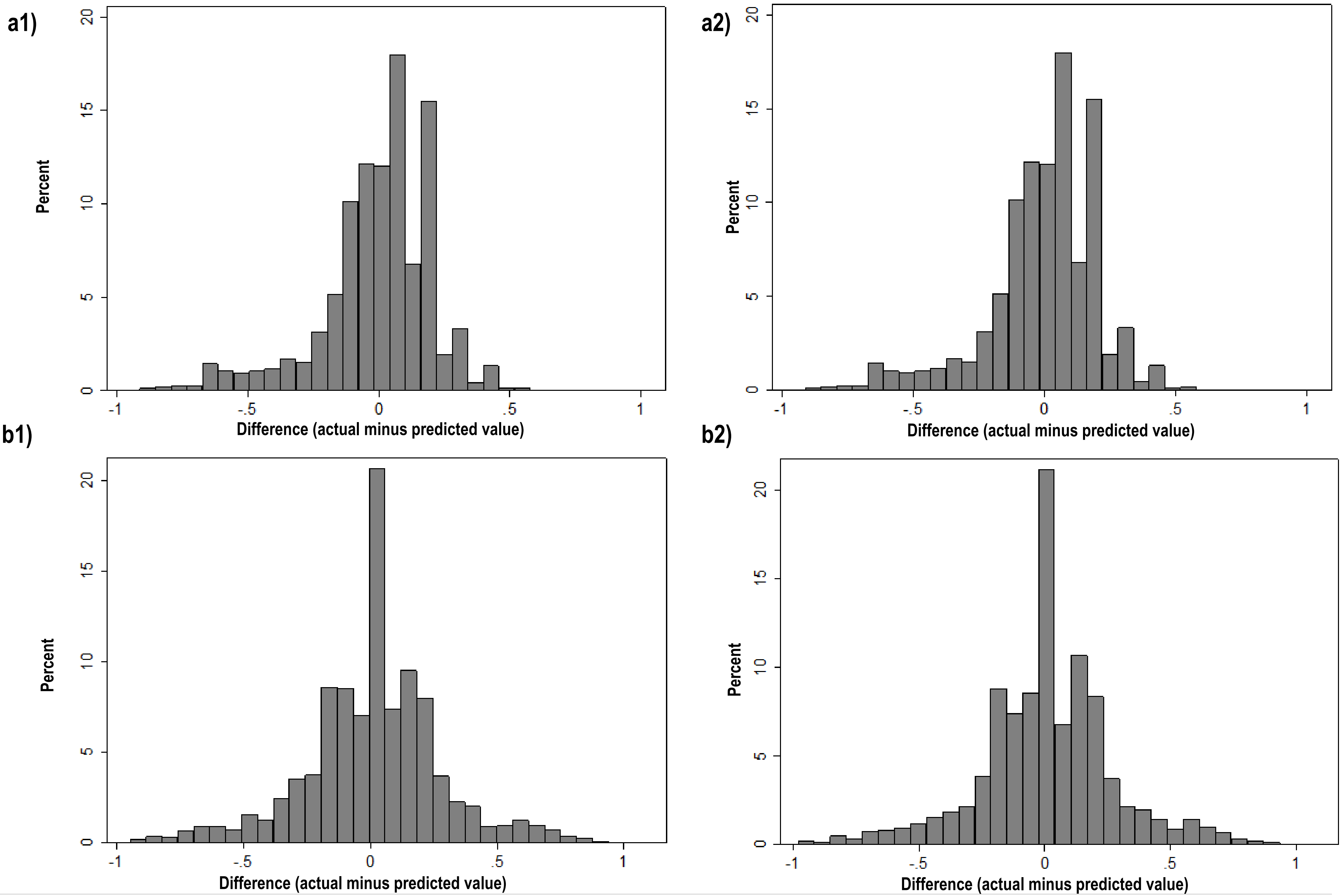

Figure 2. Frequency distribution of changes between actual and predicted responses levels

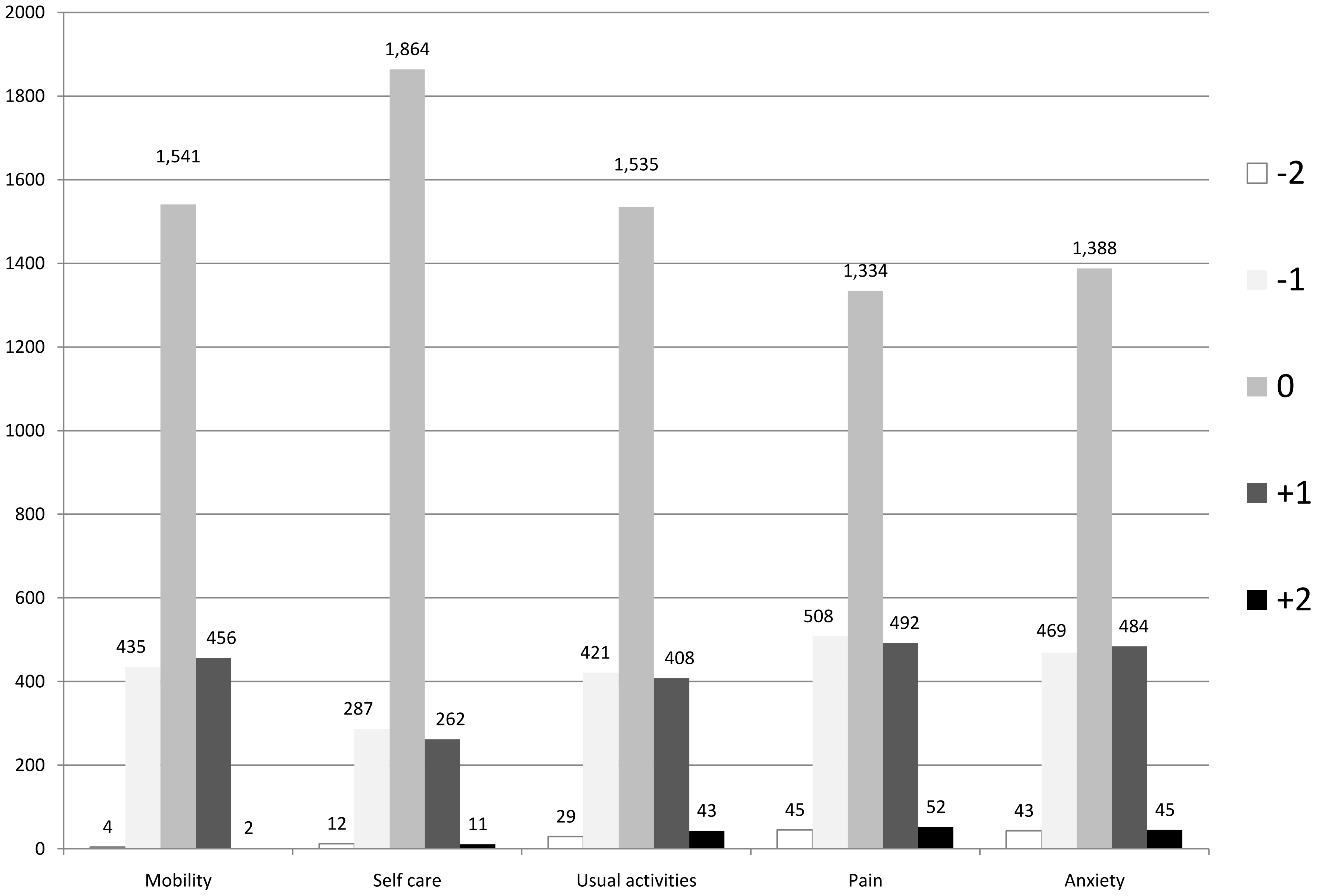

Figure3 . Quantile quantile plots: a1) OLS with mRS scores and a2) OLS with mRS scores, age, and gender, b1) Mlogit with mRS scores, and b2) Mlogit with mRS scores, age, and gender models

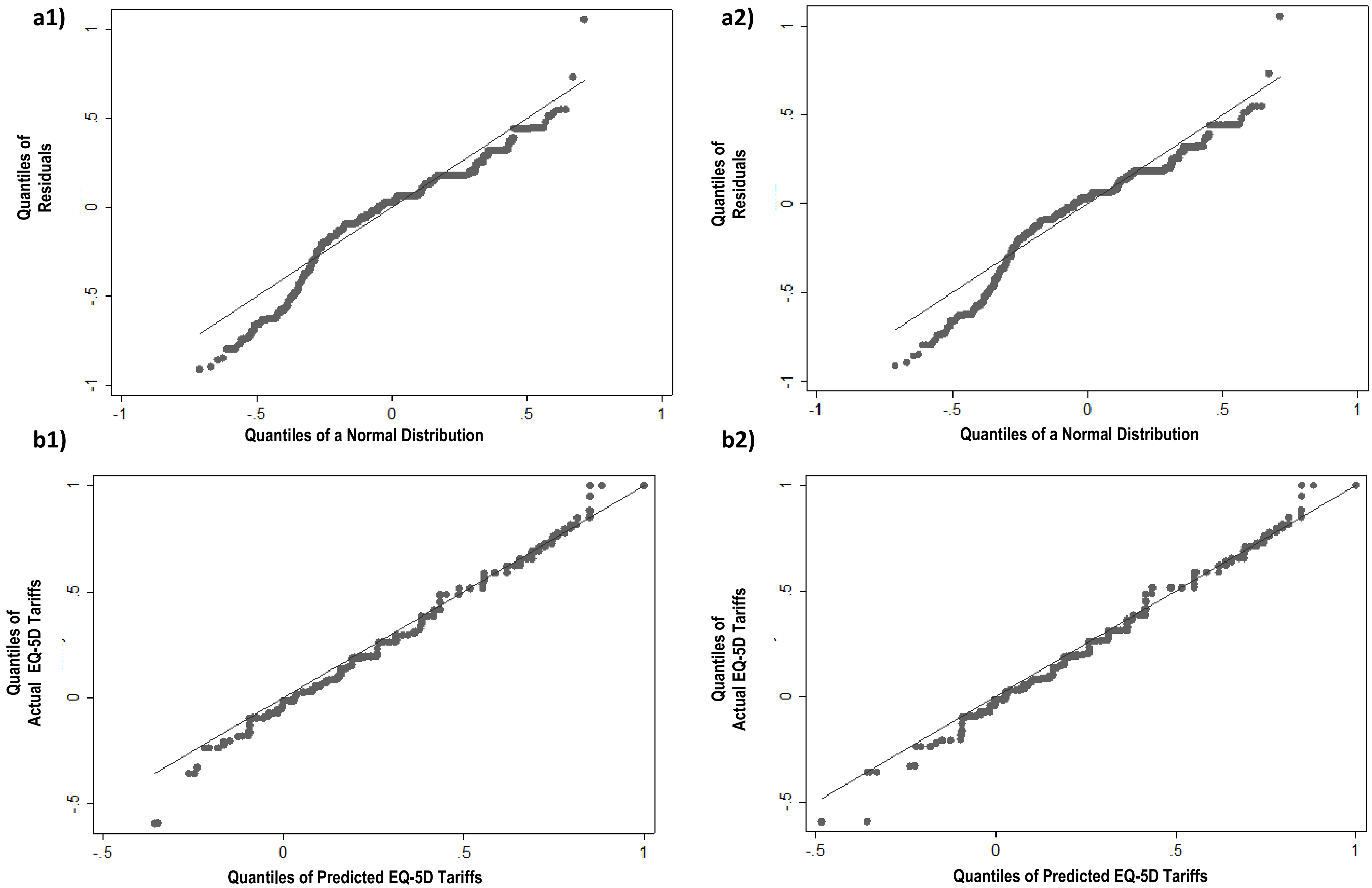

Table 10. Predicted parameter coefficients (standard errors): a) OLS with mRS, b) OLS with mRS, age, and gender, c) Mlogit with mRS, and d) Mlogit with mRS, age, and gender models †

	Coefficient
a)	
mRS 0	0.118 (0.012)*
mRS 2	-0.137 (0.008)*
mRS 3	-0.259 (0.013)*
mRS 4	-0.553 (0.013)*
mRS 5	-0.872 (0.057)*
intercept	0.817 (0.006)*

	Coefficient
b)	
mRS 0	0.121 (0.012)*
mRS 2	-0.139 (0.007)*
mRS 3	-0.262 (0.011)*
mRS 4	-0.554 (0.015)*
mRS 5	-0.872 (0.053)*
gender	0.036 (0.009)
age <45	-0.020 (0.033)
age 45 - <55	0.001 (0.015)
age 55 - <65	-0.016 (0.004)*
age 65 - <75	-0.005 (0.010)
age >85	0.030 (0.007)*
intercept	0.764 (0.012)

	Mobility		Self care		Activities		Pain		Anxiety	
	Some Problems	Extreme Problems	Some Problems	Extreme Problems	Some Problems	Extreme Problems	Some Problems	Extreme Problems	Some Problems	Extreme Problems
c)										
mRS 0	-1.831 (0.247)*	0.399 (1.620)	-1.898 (0.381)*	1.410 (1.609)	-1.585 (0.264)*	0.301 (0.987)	-1.634 (0.236)*	-2.133 (0.623)*	-0.695 (0.168)*	-0.686 (0.526)
mRS 2	1.320 (0.271)*	1.139 (1.639)	1.775 (0.260)*	0.668 (0.050)*	2.245 (0.142)*	3.384 (0.219)*	0.559 (0.115)*	0.772 (0.071)*	0.560 (0.118)*	0.891 (0.249)*
mRS 3	2.129 (0.299)*	-27.445 (1.092)*	3.539 (0.219)*	4.082 (1.181)*	2.691 (0.203)*	5.312 (0.170)*	0.118 (0.095)	1.389 (0.117)*	0.412 (0.114)*	1.294 (0.233)*
mRS 4	3.601 (0.275)*	8.187 (1.123)*	5.489 (0.164)*	8.403 (0.988)*	3.086 (0.418)*	7.483 (0.352)*	0.269 (0.224)	1.033 (0.078)*	0.410 (0.101)*	2.658 (0.176)*
mRS 5	0.589 (1.628)	9.023 (1.611)*	4.093 (0.681)*	9.234 (1.062)*	-32.019 (0.856)*	7.891 (0.524)*	-0.559 (0.492)	1.496 (0.567)*	0.598 (0.426)	3.485 (0.446)*
intercept	-0.589 (0.053)*	-6.315 (1.108)*	-2.995 (0.151)*	-6.708 (1.111)*	-1.419 (0.094)*	-5.150 (0.126)*	-0.439 (0.079)*	-3.054 (0.141)*	-1.040 (0.070)*	-4.045 (0.115)*
d)										
mRS 0	-1.758 (0.251)*	0.318 (1.596)	-1.910 (0.388)*	1.524 (1.683)	-1.605 (0.260)*	0.213 (1.000)	-1.651 (0.224)*	-2.233 (0.629)*	-0.745 (0.145)*	-0.781 (0.572)
mRS 2	1.299 (0.261)*	1.162 (1.622)	1.781 (0.259)*	0.583 (0.070)*	2.277 (0.131)*	3.436 (0.173)*	0.583 (0.115)*	0.847 (0.069)*	0.602 (0.124)*	0.979 (0.228)*
mRS 3	2.007 (0.320)*	-26.883 (1.083)*	3.530 (0.236)*	3.891 (1.228)*	2.717 (0.166)*	5.402 (0.182)*	0.153 (0.090)	1.517 (0.077)*	0.480 (0.110)*	1.444 (0.283)*
mRS 4	3.475 (0.285)*	8.274 (1.126)*	5.518 (0.190)*	8.364 (1.032)*	3.089 (0.394)*	7.572 (0.381)*	0.281 (0.246)	1.158 (0.082)*	0.465 (0.116)*	2.827 (0.198)*
mRS 5	0.242 (1.585)	9.084 (1.663)*	4.038 (0.612)*	9.140 (1.184)*	-31.73 (0.831)*	8.018 (0.412)*	-0.548 (0.530)	1.724 (0.608)*	0.679 (0.396)	3.663 (0.423)*
Gender	-0.483 (0.065)*	-0.271 (0.345)	-0.154 (0.166)	0.159 (0.270)	-0.372 (0.099)*	-0.144 (0.225)	-0.292 (0.103)*	-0.280 (0.107)*	-0.315 (0.115)*	-0.899 (0.166)*
age <45	-0.929 (0.443)	1.346 (0.491)*	0.773 (0.327)*	0.167 (2.344)	0.058 (0.208)	0.723 (0.433)	-0.146 (0.338)	0.554 (0.611)	0.127 (0.481)	-29.872 (0.937)*
age 45 - <55	-1.084 (0.322)*	-0.478 (0.714)	-0.774 (0.649)	-0.688 (0.905)	0.011 (0.070)	0.483 (0.538)	-0.130 (0.158)	-0.056 (0.335)	0.156 (0.166)	1.166 (0.636)
age 55 - <65	-0.436 (0.088)*	0.289 (0.647)	0.152 (0.121)	-0.004 (0.878)	0.004 (0.195)	0.350 (0.397)	-0.027 (0.085)	0.761 (0.241)*	0.177 (0.013)*	0.226 (0.351)
age 65 - <75	-0.446 (0.106)*	0.116 (0.394)	0.095 (0.139)	0.006 (0.484)	-0.063 (0.146)	0.068 (0.253)	0.006 (0.067)	0.206 (0.181)	0.045 (0.093)	0.550 (0.456)
age >85	0.028 (0.161)	-0.165 (0.394)	0.073 (0.109)	0.981 (0.334)*	-0.319 (0.083)*	-0.188 (0.129)	-0.448 (0.053)*	-0.568 (0.295)	-0.494 (0.171)*	-0.446 (0.722)
intercept	0.430 (0.120)*	-6.018 (1.139)*	-2.812 (0.375)*	-7.070 (1.555)*	-0.817 (0.208)*	-5.062 (0.441)*	0.058 (0.126)	-2.824 (0.234)*	-0.582 (0.187)*	-3.061 (0.539)*

† (mRS score of 1 and age bracket 75-85 were used as reference categories)
*Significant at 5%