The External Validity of Mapping MSIS-29 on EQ-5D Among Individuals With Multiple Sclerosis in Sweden

Olivia Ernstsson, MSc, Petter Tinghög, PhD, Kristina Alexanderson, PhD, Jan Hillert, MD, PhD, Kristina Burström, PhD

Background: *Mapping can be performed to predict utility values from condition-specific measures when preference-based measures are absent. A previously developed algorithm that predicts EQ-5D-3L index values from the Multiple Sclerosis Impact Scale (MSIS-29) has not yet been externally validated.* ***Aim:*** *To examine the external validity of a previously developed mapping algorithm by testing the accuracy of predicting EQ-5D-3L index values from MSIS-29 among multiple sclerosis (MS) patients in Sweden.* ***Methods:*** *Cross-sectional individual-level data were collected from population-based Swedish registers between 2011 and 2014. Health-related quality of life was assessed through MSIS-29 and EQ-5D-3L at one point in time among 767 individuals with known disability level of MS. A previously developed mapping algorithm was applied to predict EQ-5D index values from MSIS-29 items, and the predictive accuracy was assessed through mean absolute error and root mean square error.* ***Results:*** *When applying the algorithm, the predicted mean EQ-5D-3L index value was 0.77 compared to the observed mean index value of 0.75. Prediction error was higher for individuals reporting EQ-5D values <0.5 compared to individuals reporting EQ-5D values ≥0.5. Mean absolute error (0.12) and root mean square error (0.18) were smaller or equal to the prediction errors found in the original mapping study.* ***Conclusion:*** *The mapping algorithm had similar predictive accuracy in the two independent samples although results showed that the highest predictive performance was found in groups with better health. Varied predictive accuracy in subgroups is consistent with previous studies and strategies to deal with this are warranted.* ***Key words:*** *health-related quality of life; mapping; MSIS-29; EQ-5D; multiple sclerosis.* ***(MDM Policy & Practice 2017;2:1–9)***

Multiple sclerosis (MS) is a chronic disease of the central nervous system with an onset between 20 and 40 years of age. There are different variants of disease course, with approximately 85% being diagnosed with the relapsing-remitting form of MS (RRMS), which with time usually develops into the secondary-progressive MS (SPMS) course with slow but continuous deterioration. The less common primary-progressive MS (PPMS) course is characterized by gradual neurological deterioration from onset, whereas the progressive-relapsing MS (PRMS) is characterized by both worsening from onset and relapses.[1] Previous studies have described high rates of sick leave and disability pension (in this study combined and used as a proxy for work disability),[2] unemployment,[3] decreasing health-related quality of life (HRQoL),[4–6] and increasing societal costs[7] as the disease progresses. Thus, it is important to measure the burden of disease and, due to the heterogeneity of the patient population, to report variations of HRQoL across subgroups of individuals with MS (hereafter referred to as MS patients).

Functional disability due to MS is commonly assessed through the Expanded Disability Status Scale (EDSS).[8] Furthermore, there are HRQoL measures that include health aspects that are not limited to physical disability. Both generic and condition-specific patient-reported outcome measures (PROMs) can be used to assess HRQoL. A widely used example of a condition-specific measure is the Multiple Sclerosis Impact Scale (MSIS-29).[9] Generic measures can be preference-based and thus useful when conducting cost-utility analyses (CUA). In CUA, health interventions are compared regarding their costs and benefits, commonly using

DOI: 10.1177/2381468317692806

quality-adjusted life years (QALYs) as the outcome measure.[10] In a single value, QALYs combine life years and quality of life, where a quality weight of 0 is equivalent to dead and 1 is equivalent to full health. The use of a preference-based measure enables the calculation of QALYs as all health states can be assigned a quality weight (hereafter called index value). This can be exemplified by the EQ-5D-3L instrument where individual responses on five health dimensions with three severity levels can identify 243 health states, which can be assigned an index value from previously elicited valuations.[11]

As preference-based measures might not be accessible, statistical mapping techniques aiming to predict values for the quality adjustment are emerging in the scientific literature.[12] Mapping has been described as a method to predict values from other health outcomes measures that cover similar aspects of health.[13] An adequate mapping is conducted by assessing the overlap between two instruments, collecting responses on both instruments at the same time for the same individuals, developing an algorithm through statistical analyses, and by testing the performance of the algorithm.[14] Thus, responses on a condition-specific measure can be converted to responses on a preference-based measure. According to the NICE guidelines, the use of EQ-5D data is recommended as outcome measure for economic evaluation.[15] Mapping is a suggested solution that can facilitate economic evaluation when appropriate data for such purposes are not available.

The use of mapping to predict EQ-5D has been performed from several other condition-specific measures[12,16] as well as from the first[17] and second[18] versions of MSIS-29. Versteegh and colleagues developed a mapping algorithm in a sample consisting of 661 MS patients from 70 MS centers in the United Kingdom.[17] In order to study the generalizability of a mapping algorithm, its predictive performance ought to be examined in MS populations other than the one in which the mapping algorithm was developed.[17] The aim of this study was to examine the external validity of a mapping algorithm developed by Versteegh and colleagues[17] by testing the accuracy of predicting EQ-5D-3L index values from MSIS-29 among MS patients in Sweden.

METHODS

Data

Data were obtained from the clinically generated Swedish Multiple Sclerosis register (SMSreg) and register data from Statistics Sweden and the Swedish Social Insurance Agency. The SMSreg collects structured clinical information by its web-based decision support user interface and currently contains data on approximately 80% of the estimated prevalent MS population in Sweden.[19] From the SMSreg we used data on disability level of MS, type of progression, and HRQoL, that is, MSIS-29 v1 and EQ-5D-3L. In SMSreg, MSIS-29 v1 and EQ-5D-3L are typically assessed biannually for patients receiving novel disease modifying therapies (DMTs) introduced after 2006, but more occasionally for patients not on treatment or for patients on the first generation of DMTs, that is, interferon-beta and glatiramer. Calculations were made to

Received 25 November 2015 from the Health Outcomes and Economic Evaluation Research Group, Medical Management Centre, Department of Learning, Informatics, Management and Ethics (OE, KB), Division of Insurance Medicine, Department of Clinical Neuroscience (OE, PT, KA), Division of Neuro, Department of Clinical Neuroscience (JH), and Equity and Health Policy Research Group, Department of Public Health Sciences (KB), Karolinska Institutet, Stockholm, Sweden; The Swedish Red Cross University College, Stockholm, Sweden (PT); and Health Care Services, Stockholm County Council, Stockholm, Sweden (KB). Data analysis was performed at the Division of Insurance Medicine, Department of Clinical Neuroscience, Karolinska Institutet. An earlier version of the article was presented at the 36th Nordic Health Economist's Study Group meeting in Uppsala, Sweden, in September 2015. An earlier version of the article was presented during a seminar at Karolinska Institutet's Master Programme in Public Health Sciences, track Health Economics, Policy and Management. This study was presented as a poster at the European Committee for Treatment and Research in Multiple Sclerosis in Barcelona, Spain, in October 2015. Olivia Ernstsson and Petter Tinghög were funded from an unrestricted research grant from Biogen. Kristina Alexanderson has received unrestricted research grants from Biogen and from the Swedish Research Council for Working Life, Health and Welfare. Jan Hillert received honoraria for serving on advisory boards for Biogen, Sanofi-Genzyme, and Novartis and speaker's fees from BiogenIdec, Merck-Serono, Bayer-Schering, Teva, Novartis, and Sanofi-Genzyme. He has served as principal investigator for projects sponsored by, or received unrestricted research support from, Biogen, Merck-Serono, TEVA, Novartis, and Bayer-Schering. His MS research is funded by the Swedish Research Council and the Swedish Brain Foundation. Kristina Burström has no conflict of interest to declare. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Revision accepted for publication 15 December 2016.

The online appendixes for this article are available on the *Medical Decision Making Policy & Practice* Web site at http://journals.sagepub.com/home/mpp.

Address correspondence to Olivia Ernstsson, Karolinska Institutet, Tomtebodavägen 18A, Stockholm SE-171 77, Sweden; e-mail: olivia.ernstsson@ki.se.

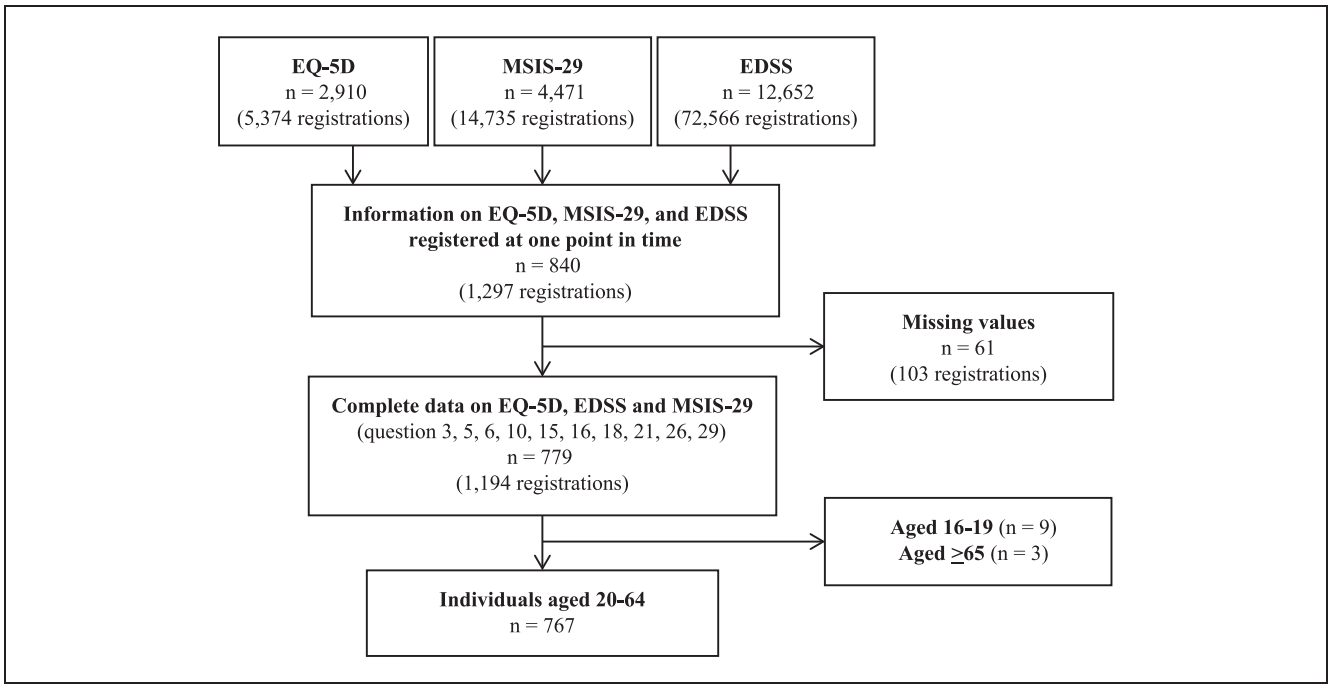

Figure 1 Flow chart of study sample selection

obtain the physical impact score (MSIS-29 Items 1–20) and psychological impact score due to MS (MSIS-29 Items 21–29)[20] and to obtain EQ-5D index values based on the value set presented in the UK MVH study.[21] Categorizations were made according to MS disability level (EDSS) normal (0), mild (1–3.5), moderate (4–5.5), and severe (6–8.5). Course of MS was categorized into RRMS, SPMS, PPMS, PRMS, and unknown. Days with work disability included both sick leave and disability pension days. Data on sociodemographic characteristics regarding year 2011 were obtained from Statistics Sweden (see Online Appendix 1 for more details on data).

Study Population

Inclusion criteria were 1) a diagnosis of MS and 2) having complete data on the three instruments (i.e., EQ-5D, EDSS, and the MSIS-29 questions relevant for validation) and all had to be recorded at the same date (Figure 1). In cases where one individual had more than one visit fulfilling these criteria, data from the earliest visit was chosen. Individuals aged 20 to 64 years were included in order to get a study population of working age, resulting in a final sample of 767 MS patients. Patient data were from year 2011 (n = 31), 2012 (n = 208), 2013 (n = 325), and 2014 (n = 203).

Instruments

The Multiple Sclerosis Impact Scale–29

The MSIS-29 is a self-administered PROM with 29 items (20 items covering physical aspects and 9 items covering psychological aspects) and aims to assess HRQoL in individuals with MS.[9] MSIS-29 (v1) measures the impact in five levels (not at all, a little, moderately, quite a bit, extremely) where respondents recall the impact on their HRQoL due to MS during the past 2 weeks.[9]

EQ-5D

The EQ-5D-3L is a generic instrument that measures HRQoL in five health dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) with three severity levels (no, moderate, and severe problems).[11] A total of 243 (3^5) health profiles can be identified. All health profiles can be assigned an index value, which facilitates the calculation of QALYs.

Expanded Disability Status Scale

The EDSS is one of the most commonly used measures to assess the degree of neurological impairment due to MS and ranges from 0 (*normal*

neurologic examination) to 10 (*death due to MS*).[8,22] The EDSS examines eight functional systems: pyramidal, cerebellar, brain stem, sensory, bowel and bladder, visual, cerebral, and other functions.[8] In contrast to the previously mentioned instruments, the EDSS assessment is performed by a physician[22] and is thus not an example of a measure of HRQoL from the patient perspective.

Analyses

Mean EQ-5D index values and mean EDSS values were presented for the total sample and in subgroups of MS patients.

The mapping developed by Versteegh and others has been described in detail elsewhere.[17] Their sample was randomly divided into a mapping development sample and two test samples in order to study the performance of several ordinary least squares (OLS) regression models. The final mapping algorithm includes 10 items from MSIS-29 (seven from the physical component and three items from the psychological component) to predict EQ-5D index values. These items concern questions about having limited ability to carry things (Item 3), and being bothered by difficulties moving indoors (Item 5), being clumsy (Item 6), spasms in your limbs (Item 10), difficulties using your hands in everyday tasks (Item 15), having to cut down on the amount of time you spent on work or other daily activities (Item 16), taking longer to do things (Item 18), feeling unwell (Item 21), feeling irritable, impatient, or short tempered (Item 26), and feeling depressed (Item 29).[9]

In the present study, the following method was applied to validate the previously developed mapping algorithm[17] in the Swedish MS sample. The dependent variable, that is, the EQ-5D index value, was calculated according to the UK MVH value set,[21] which is the most commonly used value set in Sweden. Ten items from the MSIS-29 were coded as dummy variables in accordance with the original mapping algorithm. The first number describes the item and the following numbers describes the response on that item, that is, MSI6_5 defines a response of 5 on Item 6 (extremely bothered by being clumsy) and MSI3_3_4 defines a response of 3 or 4 on Item 3 (moderately or quite a bit limited ability to carry things).[9]

The predictive performance was assessed by comparing predicted EQ-5D index values with observed EQ-5D index values, where the predictive accuracy was assessed by mean absolute error (MAE), root mean square error (RMSE), and RMSE normalized for range. MAE is the sum of individual absolute errors divided by the number of individuals, whereas RMSE is calculated by initially dividing the sum of individual squared errors by the number of individuals, and then by calculating the square root of the mean squared error.[18,23] MAE and RMSE are estimates of prediction errors, and thus, low values indicate high predictive accuracy. As RMSE is more influenced by prediction errors that considerably deviate from the mean value, RMSE is always equivalent to or greater than MAE.[17] MAE and RMSE are, nevertheless, not constructed to compare results from using instruments with different ranges. Thus, the RMSE normalized for range was reported to facilitate comparison of study results.[17,24] The RMSE was divided by the range of EQ-5D index values, where the UK value set has a possible range between -0.594 and 1. Thus, the normalized RMSE reports prediction error as a percentage of the scale size of the instrument applied.[17,18] Mean predictive accuracy was assessed for the entire sample. Additionally, in order to examine variations in how the mapping algorithm performs in subgroups with regard to health status, prediction error was also examined for groups categorized according to severity levels on EQ-5D, that is, index values ≥ 0.5 and <0.5, and MS disability level, that is, mild, moderate, and severe MS. This was performed in accordance with the suggestion to report patterns of error over the entire range of EQ-5D,[12] as several mapping studies indicate lower predictive accuracy among individuals reporting EQ-5D index values <0.5.[12,16] As EQ-5D data are absent when a conversion from MSIS-29 is performed, a strategy to identify variations in predictive accuracy between subgroups of health status, defined by MS disability level, was explored.

All statistical analyses were performed in IBM SPSS Statistics v.22, and the predictive accuracy was calculated in Microsoft Excel. The research project was approved by the Regional Ethical Review Board, Stockholm (Dnr 2007/5:6; 2011/1710-32; 2014/236-32).

RESULTS

In the total study sample of 767 MS patients, 70% were women (Table 1). The majority had mild MS disability level at the time of measure (84%), and the mean EDSS of 2.3 was equivalent for men and women. Most were diagnosed with RRMS (87%) and reported better health compared to those

Table 1 Characteristics of the Study Population (n = 767), Mean MS Disability Level (EDSS), and Mean EQ-5D Index Values

	%	*n*	EDSS, Mean (*SD*)	EQ-5D Index, Mean (*SD*)
Total sample	100	767	2.3 (1.7)	0.75 (0.25)
Men	29.5	226	2.3 (1.8)	0.79 (0.24)
Women	70.5	541	2.3 (1.7)	0.74 (0.25)
Age (years)				
20–24	4.6	35	1.1 (1.1)	0.77 (0.27)
25–34	22.4	172	1.9 (1.5)	0.76 (0.27)
35–44	37.5	288	2.2 (1.8)	0.75 (0.27)
45–54	29.1	223	2.6 (1.7)	0.75 (0.22)
55–64	6.4	49	3.8 (1.9)	0.71 (0.20)
Educational level				
Low	7.4	57	2.7 (1.7)	0.67 (0.27)
Medium	49.5	380	2.4 (1.8)	0.73 (0.27)
High	42.8	328	2.1 (1.7)	0.80 (0.22)
Missing	0.3	2	4.0 (1.4)	0.22 (0.20)
Country of birth				
Sweden	89.8	689	2.3 (1.7)	0.76 (0.24)
Other	10.2	78	2.5 (1.6)	0.67 (0.31)
MS disability[a]				
Normal	15.3	117	0.0 (0.0)	0.89 (0.15)
Mild	68.7	527	2.1 (0.8)	0.77 (0.23)
Moderate	9.0	69	4.5 (0.6)	0.58 (0.24)
Severe	7.0	54	6.4 (0.6)	0.50 (0.35)
Type of MS				
RRMS	87.1	668	2.1 (1.5)	0.77 (0.24)
SPMS	6.9	53	4.6 (1.7)	0.62 (0.28)
PPMS	0.7	5	5.1 (2.2)	0.61 (0.25)
PRMS	1.7	13	4.7 (2.5)	0.51 (0.38)
Unknown	3.7	28	2.2 (1.7)	0.73 (0.32)
Work disability[b]				
0 days	49.2	377	1.7 (1.3)	0.83 (0.20)
1–90 days	15.3	117	2.1 (1.4)	0.78 (0.20)
91–180 days	7.6	58	2.6 (1.5)	0.66 (0.24)
181–364 days[c]	17.6	135	3.1 (1.9)	0.66 (0.26)
365 days[c]	10.4	80	4.1 (2.0)	0.55 (0.32)

Note: EDSS = Expanded Disability Status Scale; MS = multiple sclerosis, RRMS = relapsing-remitting MS; SPMS = secondary-progressive MS; PPMS = primary-progressive MS; PRMS = progressive-relapsing MS.
a. Measured by the Expanded Disability Status Scale: normal, 0; mild, 1 to 3.5; moderate, 4 to 5.5; severe, 6 to 8.5.
b. Net days with sickness absence or disability pension in the previous year.
c. 181–365 days and 366 days for year 2012, due to leap year.

diagnosed with other types of MS. During the preceding year, approximately 50% had no work disability days while 10% had full work disability. More severe MS was found in the highest age group (mean EDSS 3.8), among those diagnosed with PPMS (mean EDSS 5.1), PRMS (mean EDSS 4.7), SPMS (mean EDSS 4.6), and among those with full-time work disability (mean EDSS 4.1).

The mean EQ-5D index value was 0.75 with lower mean values among women, in the highest age group, among groups with low educational level, born outside Sweden, with severe MS, and with full-time work disability (Table 1). The health dimensions with most frequent reporting of problems were pain/discomfort (56%) and depression/anxiety (46%), while a smaller proportion reported problems related to mobility (31%), usual activities (28%), and self-care (6%) (see Online Appendix 2 for more details). The age range was narrower in our study (20–64 years) compared to the UK samples in the original mapping study by Versteegh and others[17] (18–88 years). A greater proportion of the Swedish sample were diagnosed with RRMS (87%) compared to the development and test samples in

Table 2 Summary of the Accuracy of Predicting EQ-5D Index Values from MSIS-29 in the Swedish MS Sample

	n	Observed[a], Mean (SD)	Predicted[a], Mean (SD)	Observed Min-Max	Predicted Min-Max	MAE	RMSE	RMSE, %
Total sample	767	0.75 (0.25)	0.77 (0.17)	−0.36 to 1	0.03 to 0.95	0.12	0.18	11.0
EQ-5D								
≥0.5	693	0.82 (0.14)	0.80 (0.15)	0.52 to 1	0.17 to 0.95	0.10	0.12	7.6
<0.5	74	0.12 (0.17)	0.50 (0.19)	−0.36 to 0.43	0.03 to 0.87	0.39	0.42	26.6
EDSS								
Normal	117	0.89 (0.15)	0.89 (0.09)	0.19 to 1	0.59 to 0.95	0.09	0.11	6.8
Mild	527	0.77 (0.23)	0.78 (0.16)	−0.24 to 1	0.03 to 0.95	0.12	0.16	10.2
Moderate	69	0.58 (0.24)	0.59 (0.16)	−0.24 to 0.85	0.10 to 0.86	0.15	0.22	14.0
Severe	54	0.50 (0.35)	0.58 (0.16)	−0.36 to 1	0.09 to 0.95	0.21	0.30	18.6

Note: MAE = mean absolute error; RMSE = root mean square error; EDSS = Expanded Disability Status Scale: normal, 0; mild, 1 to 3.5; moderate, 4 to 5.5; severe, 6 to 8.5.
[a]EQ-5D index value.

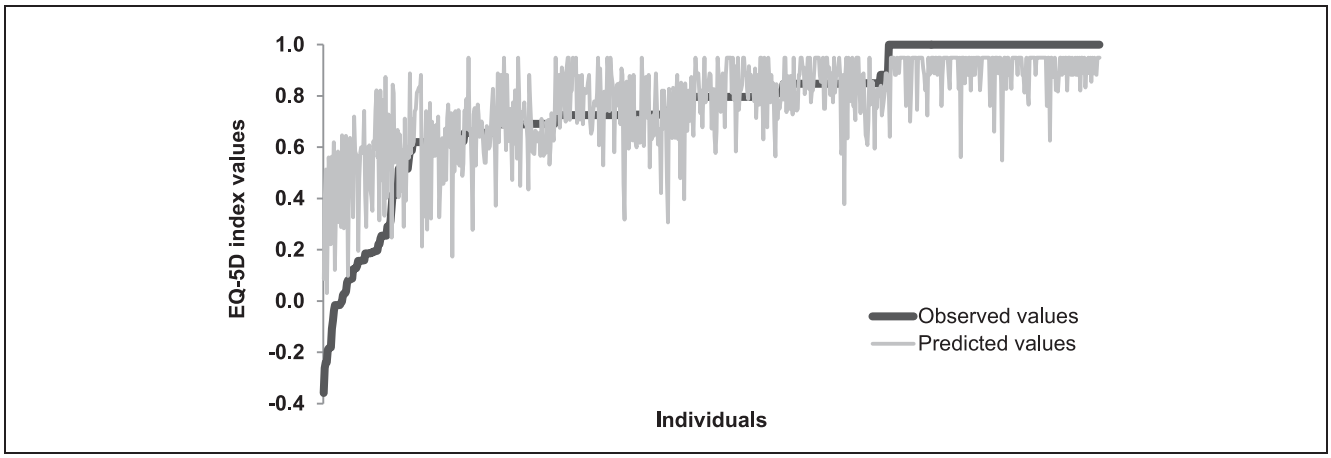

*Figure 2 Observed and predicted EQ-5D index values (UK value set), by individual responses (*n *= 767).*

the original mapping study (81% to 82%). The mean EQ-5D index value was higher in the present study (0.75 compared to ≤0.60), with a lower proportion reporting problems related to their mobility, self-care, usual activities, pain/discomfort, and depression/anxiety. The mean MSIS-29 physical impact score was 20.7 in the present study, whereas the corresponding score was 47.1 in the model development sample. Also, the mean MSIS-29 psychological impact score was lower compared to the mean score reported in the original mapping study.[17]

The mean observed EQ-5D index value was 0.75, and the mean predicted value was 0.77 for the total Swedish sample (MAE 0.12; RMSE 0.18; Table 2). These prediction errors were smaller or equal to the prediction errors found in the original mapping study. By studying absolute errors individually in the Swedish MS sample, 21% of the sample had absolute errors smaller than 0.05 and 56% had errors smaller than 0.10. Thus, 84% had an absolute difference between predicted and observed values of <0.20, leaving 16% (n = 123) with errors equal to or greater than 0.20 (not shown in tables).

The minimum and maximum values were −0.36 and 1 for the observed EQ-5D values, and 0.03 to 0.95 for the predicted EQ-5D values (Table 2). Prediction errors were on average higher for those reporting EQ-5D values <0.5 (MAE 0.39; RMSE 0.42), compared to those reporting EQ-5D values ≥0.5 (MAE 0.10; RMSE 0.12). The predicted values were generally lower than observed values among individuals reporting full health, that is, EQ-5D index = 1, whereas the predicted values were higher than the observed values among individuals reporting poor health, that is, low EQ-5D index values (Figure 2). Prediction errors were also smaller

among individuals with the lowest level of MS disability, that is, EDSS 0 (MAE 0.09; RMSE 0.11), compared to those among individuals with the highest disability level, that is, EDSS 6 to 8.5 (MAE 0.21; RMSE 0.30).

DISCUSSION

This study aimed to examine the external validity of a previously developed mapping algorithm by testing the accuracy of predicting EQ-5D-3L index values from MSIS-29 among MS patients in Sweden. When applying the previously developed mapping algorithm[17] in the Swedish MS sample, the predicted mean EQ-5D index value was 0.77 compared to the observed mean index value of 0.75. Predictions were more accurate for MS patients reporting better health (EQ-5D $\geq$0.5) compared to predictions among those reporting poor health (EQ-5D <0.5). Prediction errors were lower in the Swedish MS sample (MAE 0.12) compared to the mapping development sample (MAE 0.13).[17]

The underestimation of index values for individuals with good health and overestimation of values for individuals with poor health were also found in the original mapping study by Versteegh and colleagues and in several other mapping studies.[12,16–18,25] The results indicate that mean prediction errors were smaller in the Swedish sample compared to that in the original mapping study,[17] which appear to be an effect from the relatively smaller proportion of individuals reporting poor health (see the Online Appendixes for more details on reported problems on the EQ-5D dimensions). The large decrement of the total index value resulting from the response of severe problems in any of the EQ-5D health dimensions has previously been suggested as an explanation for overestimating low index values.[16]

One of the criteria to achieve high predictive accuracy is to have substantial overlap between the concepts of health being measured by the instruments.[12,14] The HRQoL data used in this study indicated that ratings of MSIS-29 and EQ-5D could differ substantially although both were answered at the same point in time. This variation, which results in low predictive accuracy, might be explained by patients adopting a wider perspective of health when answering the generic EQ-5D, as the MSIS-29 explicitly asks the respondent to reply on the impact due to MS. Moreover, EQ-5D measures the current state of health whereas MSIS-29 has a recall period of 2 weeks.[9,11] It is possible that responses vary depending on recall period as HRQoL can vary between time periods, especially for those with a high disability level of MS.

Having a large proportion of the sample reporting full health might contribute to that OLS regression is inadequate for capturing predictions of lower index values, and other statistical methods have been suggested to address this.[18] However, previous findings of mapping MSIS-29 on EQ-5D suggest that OLS gives equal or better predictive accuracy compared to when using methods that are more suitable for data that are not normally distributed,[18] and OLS regression has been the most commonly used statistical method to predict generic preference-based measures from non–preference-based measures.[12]

The criterion of having EDSS registered at the same point in time as the measurement of HRQoL gives information regarding the disability level of the sample. This is considered a strength in terms of implications for policy and practice as it becomes easier to compare this study population with others in which a mapping is of interest. However, a relatively small proportion of the study sample have severe MS, that is, EDSS > 6. Thus, we chose to analyze according to EDSS categories (normal, mild, moderate, severe) when analyzing predictive accuracy in subgroups, as opposed to more detailed EDSS levels presented in the original mapping study. Another related limitation is the large proportion of the study sample with a self-assessment of moderate to full health, that is, with high predictive accuracy. The study population is thus not necessarily representative of the estimated prevalent MS population in Sweden, but more likely representative of those on MS treatment. The testing of the mapping algorithm is conducted in a sample in which prediction of values for the purpose of economic evaluation is relevant. The results should, however, not be generalized to all MS patients in Sweden.

As EQ-5D index values are absent when a conversion from MSIS-29 is conducted, it is challenging to identify individuals and subgroups with low predictive accuracy. Prediction error was not found to the same extent among individuals with the most severe disability level of MS, that is, according to EDSS, as were found among those reporting poor health on EQ-5D. A more accurate mapping algorithm could possibly be developed by testing response mapping, which is done by predicting EQ-5D responses (no, moderate, and severe problems) rather than index values. This has been shown to be a successful strategy when mapping from other condition-specific instruments onto EQ-5D.[25]

However, regardless of its performance, mapping will not solve the issue of having a generic preference-based measure that fails to address health dimensions of relevance for the patient population of interest.[12] Yet another option is to develop preference-based condition-specific measures.[10,26]

The original mapping study addressed the majority of items in the MAPS Statement,[27] which facilitates an external validation of the mapping algorithm. However, as described in the checklist, there is a need to further "outline the clinical and research settings in which the mapping algorithm could be used."[27] To the best of our knowledge, there is no established threshold from which a mapping is considered adequate. Versteegh recently tested the mapping algorithm[17] through a Markov model in order to study its impact on the incremental cost effectiveness ratio (ICER).[28] The presented findings were high accuracy among individuals with mild to moderate MS (EDSS 0–7.5), and statistically significant differences between observed and mapped values (0.19 and 0.33, respectively) among those with severe MS (EDSS 8–9.5). Using mapped EQ-5D values resulted in an ICER that was 37% higher than the ICER based on observed EQ-5D values. Another study examining the effect of using mapped EQ-5D values from a condition-specific measure developed for osteoarthritis patients found that although mapped values and observed values differed, both led to the same conclusion of the most cost-effective intervention.[29] The issue of systematic differences in accuracy depending on self-reported health status is a concern for the practical use of mapping algorithms. In order to provide well-informed recommendations, further studies are needed to investigate the potential effects on different subgroups of patients in decision making and to investigate whether the systematic differences reveal a general issue of mapping onto the EQ-5D index.

The findings from examining the external validity of a mapping algorithm developed by Versteegh and others[17] among MS patients in Sweden indicate that the mapping algorithm has similar accuracy in two independent samples, although the highest predictive performance was found in groups with better health. Measurements from generic preference-based measures are preferred as the conversion between two instruments inevitably leads to prediction error. In cases of not having access to preference-based measures in already collected data, the mapping algorithm is preferably used for individuals with moderate to full health. However, strategies to deal with low predictive accuracy for those in poor health are warranted.

ACKNOWLEDGMENTS

We thank Matthijs Versteegh for kindly responding to questions about the original mapping study, and Michael Wiberg for valuable discussions on statistical methods. We thank all participants in the discussions of a previous version of the article at the 36th Nordic Health Economists' Study Group Meeting. Comments on an earlier version of this article during seminars held as part of the Karolinska Insitutet's Master Programme in Public Health Sciences track Health Economics, Policy and Management are gratefully acknowledged.

REFERENCES

1. Vollmer T. The natural history of relapses in multiple sclerosis. J Neurol Sci. 2007;256(Suppl. 1):S5–13.

2. Tinghog P, Hillert J, Kjeldgard L, Wiberg M, Glaser A, Alexanderson K. High prevalence of sickness absence and disability pension among multiple sclerosis patients: a nationwide population-based study. Mult Scler. 2013;19(14):1923–30.

3. Julian LJ, Vella L, Vollmer T, Hadjimichael O, Mohr DC. Employment in multiple sclerosis. Exiting and re-entering the work force. J Neurol. 2008;255(9):1354–60.

4. Karampampa K, Gustavsson A, Miltenburger C, Eckert B. Treatment experience, burden and unmet needs (TRIBUNE) in MS study: results from five European countries. Mult Scler. 2012; 18(2 Suppl.):7–15.

5. Naci H, Fleurence R, Birt J, Duhig A. The impact of increasing neurological disability of multiple sclerosis on health utilities: a systematic review of the literature. J Med Econ. 2010;13(1):78–89.

6. Kobelt G, Berg J, Lindgren P, Fredrikson S, Jonsson B. Costs and quality of life of patients with multiple sclerosis in Europe. J Neurol Neurosurg Psychiatry. 2006;77(8):918–26.

7. Naci H, Fleurence R, Birt J, Duhig A. Economic burden of multiple sclerosis: a systematic review of the literature. Pharmacoeconomics. 2010;28(5):363–79.

8. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983; 33(11):1444–52.

9. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain. 2001;124(Pt 5):962–73.

10. Brazier J. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford, NY: Oxford University Press; 2007.

11. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337–43.

12. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ. 2010;11(2):215–25.

13. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. Value Health. 2013;16(1):202–10.

14. Petrillo J, Cairns J. Converting condition-specific measures into preference-based outcomes for use in economic evaluation. Expert Rev Pharmacoecon Outcomes Res. 2008;8(5):453–61.

15. National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal 2013: Process and Methods Guide. London: National Institute for Health and Care Excellence; 2013.

16. Versteegh MM, Rowen D, Brazier JE, Stolk EA. Mapping onto Eq-5 D for patients in poor health. Health Qual Life Outcomes. 2010;8:141.

17. Versteegh MM, Leunis A, Luime JJ, Boggild M, Uyl-de Groot CA, Stolk EA. Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D. Med Decis Making. 2012;32(4):554–68.

18. Hawton A, Green C, Telford C, Zajicek J, Wright D. Using the Multiple Sclerosis Impact Scale to estimate health state utility values: mapping from the MSIS-29, version 2, to the EQ-5D and the SF-6D. Value Health. 2012;15(8):1084–91.

19. Hillert J, Stawiarz L. The Swedish MS registry—clinical support tool and scientific resource. Acta Neurol Scand Suppl. 2015; 132(199):11–9.

20. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. Health Technol Assess. 2004;8(9):iii, 1–48.

21. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095–108.

22. Hoogervorst EL, Zwemmer JN, Jelles B, Polman CH, Uitdehaag BM. Multiple Sclerosis Impact Scale (MSIS-29): relation to established measures of impairment and disability. Mult Scler. 2004;10(5):569–74.

23. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Res. 2005;30(1):79–82.

24. Kontodimopoulos N, Aletras VH, Paliouras D, Niakas D. Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15D instruments. Value Health. 2009; 12(8):1151–7.

25. Arnold DT, Rowen D, Versteegh MM, Morley A, Hooper CE, Maskell NA. Testing mapping algorithms of the cancer-specific EORTC QLQ-C30 onto EQ-5D in malignant mesothelioma. Health Qual Life Outcomes. 2015;13:6.

26. Versteegh MM, Leunis A, Uyl-de Groot CA, Stolk EA. Condition-specific preference-based measures: benefit or burden? Value Health. 2012;15(3):504–13.

27. Petrou S, Rivero-Arias O, Dakin H, et al. The MAPS reporting statement for studies mapping onto generic preference-based outcome measures: explanation and elaboration. Pharmacoeconomics. 2015;33(10):993–1011.

28. Versteegh M. Impact on the incremental cost-effectiveness ratio of using alternatives to EQ-5D in a Markov model for multiple sclerosis. Pharmacoeconomics. 2016;34(11):1133–44.

29. Barton GR, Sach TH, Jenkinson C, Avery AJ, Doherty M, Muir KR. Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores? Health Qual Life Outcomes. 2008;6:51.