PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Mapping the 12-Item Multiple Sclerosis Walking Scale to the EuroQol 5-Dimension Index Measure in North American Multiple Sclerosis Patients
AUTHORS	Coleman, Craig; Sidovar, Matt; Limone, Brendan; Lee, Soyon

VERSION 1 - REVIEW

REVIEWER	Olivia J. Phung, PharmD Assistant Professor Pharmacy Practice and Administration Western University of Health Sciences No conflicts of interest to disclose.
REVIEW RETURNED	06-Mar-2013

GENERAL COMMENTS	General: This is a well-conducted validation study which follows standard methods. This provides valuable information which will allow the use of MSWS-12 in future economic analyses. Introduction: The UK study was pointed out as having a small sample size; please include the number of patients in this sentence. Please specify how UK patients would be unique from North American patients and how this would decrease applicability of the existing study to the target population. Results: Please specify that Table 1 describes both the derivation and validation cohorts. Please comment on if there were any differences between the derivation and validation cohorts. Table 1: It may be useful to provide the descriptions of Items 1 to 12 somewhere in the paper. A footnote to this table may be an appropriate location.

REVIEWER	Annie Hawton Research Fellow in Health Economics Health Economics Group University of Exeter Medical School UK I have no competing interests.
REVIEW RETURNED	26-Mar-2013

THE STUDY	A little more information on the clinical characteristics of the participants would be useful e.g. the proportions with different sub-types of MS. Also, please can the authors provide some information on the representativeness of the 3,505 of the 34,000 people with MS on the registry, which may support the generalisability of the mapping algorithm to other individuals with MS.
GENERAL COMMENTS	Thank you for the opportunity to review this article. I found it clear, accessible, and to the point. I also found it very useful as it represents a development of research related to work I have previously conducted. The specific comments I have are fairly minor: - Those wishing to use a mapping algorithm to estimate EQ-5D scores from the MSWS-12 may not have individual item scores available, but may have total MSWS-12 scores. As the MAE for the 'total score' algorithm is similar to that of 'all individual item scores' (0.115, as compared to 0.111 in the validation sample), it would seem appropriate and useful to also provide the 'total score' algorithm. - Our (Hawton et al.) study was based on a prospective, longitudinal, cohort study (rather than registry data). - Was there any evidence of multi-collinearity and, if so, how was this addressed? - A proportionally small number of respondents (n=80) had severe mobility impairment according to PDDS. This is unsurprising considering that sampling targetted participants who had previously reported a PDDS of less than or equal to 7 (p.5, line 9). Might a different categorisation be more useful? - p.5 line 43. Health states are valued by the general population, rather than specifically patients. - p.7, line 33. Why is the correlation between the EQ-5D and the MSWS-12 reported as Spearman's rho (rather than Pearson's) when all the following statistics are parametric? - p.7, line 39. It would be helpful to make clear in the text that these are the statistics for the validation sample. - p.9, line 7 'used data from people with MS in the South West of England'.

REVIEWER	Myla Goldman Department of Neurology University of Virginia
REVIEW RETURNED	02-Apr-2013

THE STUDY	Authors need to explain in more detail why the EQ-5D is the right measure to address the question. Further, they need to do a better job setting the stage for why this is important work. They need to address the "who cares" issue in further detail. Reviewer recommends statistical review to comment on the

	appropriateness of the approach overall and interpretation of the results.
RESULTS & CONCLUSIONS	Authors need to address the failure of this approach on the ends of the spectrum. While the estimation works in some of the population - it does not appear to be within boundaries of an acceptable "good-of-fit" (0.0011 - 0.19) in the Mobility - PS Normal and EQ-5D < 0.50. What does this mean for the application of the approach. Further, authors should provide better understanding/discussion of the meaningfulness or application of these results. Reviewer recommends a hypothetical example to provide context. As manuscript is currently presented - the import and relevance is missing.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Olivia J. Phung, PharmD
Assistant Professor
Pharmacy Practice and Administration
Western University of Health Sciences
No conflicts of interest to disclose.

General: This is a well-conducted validation study which follows standard methods. This provides valuable information which will allow the use of MSWS-12 in future economic analyses.

Thank you.

Introduction:
The UK study was pointed out as having a small sample size; please include the number of patients in this sentence.

Done.

Please specify how UK patients would be unique from North American patients and how this would decrease applicability of the existing study to the target population.

Hawton et al. conducted such an analysis, mapping the MSWS-12 onto the generic, preference-based EuroQol 5-dimension (EQ-5D) health-utility measure. However, their analysis was conducted on a small sample of patients (n=560) with MS from a prospective, longitudinal, cohort study in the United Kingdom. Since patients in different geographic regions of the world may perceive the impact of walking impairment on HRQoL differently, and Hawton et al. utilized EQ-5D preference weights derived in a general United Kingdom and not a United States population; the external validity of their results to an MS population in the United States is unknown. This latter statement has been added to the introduction.

Results:
Please specify that Table 1 describes both the derivation and validation cohorts.

Done.

Please comment on if there were any differences between the derivation and validation cohorts.

There were not. This was expected due to the random splitting of the full cohort as described in the

methods. We now state this in the results section.

Table 1: It may be useful to provide the descriptions of Items 1 to 12 somewhere in the paper. A footnote to this table may be an appropriate location.

Added to the legend of Table 1 as suggested.

Reviewer: Annie Hawton
Research Fellow in Health Economics
Health Economics Group
University of Exeter Medical School
UK

I have no competing interests.

A little more information on the clinical characteristics of the participants would be useful e.g. the proportions with different sub-types of MS.

NARCOMS does not collect data on MS subtype (RRMS, primary and secondary progressive), so this cannot be provided. Sixty percent of the 3,505 patients were receiving a disease-modifying drug, which can serve somewhat as a surrogate for RRMS. This data has been added to the first paragraph of the results section.

Also, please can the authors provide some information on the representativeness of the 3,505 of the 34,000 people with MS on the registry, which may support the generalisability of the mapping algorithm to other individuals with MS.

We apologize for any confusion. In the history of the NARCOMS registry, over 34,000 unique patients have participated; however, only 9,899 patients responded to the spring 2010 update survey. Of these, 3,505 also responded to the supplemental survey sent out a few months later. We now describe this more completely in the results section and provide some basic demographics characteristics of the 9,899 person cohort for comparison. As can be seen, no major differences appear to exist between patients the cohort responding to the NARCOMS update and supplemental surveys. We also have added the fact that the NARCOMS population may not be representative of all North American MS patients a limitation to the discussion section.

Thank you for the opportunity to review this article. I found it clear, accessible, and to the point. I also found it very useful as it represents a development of research related to work I have previously conducted.

Thank you.

- Those wishing to use a mapping algorithm to estimate EQ-5D scores from the MSWS-12 may not have individual item scores available, but may have total MSWS-12 scores. As the MAE for the 'total score' algorithm is similar to that of 'all individual item scores' (0.115, as compared to 0.111 in the validation sample), it would seem appropriate and useful to also provide the 'total score' algorithm.
In acknowledgement that not all who wish to use a mapping algorithm will have access to individual item score data, the “total score” algorithm which did not perform as well as the preferred individual MSWS-12 item-score equation is now provided in the legend of Table 2. As it was not the best performing/preferred model, no additional validation of this model was conducted.

- Our (Hawton et al.) study was based on a prospective, longitudinal, cohort study (rather than registry data).

Text has been revised.

- Was there any evidence of multi-collinearity and, if so, how was this addressed?

Multicollinearity of the MSWS-12 items was said to be problematic when tolerance was <0.10 or the variance inflation factor was >10. Items demonstrating multicollinearity would have been excluded from the model. This has been added to the methods section. No multicollinearity was observed meeting the above-mentioned criteria was detected in the best fit model.

- A proportionally small number of respondents (n=80) had severe mobility impairment according to PDDS. This is unsurprising considering that sampling targetted participants who had previously reported a PDDS of less than or equal to 7 (p.5, line 9). Might a different categorisation be more useful?

The categorization chosen was not random, but rather was designed to best fit how the PDDS depicts mobility/walking impairment. According to the PDDS, a score of 1-2 would not be associated with mobility impairment, scores of 3-6 would be some walking impairment, and scores of 7 or more would depict being wheel chair bound or bedridden. We have concerns about arbitrarily stratifying the population. This is why we augmented the PDDS analysis, by also looking at the performance of the equation by disease severity using the NARCOMS mobility performance scale and the EQ-5D (in the same fashion as Hawton et al. and consistent with the NICE Decision Support Unit's recommendations). We believe the reporting of results stratified by these additional tools should overcome concern about the highest PDDS group having a small sample size; as this was not a problem with these other measures.

- p.5 line 43. Health states are valued by the general population, rather than specifically patients.

Corrected.

- p.7, line 33. Why is the correlation between the EQ-5D and the MSWS-12 reported as Spearman's rho (rather than Pearson's) when all the following statistics are parametric?

Now reported as Pearson's r as requested.

- p.7, line 39. It would be helpful to make clear in the text that these are the statistics for the validation sample.

Done.

- p.9, line 7 'used data from people with MS in the South West of England'.

Changed.

Reviewer: Myla Goldman
Department of Neurology
University of Virginia

Authors need to explain in more detail why the EQ-5D is the right measure to address the question.

Beyond the obvious reason for mapping to the EQ-5D – the only health utility measure used in the NARCOMS database this analysis was run in – the EQ-5D is also a "prominently" used generic,

preference-based health-utility measure. This justification has been added to the introduction.

Further, they need to do a better job setting the stage for why this is important work. They need to address the "who cares" issue in further detail.

Again, additional text has been added to the introduction to further explain the frequent use of the EQ-5D and the importance of an MSWS-12 to EQ-5D mapping equation to enable the conduction of health economic evaluations.

Reviewer recommends statistical review to comment on the appropriateness of the approach overall and interpretation of the results.

Authors need to address the failure of this approach on the ends of the spectrum. While the estimation works in some of the population - it does not appear to be within boundaries of an acceptable "good-of-fit" (0.0011 - 0.19) in the Mobility - PS Normal and EQ-5D < 0.50. What does this mean for the application of the approach.

As noted on page 9 of the discussion, the failure of this approach, particularly at the lower end of the disability spectrum (as measured by the observed EQ-5D or PDDS) suggests generalizing our model to severe MS patients (PDDS ≥ 7, NARCOMS mobility scale score ≥6) may be questionable.

Further, authors should provide better understanding/discussion of the meaningfulness or application of these results. Reviewer recommends a hypothetical example to provide context. As manuscript is currently presented - the import and relevance is missing.

We have added text to the introduction and discussion to better explain that the creation of such a mapping equation may enable to ability for cost-utility analyses to be conducted on MS interventions that do not have health utility data. As suggested, we have added a worked example of the mapping equation to the manuscript as Appendix 1.

VERSION 2 – REVIEW

REVIEWER	Annie Hawton Research Fellow in Health Economics University of Exeter Medical School University of Exeter Exeter UK
REVIEW RETURNED	23-Apr-2013

- The reviewer completed the checklist but made no further comments.