Author's response to reviews

Title: Longitudinal proxy measurements in multiple sclerosis: patient-proxy agreement on the impact of MS on daily life over a period of two years.

Authors:

Femke FAH van der Linden (fah.vdlinden@vumc.nl)
Jolijn JJ Kragt (jj.kragt@vumc.nl)
Margarethe M van Bon (msk700@student.vu.nl)
Martin M Klein (m.klein@vumc.nl)
Alan AJ Thompson (a.thompson@ion.ucl.ac.uk)
Henk HM van der Ploeg (vanderploeg-stapert@planet.nl)
Chris CH Polman (ch.polman@vumc.nl)
Bernard BMJ Uitdehaag (bmj.uitdehaag@vumc.nl)

Version: 2 Date: 13 January 2008

Author's response to reviews: see over
Dear Sir,

Please find attached a revised version of our manuscript entitled:
‘Longitudinal proxy measurements in multiple sclerosis: patient-proxy agreement on the impact of MS on daily life over a period of two years’ (MS: 1430757510156619)

First, we would like to express our gratitude for the fact that you gave us the opportunity to rewrite the initial manuscript. We appreciate the valuable comments of the reviewers, which we used to improve our manuscript. The comments of each reviewer will be addressed in the following pages of this letter.

We hope that the manuscript in its current version can be accepted for publication in BioMed Central Neurology.

With kind regards,

Yours sincerely,

Ms. F.A.H van der Linden
Reviewer 1: Michael Hutchinson

Major Compulsory Revisions:

1. We are told EDSS was estimate at both time points but that EDSS and cognition and HADS had no effect on the patient/proxy assessments. We are not told what happened to the EDSS over time other than the median change from 4.5 to 5.5. If one groups the patients by using change/no change in the EDSS as an anchor – using the physician’s assessment as a standard then what happens to proxy and patient measures?

2. There is the difficulty that neurological experts measured the EDSS at baseline but a medical student measured it at follow-up. The finding that stable patients reported worse MSIS-29 scores than changed patients is counter-intuitive—one might expect stable patients to show a response-shift phenomenon—but it is the deteriorated patients who are showing this. Response-shift should at least be discussed as an explanation of the reporting difference between proxy and patient.

The reviewer has two valid points. We acknowledge the difficulty of the different methods of EDSS assessment at baseline and follow-up. Home-visits were chosen at follow-up to keep the response rate of this already small number of patients as high as possible. Practical considerations obligated us to use a medical student for these home-visits. However, for further data analyses we decided to categorize the EDSS scores into three distinctive categories: 0.0 – 3.5 (no problems with ambulation), 4.0 – 6.0 (problems with ambulation up until the requirement of a cane) and 6.5 – 8.0 (the need of two canes or a wheelchair). We believe that by doing that the reliability of the scores are acceptable, because the boundaries between the categories are clearly noticeable for a non-expert such as a medical student. We therefore trust the use of the cross-sectionally assessed EDSS at baseline and follow-up in the data analyses concerning the possible influence of disability on patient-proxy agreement. However, we specifically chose not to use the EDSS as an anchor for the longitudinal data-analyses because of the two different methods of assessment at baseline and follow-up. We realize that we did not justify these considerations in the manuscript, but we hope this will answer the reviewer’s questions.

We agree with the reviewer that response shift could be of influence and we therefore included this in the discussion.

4. Some of the discussion drifts in the methods section which should be considerably shortened (page 7, lines 9-11)??. The results section in the abstract is too short; the background section too long. OVERALL EVERYTHING IS TOO LONG- GIVEN THE RELATIVELY MODEST FINDINGS AND THE POSSIBILITY OF TYPE 2 ERRORS WHICH THE AUTHORS ACKNOWLEDGE

The background in the abstract was shortened and the results section was extended. The discussion section was also shortened and we reduced the number of references. However, the additional information on response shift, as well as requirements of other reviewers, added additional text.
Reviewer 2: Alessandra Solari

Major points:

1. The study was performed as part of another one. The authors should give details about the ongoing study in order to understand how the study population was selected, instead of simply writing patients and proxies respondents were recruited from an ongoing study. Additional information about the inclusion criteria of the ongoing study was added to the method section. We hope this clarifies the way that patients were selected.

2. The Dutch version of the MSIS-29 is not documented in any peer-reviewed publication, but only by an abstract dating back to four years ago. We therefore have no peer-reviewed data documenting the properties of the Dutch MSIS-29 and in particular with it has similar properties to the original. Although the validity of the translation of the scale into several European languages, including Dutch, was thoroughly assessed in a working group lead by the designer of the scale and, we regret to say that this data has not (yet) been published as a full paper in a peer-reviewed journal.

3. Proxies were partners of MS patients. In the Discussion the authors acknowledge this can be a limitation. However we have very little information about these proxies except their sex and age. In my opinion more information is required, such as proportion of partners actually living with their patients, how long they have known the patient. Given that the some partners are also the care giver, psychological status is also highly relevant to the responses partners are likely to give on the patient MSIS-29. It seems that none of this information was collected and this is a serious weakness.

We agree with the reviewer that proxy characteristics can have an influence on patient-proxy agreement. The proxy respondents were all partners and had daily contact with the patient. All the partners lived together with the patients, although one couple was separated during the week since she patient stayed in a rehabilitation center. We added this information to the method section. Unfortunately, no additional information of the proxy respondent was included in the study. We will certainly keep this in mind for future studies.

4. Reliability should be assessed not only on the two subscores, but also on individual items of the MSIS-29. This because it is possible that a given subscore can be obtained as a result of differing scores on the individual items.

We agree that the subscores represent mean scores of the individual items and that patient-proxy agreement might differ between certain items. We choose to analyse our results in line with the standard use of the scale. Assessing the reliability of individual items requires a more detailed approach and a larger sample to have sufficient power allowing to correct for multiple testing. Also, this approach would mean adding additional analyses to the length of the manuscript. We would therefore like to keep our focus on the subscores.

5. Other limitations are the limited power of the study and uneven retesting.

We agree with the reviewer that there is limited power due too the small sample size which could lead to Type II errors, but this was acknowledged in the discussion section. We agree that the results could have been influenced by uneven retesting, we added this to the discussion section.

6. On Table 3 there seems to be a typing error (0.992??).

We are grateful for the precision of the reviewer and we adjusted the typing error.
Reviewer 3: Allyson Jones

Comments and Minor Essential Revisions:

1. Introduction: The introduction presents relevant information to support the significance of such a study.

2. Methods: The methods provide satisfactory detail. It is unclear, however, what relative timeframe that baseline represented. Was baseline relative to when patients were first seen at the center? How were patients recruited?
   
   Baseline data was collected over a period of 4 months. Baseline was not relative to when patients were first seen at the center. We would also like to refer to point 1 of reviewer 2.

3. Results: Although the authors state no differences were seen between subgroups of the HADS or BRB-N scores, it was unclear whether patients represented a spectrum of severity of cognitive deficits and/or emotional states. Education levels of the patient and proxy are other factors that may affect the level of agreement and should be reported.
   
   We believe that Table 1 provides data on the spectrum of severity of cognitive deficits and/or emotional states. The number of patients in each of the subgroups of the BRB-N and the HADS are shown. Most of the patients were not suspected of depression (n=42) and anxiety (n=37) and most of the patients did not have cognitive problems (n=46).
   
   Unfortunately no additional data was collected on educational level of the patients and proxies. We will certainly keep this in mind for future studies.

4. Discussion: The Discussion provides a good summary of the results; however, further emphasis is needed on the comparison findings to previous research, particularly with possible influencing factors. Further discussion is needed regarding the practical application of the information.
   
   We agree with the reviewer that further emphasis on the comparison findings to previous research was needed and we incorporated this in the discussion section. However, we believe that conclusions on the practical applications based on our results might be premature in light of the limitations of the study, including the size and composition of the sample.

5. This is the first study of MS patients to examine agreement longitudinally. This is an important observation to researchers who wish to use proxy respondents when measuring quality of life in MS patients who might not be able to be interviewed. Overall, there are grammatical errors throughout the text which impede the flow of the reading. The study strengths include a carefully conducted analytic strategy, the inclusion of validated measures, and the prospective design. The limitations consist of exclusion of significantly cognitively impaired patients which limit the results to the patients least likely to need proxy reporting.
   
   We could not agree more on the fact that significantly cognitively impaired subjects are the sample in which proxy respondents might be most useful. The reason for not (yet) including that specific sample was that we first wanted to get more insight into longitudinal proxy measurements in a sample who were not significantly impaired. This gave us insight into several methodological problems concerning proxy measurements, which can be used to improve studies in samples with cognitively impaired patients. We are currently working on patient-proxy agreement in a sample with cognitively impaired patients. This will be the main focus in our next paper.