Reviewer's report

Title: The size of the treatment effect: do patients and proxies agree?

Version: 2 Date: 31 July 2008

Reviewer: Alessandra Solari

Reviewer's report:

The manuscript addresses the agreement between MS patients and proxies on changes in the 29-item MS impact scale (MSIS-29) after a cycle of i.v. steroids. Patient self-reported rating scales are increasingly used as outcome measures in clinical trials on MS, and determining the responsiveness of these instruments is fundamental. It is also important to consider surrogate responders in cases where patients assessment is not reliable, usually because of cognitive compromise, provided that agreement between patient and proxy responses is acceptable. Thus, the topic of this manuscript is of interest.

However the manuscript has the following major limitations:

1. The study is of limited size (55 patient-proxy couples) and the population is poorly delineated. It is not at all clear whether these are preliminary results of an ongoing study and, if so, it is not specified the target number of patients to be enrolled, or how this number was determined. No information is given regarding the recruitment period.

2. Attrition was high (24%), with analyses performed on only 42 patient-proxy couples. This high proportion of missing subjects seriously undermines the internal validity of the results.

3. The authors compared measures of internal responsiveness (“effect size”) and external responsiveness (receiver operator characteristic – ROC – analyses) in people with MS and their proxies. This comparison is useful but only complimentary to the stated objective of the study: not to assess responsiveness of the MSIS-29, but rather the agreement between MS patients and proxies respondents on changes in the 29-item MS impact scale (MSIS-29). It is thus vital to assess inter-rater agreement on physical and psychological score changes, by estimating for example the intraclass correlation coefficient.

Minor points:

1. The abstract is uninformative: no figure is given regarding number or participant characteristics; numerical data are totally absent, even from the results section.

2. In the results section, the descriptive data are insufficient: for the EDSS only median values are given, without any indication of dispersion (range or interquartile range).

3. A table reporting the number of patients and proxies scoring minimum and
maximum; and the number of missing items at the two MSIS-29 scales is necessary.

4. The fact that only those with more than 50% missing items on both MSIS-29 scales were excluded from analyses (page 8) is a questionable choice, and should be justified with documentation.

5. Finally, ROC curves should be provided, and the area under the curves reported together with their confidence intervals (or statistical significance). This is important in consideration of the limited numbers and (as stated by the authors) variability of the findings.

In consideration of the three major drawbacks of the study (lack of the principal analysis, limited sample size and high attrition), my position is that the manuscript unacceptable for publication in BMC Neurology.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.