Reviewer’s report

Title: A critical analysis of test-retest reliability in instrument validation studies of cancer patients under palliative care: a systematic review

Version: 2 Date: 21 November 2013

Reviewer: Camilla Zimmermann

Reviewer’s report:

This is an interesting and unique analysis of test-retest reliability studies of instruments developed for palliative oncology care. It is an important study, given the increasing number of validation studies in this area, and the necessity of establishing rigorous standards for the conduct of such studies. I have the following suggestions for the improvement of this manuscript.

Major

1. Under search strategy, on page 6, it is not clear which discrepancies were rectified by referring to the original articles. Are you referring to differences in opinion regarding which articles should be reviewed?

2. Under Exclusion criteria, page 7, “the study was a validation of a needs instruments” – incorrect grammar and not clear what a needs instrument is. Please clarify. As well, it is redundant to list all the various types of studies that were excluded, which did not fall into the inclusion criteria. In other words, if inclusion criteria are “validation study of a multidimensional quality of life instrument or a multidimensional symptom assessment instrument”, then it is obvious that studies were excluded if that were not validation studies, focused on only one symptom, or evaluated a quality of care instrument. Likewise, if the inclusion criterion is a population composed mainly of advanced cancer patients, then there does not need to be an exclusion criterion for “not composed mainly of cancer patients”. Therefore please edit the “exclusion criteria” section accordingly.

3. Under “scores of retest” (page 11), there should not should be an explanation in the results for why the trend for the retest results favoring shorter time intervals were nonsignificant (ie due to small numbers and lack of power). This is an interpretation, and should be mentioned in the discussion section (as it already is). Should just be mentioned in the results that these were nonsignificant trends, without offering an explanation here. As well, in the discussion it should be mentioned that another reason contributing to the nonsignificant results might be the large interquartile range for some of the domains.

Minor

1. Under exclusion criteria on page 7, there seems to be a number missing (number 3, before “the population studied was not composed mainly of advanced cancer patients…”)
2. It is mentioned on page 14 that “we believe that symptoms are more prone to fluctuations than overall HRQoL scores”. I agree with this statement, but it would carry more weight if it was rephrased as less tentative and followed by a reference. At the very least, a reason for this belief should be given.

3. Is there a time interval that you would suggest for test-retest analysis for symptom and quality of life measures, respectively, for palliative populations? This would be useful for future research.

4. Table 4: Suggest renaming the table so that it better describes the results. Also, suggest moving the heading “median (Q1-Q3)” to right underneath “Time interval (hours)”, because the median

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

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests