Reviewer’s report

Title: One-minute stair climbing, 50-foot walk, and timed up-and-go were responsive measures for patients with chronic low back pain undergoing lumbar fusion surgery

Version: 0 Date: 09 Nov 2018

Reviewer: Lina Holm Ingelsrud

Reviewer's report:

The aim of this study was to evaluate the responsiveness and define the MIC for four well established physical capacity tasks to evaluate outcome in patients undergoing lumbar fusion surgery for motion-elicited chronic LBP with degenerative changes. The study is overall a well performed and high-quality clinimetric study, using contemporary described (COSMIN) methodology for the responsiveness analyses. I have some minor suggestions/concerns, primarily regarding the MIC analyses, for the authors to consider.

Overall, since the MIC may vary depending on the direction of change, I suggest being clear throughout the manuscript that in this study, the MIC for improvement was evaluated.

Background:
Line 111-113. Quote: "(…) other patient groups in the Swedish Spine Registry (...)"). It is not perfectly clear to me what the "other patient groups" refer to. Do you mean that patients with lumbar spinal stenosis and disc herniation have superior outcome to patients with chronic LBP due to degenerative disc disease? If so, I would specify that.

Line 131. Quote: "Physical capacity tasks also have clear face validity (...)"). I disagree that physical capacity tasks always have clear face validity for all patient groups and settings. Suggest moderating the statement.

Methods
Table: 1, hypothesis 1: It is not clear which levels on the GPE were used to define patients with and without meaningful change.

Lines 287-289: I just have a little wording suggestion about the ROC curve interpretation. Quote: "(…) with 1 indicating perfect responsiveness". I would defer from stating that the area under the ROC curve value of 1 in itself serve as proof of perfect responsiveness. When the value is 1, there is perfect discrimination between the unchanged and the improved patients. The proof of responsiveness is based on the hypothesis testing method, as presented in this study.

It would be good as a reader to have access to the actual wording of the construct-specific GPE
There is no consensus on the minimal number of observations needed to perform robust MIC analyses. For the MIC values, it would be beneficial to calculate confidence intervals, which may be calculated using bootstrapping. The number of observations in this study is limited (n=57), why the confidence intervals around the MIC values may be large.

Results
Table 4: Could the table 4 also include the distribution of change in physical capacity task scores for the total 57 patients also? The purpose would be to put the MIC values into clinical context (what is the change that is "normal" to find in these physical function scores, and which degree of change is needed to perceive the change as clinically relevant)? Furthermore, it would be relevant to include in this table the correlation between the GPE and the Physical function change scores, to serve as further justification of the validity of the GPE scores for use as anchors in MIC analyses (reference: Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61(2):102-9).

Discussion
Lines 422 to 426. It is not clear to me what the difference in study populations is. Which populations were studied by Andersson et al. (51) and Strand et al., 2011 (52)?

Lines 426 to 428: Did the authors Andersson et al. and Strand et al., 2011 also use the optimal hypothesis testing procedure to evaluate responsiveness? If they rather based their conclusion on (probably more frequently used) solely effect size parameters, which also may explain why their conclusion was different.

Lines 472-481. Clinical implications. It seems the focus of the study is to establish the usefulness of these physical function measures in a clinical setting for individual patients. It is however important to acknowledge that the MIC is a group-based statistic, and the MIC value is therefore of limited value for individual patients in the clinic. From my point of view the MIC has is more appropriate to evaluate group-based change, as either 1) the proportion of "responders" to treatment, where patients with change scores larger than the MIC values are classified as responders, or 2) interpreting the mean change improvement for a group of patients. For individual patients in the clinic, however, that individual patient is the only expert on whether the improvement he/she experience is of importance to him/her. The MIC values may in this context serve as a reference for what the average patient finds important, to be used in the shared-decision making process when evaluating the treatment outcome.

Another point that is worthy to mention in the discussion is that, even within anchor-based methods, MIC values vary markedly based on the statistical methodology used. A study from 2015 suggested that ROC estimates are less precise than using a predictive modeling method and also biased when the proportions of improved patients is higher than 50% (which is in fact the case in this present study) (references: 1) Terluin B, Eekhout I, Terwee CB, de Vet HC. Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. J Clin Epidemiol. 2015;68:1388-96., and 2) Terluin B, Eekhout I, Terwee CB. The anchor-based minimal important change, based on receiver operating characteristic analysis or predictive modeling, may need to be adjusted for the proportion of improved patients. J Clin Epidemiol. 2017;83:90-100). It would be relevant for the authors to acknowledge, in the discussion, that the optimal MIC methodology is debated in the literature, and no consensus exists on the best Mic methodology. Consequently, the MIC
values may prove different if other methodology is applied in the future.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

Yes

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or
has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal

Do you want to get recognition for reviewing this manuscript? Add a record of this review to Publons to track and showcase your reviewing expertise across the world’s journals. Signing up is quick, easy and free!

Yes