Reviewer’s report

Title: Elaborating on the Assessment of the Risk of Bias in Prognostic Studies in Pain Rehabilitation using QUIPS – Aspects of Interrater Agreement

Version: 0 Date: 05 Nov 2018

Reviewer: Jill Hayden

Reviewer's report:

Thanks for the opportunity to review this manuscript. The paper presents a welcome assessment of interrater reliability and critique of the QUIPS tool. Overall the paper is well written and clear. I have comments and suggestions to clarify study methods and guidance recommendations provided, as well as some general suggestions for more consistent language and descriptions for prognosis systematic reviews.

1. The terms 'prognostic factors' and 'predictors' are each used. It is unclear if these are intended to have different meanings (e.g. factors associated (independently?) with outcome, predictive models to identify risk groups, or factors acting as treatment effect modifiers) - since language is still so inconsistent in the prognosis field, it warrants early definition of terms followed by consistent use.

2. It appears that the systematic review is a broad review of all prognostic factors/several outcomes. This should be clearly described in the methods, including clear description of how QUIPS was applied, and challenges in this context. Was QUIPS operationalized and assessed separately for each prognostic factor/outcome under consideration? For example, the judgement for assessment of confounding may differ for various prognostic factor/outcome associations; the risk of bias related to prognostic factor measurement will vary considering different prognostic factors. How was this considered? If not, how might this have impacted your results and conclusions? Challenges and recommendations for future reviews?

3. Language related to 'quality' vs 'ROB' seem to be used interchangeably (page 4), but should make distinction for clarity (e.g. page 12 bottom; page 9, line 1). It will likely add clarity to reserve use of 'quality' language for the broader GRADE assessment of the overall quality of evidence available, with 'risk of bias' used to describe study validity and impact on potential bias due to study methods.

4. Description of the Hayden 2013 study should be clarified. This study reported results from various published systematic reviews that had used the QUIPS tool in various ways. I suggest reporting methods (page 5, line 21) as '…kappa values evaluating QUIPS WERE
REPORTED TO VARY from…’. On page 10 this should be clarified ('raters' were not selected, they were simply systematic review authors who had used the QUIPS tool, with additional detail from those who responded to a survey). It seems that it would be most relevant in this study to limit comparison of your results to those systematic reviews that were most comparable (i.e. broad prognostic factor reviews that assessed by 6 domains as recommended).

5. I think that a particularly useful contribution to the literature is a more detailed specific discussion of issues relevant to the chronic pain rehab field. Since the recommendation is that QUIPS should be operationalized for the review question, I would like to see more elaboration and examples of these specific issues; how you resolved and what are your recommendations for full operationalization relevant to this chronic pain rehab field.

6. Clarify - were the 5 studies used for pilot testing from the same review? Were they included in the interrater reliability testing?


8. Page 7, line 16. 'All ROB domains in all STUDIES (n=43) were separately judged by the raters as having low, MODERATE, or high ROB'. (correct wording)

9. Be clear and purposeful about use of 'papers' or 'studies'. Were 43 studies assessed, or 43 papers (i.e. these may be different if there are multiple publications of the same study). Since this is common in prognostic factor reviews, it will be good to be consistent with the language used.

10. There are a few references to differences in assessment between studies with high and low ROB (e.g. Page 8, line 13). How was this assessed? How were high and low ROB studies defined? Was this before consensus or after? Overall or by domain? This should be described clearly in the methods.

11. Page 8, lines 17-22 describe number of domains rated similarly. It seems that it would be more useful to describe which specific domains tended to be more challenging/less agreement, rather than the number of domains.

12. The distinction between Round 1 and Round 2 in tables and text of results is of unclear value. If there is an important distinction in the methods between Round 1 and Round 2 (e.g. were the QUIPS items operationalized differently or more fully for Round 2? Were different reviewers involved? Were the studies newer?), any different methods should be clearly
described in the methods section, then implications discussed. If not, I suggest dropping these subgroups from results tables.

13. It is useful to point out where there were particular challenges with application of the QUIPS tool. The manuscript will be further strengthened by describing what your solution was. What was the additional operationalization that you included - did this make a difference in the Round 2 assessments? Table 3 is helpful for further operationalization of the QUIPS items. Did you also further operationalize the tool for your domain judgments? This will be helpful to add to the elaboration specifically for the chronic pain rehab field.

14. Page 14, lines 15-20. It is a useful example to describe a potential a cut-point to help inform judgements about risk of bias due to participation and attrition. However, related discussion should include caution that while it may be useful as a guide, reviewers still need to consider this number in the context of other study information (e.g. a study may have less than 33% loss to follow-up, but if this loss was systematic based on a characteristic that could confound the association between the prognostic factor and outcome of interest, then the study still could be at high risk of bias).

Minor editing suggestions:

15. Abstract, line 6. I would not call QUIPS a 'new tool' since it has been around since 2006.

16. Abstract, line 10. It seems that goal is to elaborate on the OPERATIONALIZATION of the instrument (or use of the instrument?).

17. Abstract, line 13. We performed a SYSTEMATIC REVIEW and meta-analysis…

18. Page 4, line 13. '…whether a potential prognostic factor is associated with an outcome and secondly to ESTIMATE the strength…'

19. Page 5, line 5. Cochrane Prognosis Methods Group (drop 'The' and 'Collaboration')

20. Page 9, line 11-12. 'Maximal disagreement was not caused by one of the raters judging CONSISTENTLY lower or higher…'.

21. Page 12, line 11. 'Most of the studies included in the present study were classified as having multiple DOMAINS WITH HIGH risk of bias'.

22. Page 13, line 12. 'modification/clarification of the prompting items for each specific REVIEW QUESTION'. 
23. Page 13, line 22. 'discussed before performing ROB ASSESSMENT IN a systematic review'.

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

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.

Non-financial competing interests in relation to this paper: Author of the QUIPS tool that is critiqued in this paper :)

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