**Reviewer’s report**

**Title:** Systematic reviews and meta-analyses addressing comparative test accuracy questions

**Version:** 0  **Date:** 15 Apr 2018

**Reviewer:** Jon Deeks

**Reviewer’s report:**

Systematic reviews and meta-analyses addressing comparative test accuracy questions

I strongly agree with the sentiment of this paper, and believe that it is important to ensure that the profile of comparative test accuracy studies is increased. I include some suggestions below for the authors to consider.

1. This paper does not lend itself to the structured abstract format included as there are no methods or results in this commentary. However there is a more appropriate abstract included on the pdf so I am not certain whether this has been changed already.

2. The clarity of the paper might be helped by including some real examples. For instance, how does a clinician decide whether to use Ultrasound or CT to diagnose appendicitis? The importance of comparative evidence is obvious when you focus on a management decision which involves making a choice between two technologies.

3. The authors focus on the importance of primary comparative accuracy studies, but not so much on ensuring systematic reviews address comparative accuracy when it is appropriate. Many reviews also only evaluate an individual technology, which compounds the problem. It is often hard to convince review authors that they should evaluate more than one technology in a review due to the magnitude of the work involved. The authors could give greater profile to the importance of systematic reviews being comparative as well as the primary studies.

4. The authors state that primary comparative studies of diagnostic tests are rare (abstract). However, they made up around one third of the included studies in the comparative Cochrane Reviews, and I recall that Yemisi Takwoingi reported a similar proportion in her larger review of all DTA systematic reviews (ref 1). Even then these figures will be underestimates as some of the apparent "1 test" studies will actually have compared with one or other test, but not include the comparison of interest. Thus I would ask the authors to reconsider the language that they use here. I think it is important to be aware that there are many comparative studies which exist which can be used to include in systematic reviews.

5. There is often confusion between a comparator test and the reference standard in DTA discussions. It might be helpful for the authors to highlight the difference. Also, there
are occasions where the required comparator test is the reference standard which also deserves comment.

6. Extending the questions to prognostic models seems appropriate. However, I find the language here confusing. The authors talk about finding out whether "a predictor has incremental value". I am confused as to whether the authors are talking about a single predictor variable (which often is investigated during the model development process), or comparing two different predictive models (which is less likely to be considered). I would expect their intent to be the latter but this is not clear. Again it would be good to have a real comparison to consider e.g. comparing QRISK2 with QRISK3 or some other real health problem. Citing a review or study which had done this would be useful.

7. The paper concludes by talking about clinical utility comparisons. My observation is that there is not the same issue when researchers start talking about clinical utility - they naturally think about building comparisons into the study, whether they are looking at changes in diagnostic thinking, treatment decisions, patient outcomes or cost effectiveness.

8. The paper would benefit from a grammatical check.

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

An article of importance in its field

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

Needs some language corrections before being published

**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 have been working and continue to work on related research questions. Individuals in my team are funded to work on projects on comparative test accuracy.

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