Reviewer’s report

Title: The Brier score does not evaluate the clinical utility of diagnostic tests or prediction models

Version: 1 Date: 19 Sep 2017

Reviewer: Werner Vach

Reviewer’s report:

Unfortunately, the authors did not address adequately the points raised by the reviewers, and the paper has not improved.

There are still several major issues with the paper.

1) The upper part of Table 1 mainly illustrates mainly that the overall classification rate is a poor idea to describe the clinical utility of a diagnostic test, as it weights FP and FN decisions equally. Every measure, which allows to reflect our knowledge about the clinical consequences of FP and FN decisions, is preferable. It is hence in no way surprising, that net benefit behaves better, as the net benefit allows this.

2) The lower part of Table 2 is rather meaningless. According to the authors, there is no functional or logical relation between the test evaluated by the net benefit and the prognostic model evaluated by the Brier score. So why should this comparison make sense?

3) The authors try to damn the Brier score overall. However, the Brier score is not designed as a measure of the clinical utility of a binary test. So the negative conclusions drawn by the authors are just unjustified. The question, whether the Brier score is useful to measure the clinical utility of a prognostic model would require a framework, where the model is not reduced to a binary test. For example, when a patient is interested in knowing his or her individual risk in order to make an informed decision. Hence the paper cannot make any statement about whether the Brier score is useful to evaluate the clinical utility of a prognostic model.

4) The authors still present no reference for a paper, where the Brier score is explicitly used to evaluate a binary test.

5) The formulas for E[BS] on page 6 are incorrect.

As pointed out in my first review, there may be a useful point in this paper in the sense that researchers should be aware of that a good prognostic model does not necessarily define a diagnostic test with good clinical utility, and a poor prognostic model may define a test with good clinical utility - at least if we use the Brier score for assessment of the quality of the
prognostic model. Hence whenever a test is derived from a prognostic model, the value of the test has to be evaluated in addition of evaluating the prognostic model.

**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.

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