Reviewer’s report

Title: Quantile-based Fecal Hemoglobin Concentration for Assessing Colorectal Neoplasms with 1,263,717 Taiwanese Screenees

Version: 0 Date: 29 Oct 2018

Reviewer: Iris Lansdorp-Vogelaar

Reviewer's report:

The authors have assessed the association between fHb concentration and detection of colorectal neoplasia using a novel methodology that they claim avoids some of the bias associated with standard approaches. Although I see merit in their work, I have some major concerns with the manuscript in its current form:

1. The authors description of the methods and the work is not intuitive. Based on the description I would expect survival curves. Yet Figure 3 portrays cumulative percentage curves. It is not until reading the discussion that it becomes clear that indeed fHb concentration is considered as time. This intuitive explanation (bottom of page 20) should be added to the methods section.

2. I don't see the advantage of using this method over previous methods. The authors claim that rank-based methodology is needed because fHb concentration data are skewed. Is ranking ordinal fHb the only solution to the problem? The discussion says that skewedness could be easily solved using log-transformation. Can we not just use the regular risk-prediction techniques on log-transformed data that are more intuitive?

3. Why are tail end outcomes and skewedness even a problem for unbiased estimation? How would this affect your outcomes? This should be explained more clearly in the introduction to better set the stage for the authors work.

4. The methods are not described clearly enough. A lot of necessary information is lacking, such as definitions of who are considered neoplasm free, description of the Bayesian inversion method, how people are followed and treated in the analysis: how are people without neoplasia detected in round 1 treated if they develop neoplasia later in time. There is 1-2 sentences somewhere, but the information is too concise to get the full picture. On the other hand, the conceptual framework in the methods is very obvious and in my eyes largely redundant. It takes up a lot of space. I suggest to remove most of this framework to allow for more detailed explanation of the method plus its rationale. Some of these intuitive findings could actually be used to start of the results section.
5. The authors performed this analysis to address bias from skewedness of the data. However, I am concerned that this method is even more prone to detection bias than the standard methodologies. All people with low fHb concentrations are classified in the "no neoplasm" group. Yet many of these individuals may actually have neoplasms. It may very well be that if these people were to be included in the correct group, that the differences in fHb concentrations between the groups would be much smaller, theoretically even non-existent. The authors can try to assess this bias by only looking at individuals that underwent colonoscopy and evaluate if they see the same rank-ordering of fHb concentrations.

6. Page 12, line 49: states that interval cancers are censored. However, subsequent lines seem to suggest that missing values are imputed and so these cases are not censored, but included in the analysis. I have concerns about the imputation method: I strongly suggest to also include actual measured fHb in the imputation, because to me that seems most reliable information you have about concentration value.

7. Abstract: from reading the abstract, I was very unclear about the analysis performed. Also, the word rank is never used in methods, but is included in conclusion and title. Unclear what is meant by rank-based. Suggest to reformulate or include also in methods and results so that it is clear what is meant.

8. Table 3: risks at 90th percentile add up to >100%. Does that make sense? Can the authors explain?

Minor comments:

1. Figure 3: does this figure not sort of represent a sensitivity/specificity estimate? To me it is more intuitive to describe it that way.

2. Line 24: What does 83% reflect. Suggest to delete it or fully explain it.


4. Suggest to drop the non-parametric model, since it does not add anything to the analysis.

5. From the manuscript fHb concentration for different percentiles of normal individuals seem to be lacking. Why?
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

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

Unable to assess

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 recommend additional statistical review

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 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.