Author’s response to reviews

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

Authors:

Szu-Min Peng (d05849012@ntu.edu.tw)
Han-Mo Chiu (hanmochiu@ntu.edu.tw)
Hsiao-Hsuan Jen (d05849010@ntu.edu.tw)
Cheng-Yang Hsu (bacilli65@gmail.com)
Sam Li-Sheng Chen (samchen@tmu.edu.tw)
Sherry Yueh-Hsia Chiu (sherrychiu@mail.cgu.edu.tw)
Amy Ming-Fang Yen (amyyen@tmu.edu.tw)
Jean Ching-Yuan Fann (jeanfann@email.knu.edu.tw)
Yi-Chia Lee (yichialee@ntu.edu.tw)
Hsiu-Hsi Chen (chenlin@ntu.edu.tw)

Version: 2 Date: 01 Feb 2019

Author’s response to reviews:

Reply to Reviewer 2 (Iris Lansdorp-Vogelaar, PhD):

1. (Q) Add the importance of using quantile-based fHb to the abstract. Currently the reader cannot appreciate why it would be important to study the statistical properties of this method. Suggest to add that current studies on fHb association with risk of colorectal neoplasm are hampered by skewedness of the data.

   (A) This sentence “current studies on this subject are hampered by skewedness of the data and the ordinal property of f-Hb” has been added to the background of abstract to address the importance of using quantile-based fHb method. (see Abstract)

2. (Q) The results of the abstract do not align well with the methods: the Results start with the term "adjusted", however nowhere do the methods mention anything about adjusting. The
results of the abstract mention fHb for the normal group, however no value is provided for this group. Suggest to add value of 0 there.

(A) The term “adjusted” together with the use of “accelerated failure time multivariate analysis” has been specified in the method section of abstract. The value of 0 μg/g has been delineated for the normal group. (see Abstract)

3. (Q) I find the sentence with colorectal neoplasia risk hard to follow: do these numbers reflect the risks at their respective median values. i.e. risk of 4.0% of non-advanced adenoma at 57; risk of 4.8% of AA at 82.4 and risk of 29.5% for CRC at 163.1? I suggest to present risks at one fixed value for all three types and then may be present different types. Would be easier to follow for the reader. For example, similar to how you have presented the results in appendix table 5.

(A) This has been amended as suggested by the reviewer in the abstract (Page 4, Line 3 from the bottom) and the result section (Page 19, Line 3 from the bottom) with the following statement “At 90 ug/g of f-Hb, the highly suspected cut-off for colorectal disease, the risks were 17% for non-advanced adenoma, 6% for advanced adenoma, and 9% for CRC.”.

4. (Q) Conclusion of the abstract is not really a conclusion, but rather a method description.

(A) The conclusion has been re-written to conclude the risk assessment for the risk of multistage of colorectal neoplasia using the quantiles of f-Hb with the flowing sentence “Covariate-adjusted risk stratification for multistage outcomes of colorectal neoplasia were provided by using the quantiles of fecal hemoglobin concentration, yielding the estimated life-time risks of 25th to 75th quantiles, ranging from 0.5% to 44% for colorectal cancer, 0.2% to 46% for non-advanced adenoma, and 0.1% to 20% for advanced adenoma.” (Page 5, Line 2)

5. (Q) The introduction is very long and some methods are intertwined.

(A) The introduction has been abridged by pruning unnecessary intertwined sentence as suggested.

6. (Q) I find the inclusion and exclusion of cases at baseline extremely confusing. Do I understand it correctly that for the fHb AFT model you include cases at baseline, but for predicting future risk with the Bayesian inversion method you exclude them?
(A) The simple sentence has been provided for clarification as follows “That means quantile-based f-Hb AFT model included colorectal neoplasia at baseline, but for predicting future risk with the Bayesian inversion method these cases at baseline were excluded.” (Page 10, Line 2)

7. (Q) Methods, page 10, final two paragraphs contain results.

(A) These have been moved to the results section (Page 15, Line 3 from the bottom, and Page 16, Line 9).

8. (Q) I continue to have difficulty with imputing fHb values of screendetected cancers for interval cancers. Biologically interval cancers could be non-bleeding cancers that are therefore missed at screening. It is agreed that there are some cancers that never bleed. If you would do a FIT on these cancers just before clinical diagnosis you would still get a zero FIT result. Imputing values of screendetected cancers in that case, in my eyes would bias the analysis.

(A) The definition of interval cancer in this article is to follow the pathway of adenoma-carcinoma leading to the bleeding phenotype of interval cancer. Given this premise, the value of FIT just prior to the occurrence of interval cancer resulting from bleeding phenotype is very unlikely to have a zero FIT value as mentioned by the reviewer. Those interval cancers may be missed in the previous screen due to the undetectable bleeding phenotype stage. Here, we assume these undetectable bleeding phenotype interval cancers would grow up to detectable bleeding phenotype as they are detected in the screen. This form the basis for imputation of interval cancer with those accrued from screens. Regarding non-bleeding phenotype of interval cancer, it is true that there are some colorectal cancer with De novo pathway with flat and possible non-bleeding pathway. They may not lead to the elevated f-Hb. However, they are also not detected by screening with FIT. The percentage of De Novo pathway is rare and may not substantially affect the results. The assumption of bleeding phenotype and the limitation of the failure of considering de novo has been addressed by one of limitations. (Page 24, Line 7)

9. (Q) Suggest to omit the univariable analysis. The manuscript is already quite long, and the univariable analysis does not add much.

(A) The descriptions on univariate analysis have been deleted from the results section.

10. (Q) I don't agree with the result that median fHb is so much higher for non-advanced adenomas than for normals (Table 1). Previous studies have shown that sensitivity of FIT for
non-advanced adenomas is very low and that most of them will be found by chance rather than by the test. The fact that the median f-Hb concentration for non-advanced adenomas is so much higher than for normals is in my opinion due to detection bias. Non-advanced adenomas with f-Hb concentrations below 20, would be misclassified as normals. I understand the reply of the authors that they also included interval cancers and findings at later exams, but we all now that many people with adenomas (40% of the population!) would never develop symptoms or have their adenomas found. So all these people would be classified in your analysis as normal, while in fact they actually have non-advanced adenomas. This limitation of the study should be addressed in the discussion.

(A) This has been addressed as one of limitations at the end of discussion with the following statement “The second limitation is that as previous studies have shown that sensitivity of FIT for non-advanced adenomas is low the median f-Hb concentration for non-advanced adenomas is far higher than that for the normal group may be due to detection bias. That means that non-advanced adenomas with f-Hb concentrations below 20 would be misclassified as the normal subjects. The quantile-based method making allowance for such a misclassification is the subject of ongoing research.” (Page 24, Line 6 from the bottom)