Author's response to reviews

Title: Evaluating screening approaches for hepatocellular carcinoma in a cohort of HCV related cirrhosis patients from the Veteran's Affairs Health Care System

Authors:
Nabihah Tayob (ntayob@mdanderson.org)
Peter Richardson (peterr@bcm.edu)
Donna White (dwhite1@bcm.edu)
Xiaoying Yu (xiyu@utmb.edu)
Jessica Davila (jdavila@bcm.edu)
Fasiha Kanwal (kanwal@bcm.edu)
Ziding Feng (zfeng3@mdanderson.org)
Hashem El-Serag (hasheme@bcm.edu)

Version: 1 Date: 14 Nov 2017

Author’s response to reviews:
We thank the reviewer for their valuable comments that have improved our manuscript. The point-by-point response to their comments has been included here. In addition, the authors' response letter has also been included within the Cover letter (attached as a supplementary file).

Massimiliano Orso (Reviewer 1): This paper is aimed to evaluate two biomarker screening approaches, a six-month risk prediction model and a parametric empirical Bayes (PEB) algorithm, in terms of their ability to improve the likelihood of early detection of hepatocellular carcinoma (HCC) compared to current alpha-fetoprotein (AFP) alone when applied prospectively in a future study.

The paper is well planned and written; I have just some comments to do.

Major comments:
- The cohort of patients was randomly split into the training and validation cohorts. I wonder if this is a robust approach. Should be a cross-validation approach a better option in this case?
Response: Thank you for raising this important question. We have chosen a split sample approach for our primary analyses because we have very large cohort and sufficient sample size (especially HCC cases) to create strict separation of the training and validation cohorts. However we acknowledge that there is disagreement over which approach is most appropriate, and so we have now included in the revised paper an out-of-bag bootstrap validation approach as a sensitivity analyses in Table B in Appendix B. The following discussion of these results has been included on page 11-12 of the main manuscript.

“We chose a split-sample approach with training and validation cohorts to evaluate our HCC screening algorithms since we have a large cohort with 902 HCC cases and 11,222 controls. In a sensitivity analysis, we utilized an out-of-bag bootstrap validation approach where each bootstrap training cohort consisted of 12,124 patients drawn with replacement from the full analysis cohort and each bootstrap validation cohort consisted of all the patients not included in the bootstrap training cohort. The model parameters for each of the HCC screening algorithms were estimated using the training cohort, the screening algorithms were implemented in the validation cohort and the patient-level TPR at 10% screening-level FPR was estimated. This procedure was repeated 300 times and the results were averaged over the bootstrap iterations. In Table B in Appendix B, we observe that the results are mostly consistent; both the laboratory-based algorithm and the PEB approach showing improved TPR over the standard thresholding approach with AFP only across all the definitions of true positive screenings in HCC cases except one. For definition A2 with a restrictive time frame (only positive screens within 3-6 months prior to HCC diagnosis are true positives) and fewer HCC cases, the PEB algorithm and AFP only algorithm are approximately equivalent. In the two-years prior to HCC diagnosis (definition C1), the TPR of the PEB algorithm was 5.03% greater than the standard thresholding approach with AFP only (61.26% vs 56.23%) and 1.57% greater than the AFP+Lab+ΔAFP approach (61.26% vs 59.69%).”

- Why only the training cohort was restricted to only those with AFP < 400 ng/ml?

Response: Thank you for this question as it highlights that we did not provide sufficient details in our manuscript to address this concern. We had previously stated “The training cohort was then
further restricted to only those with AFP<400ng/ml because the algorithms will only be applied in those patients.” We have replaced the sentence with the following on page 10 of the main manuscript.

“Our goal is to assess the performance of each of the screening algorithms within the OR rule, i.e. the patient has a positive screen if either AFP≥400ng/ml or the screening algorithm indicates a positive screen. Therefore the training cohort was further restricted to only those with AFP<400ng/ml since the screening algorithms will only be applied in those patients. We do not restrict the validation cohort since our goal is to assess the performance of the screening algorithms as they would be used in clinical practice, which includes the OR rule.”

Minor comments:
- page 2, line 30: delete the word "trial";

Response: Thank you. This has been corrected.

- page 2, lines 37-38: "…that is widely used in screening…": Could you provide a reference for this statement or some examples of Countries in the world using AFP for HCC screening? Bruix and Sherman (AASLD PRACTICE GUIDELINE Management of Hepatocellular Carcinoma: An Update) do not recommend the measurement of AFP or other serum biomarkers alone or in combination with USG for surveillance.

Response: Thank you for raising this question. We have included results from a population-based cohort study that found 98% of HCC patients with a prior diagnosis of cirrhosis who received regular surveillance had an AFP testing. In addition, a description of the 2017 update of the AASLD guidelines has been included, which recommend ultrasound with or without AFP. The paragraph of page 2 of the main manuscript now reads:

“Serum α-Fetoprotein (AFP) is a well established diagnostic biomarker for HCC that is widely used in screening despite the wide variation in its reported performance. A population-based US cohort study found that among HCC patients with a prior diagnosis of cirrhosis who received
regular surveillance, 52% received both ultrasonography and AFP, 46% received AFP alone and 2% received ultrasonography alone [4]. A 2017 update of the AASLD guidelines recommends surveillance using ultrasonography, with or without AFP, every six months [5].”

- page 9, line 5-6: change "cummulative" with "cumulative".

Response: Thank you. This has been corrected.