Author’s response to reviews

Title: Systematic characterization of germline variants from the DiscovEHR study Endometrial Carcinoma population

Authors:

Jason Miller (mijason@pennmedicine.upenn.edu)
Raghu Metpally (rmetpally@geisinger.edu)
Thomas Person (tnperson@geisinger.edu)
Sarathbabu Krishnamurthy (sbkrishnamurthy@geisinger.edu)
Venkata Ramesh Dasari (vrdasari@geisinger.edu)
Manu Shivkumar (mkshivakumar@geisinger.edu)
Daniel Lavage (drl5070@yahoo.com)
Adam M Cook (amcook1@geisinger.edu)
David Carey (djcarey@geisinger.edu)
Marilyn Ritchie (marylyn@pennmedicine.upenn.edu)
Dokyoon Kim (dkim@geisinger.edu)
Radhika Gogoi (radhika.gogoi@gmail.com;rpgogoi@geisinger.edu)

Version: 2 Date: 05 Apr 2019

Author’s response to reviews:

Reviewer reports:

Mari Kylleso Halle, Ph.D. (Reviewer 2): General comments:

The authors have made a clear, appropriate and comprehensive response to the review and the suggested changes have been incorporated appropriately into the manuscript. However, the text, and especially the results part, should be revised once more in order to detect typos and to clarify.
I appreciate that you have incorporated a more detailed description of the several patient/non-patient cohorts, although preferably percentages and p-values should be included instead of just patient numbers (Supplementary Table S1, S2). The differences in BMI between the cohorts are described in the results. Still, the potential bias that obesity may introduce is not evaluated. As far as I can see, the detected germ line variants in the DiscovEHR EMCA cohort could just as likely be associated with obesity. In my opinion, this should be mentioned in the discussion.

Minor comments:

1. Page 9, Line 18: Please rephrase the sentence: The OHRM cohort has an obese BMI which is consistent with the association between obesity and cancer'. My suggestion: 'The average BMI of the OHRM cohort is XX, which is classified as obese and consistent with the association between obesity and cancer.'

2. Page 9, Line 24: 'similar distributions grade two samples'. There appears to be a typo here. To clarify, I suggest the following change: 'similar distributions of grade 2 samples'.

3. Page 9, Line 33: the sentence seems to be incomplete. Consider adding 'are performed': 'Downstream analysis of all grades and stages are performed together unless otherwise noted'.

4. Page 11, Line53: Please remove 'the'; it appears to be a typo.

5. Page 16, Line29: Please remove 'this work'; it is written twice and appears to be a typo.

Response:

We appreciate the reviewer’s comments and suggestions. Since we did not make claims about the statistical significance of the sample sizes we felt it was unnecessary to include p-values, however we did add percentages to tables S1 and S2 as the reviewer requested. We have now included discussion concerning obesity being a risk factor, but also having the potential for being independently associated with the variant burden (see last paragraph of the discussion). We have also made the minor changes requested and checked the manuscript for typos. We hope that these edits address the concerns of the editor and reviewer.