**Author’s response to reviews**

**Title:** Validation of the CancerMath prognostic tool for breast cancer in Southeast Asia

**Authors:**

Hui Miao (ephmh@nus.edu.sg)

Mikael Hartman (mikael_hartman@nuhs.edu.sg)

Helena Verkooijen (H.M.Verkooijen@umcutrecht.nl)

Nur Aishah Taib (nuraish@gmail.com)

Hoong-Seam Wong (whseam@yahoo.com)

Shridevi Subramaniam (shridevi@crc.gov.my)

Cheng-Har Yip (chenghar.yip@gmail.com)

Ern-Yu Tan (ern_yu_tan@ttsh.com.sg)

Patrick Chan (patrick_chan@ttsh.com.sg)

Soo-Chin Lee (csilsc@nus.edu.sg)

Nirmala Bhoo-Pathy (ovenjjay@gmail.com)

**Version:** 3  **Date:** 30 Sep 2016

**Author’s response to reviews:**

Dear Editor and Reviewers,

We would like to thank you for the positive feedback and conditionally accepting our manuscript. In this round of revision, we addressed all the points made in your letter. Below, you will find our clarification and actions taken in response to reviewer’s comments. Your and the reviewers’ suggestions have again further improved the paper. We look forward to hearing from you and appreciate your time and effort taken to review our manuscript.
Reviewer #3: Most of my comments have been addressed. However, there seems to be some confusion regarding some of the statistical methods.

The researchers have complete 5- and 10-year follow-up data on all patients in their analysis and this allows them to perform a simplified validation exercise (since they have no censoring to deal with). That is, they can easily calculate the numbers (and proportions) of patients who survive at these time-points and compare them with the predicted proportions. With this in mind, I do not understand why Kaplan-Meier is being used to calculate the 5- and 10- survival rates (page 11, line 42). KM is not required as there is no censoring.

We agree with the reviewer that calculating the proportion of patients who survived will provide the same survival estimates as Kaplan-Meier method since there is no censoring to deal with. We have validated both methods and the survival rates and standard error remained the same, which confirms that there is no censoring in the subsets selected for 5-year and 10-year survival analysis. Therefore we removed ‘Kaplan-Meier analysis’ for survival estimation in our manuscript (method section, page 8, line 43-48, 53-55).

Furthermore, the researchers state that discrimination was assessed using both the AUC and the c-statistic (page 12, line 7-9). I suspect that the researchers mean the "c-index" which is commonly used with survival data (the AUC and c-statistic are the same thing). I note that the interpretation of the c-index is not quite the same as the AUC/c-statistic (page 12, line 18).

As the reviewer correctly pointed out, c-statistics in our manuscript is actually Harrell's c-index. We updated this term in our manuscript (method section, page 9, line 2-4 and results section, page 10, line 53). What we meant by same interpretation as the AUC is that a c-index of 0.5 indicates no discrimination and a c-index of 1.0 means perfect discrimination (method section, page 9, line 7-12).

The new Kaplan-Meier plot (Figure 2) is welcomed but could be improved by showing survival for different sub-groups, e.g. tertiles or quintiles of predicted risk, etc.

We would like to thank the review for the clarification. Kaplan-Meier curves by quintiles of predicted 5-year survival (new Figure 3) is added as suggested. We also reported the log-rank test result to demonstrate the significant difference in survival experience between the five prognostic groups (results section, page 10, 41-48).
The discussion regarding calibration slope (page 16, line 17) is a little cryptic ("…predicted survival do not vary enough"). I suggest that this is spelled out more.

Thanks for the comment. We have elaborated more on the calibration slope as suggested (discussion section, page 13, line 17-22).

Yours Sincerely,

Hui Miao