Reviewer’s report

Title: Risk Prediction for Breast Cancer in Han Chinese Women Based on a Cause-Specific Hazard Model

Version: 0  Date: 10 Sep 2018

Reviewer: Yubei Huang

Reviewer's report:

The authors have developed a risk prediction model to define high-risk population of breast cancer among Han Chinese women. However, the following revisions should be considered to improve their work.

1. In the part of "Study population" at page 5, the authors should give detailed information on the sources of breast cancer patients in the Shandong Case-Control Study, for example, from one hospital, several hospitals, or from local cancer registries. Additionally, the authors should also address whether these cases are only newly diagnosed breast cancers patients, or included patients diagnosed as the second cancer after primary cancer.

2. In the part of "Study population" at page 5, the authors should also give detailed information on Taixing Prospective Cohort Study. First, as described in the paper, 18831 participants completed questionnaire, and 18564 participants who have not been diagnosed as breast cancer at baseline were selected in the study. Namely, 267 (18831-18564) participants had been diagnosed as breast cancer at baseline, which led to a prevalence of breast cancer as 1.4%. This prevalence of breast cancer in China is so high. How were these baseline breast cancer patients found? Self-reported, or screening by several modalities? Second, 23 participants were excluded due to missing information of risk factors. However, the authors did not give a clear list of these risk factors. Third, 5365 subjects from 18541 eligible subjects were further excluded due to lost to follow-up. The author should clearly mark the rate of lost to follow-up as 28.9%, since the rate of lost to follow-up is also so high in a 7-year cohort. Fourth, after excluding baseline breast cancer patients, how were the newly incidence cancers were identified during the follow-up, from local cancer registries, or from active follow-up with telephone, face-to-face interview, or reviews from case report forms in hospitals.

3. In the part of "Measurements and definition of risk factors" at page 6, the author mentioned only two reproductive factors (such as Number of abortions and age at first live birth) were collected in the standardized questionnaire, several other important reproductive factors, such as number of living births and months of breast feeding, were not collected. Moreover, several other risk factors, including age at menarche, age at postmenopause, oral...
contraceptive, and hormone replacement therapy, were also not collected in the questionnaire. The author should address these limitations in the end of discussion. Additionally, BMI should be given complete spelling when it first appeared at page 6. The author should also address how weight and height were collected, self-reported or measured. Lastly, how was life satisfaction scores scale developed or revised from previous similar scales, and the author should address why 13 were selected as the cut-off values of satisfaction or unsatisfaction.

4. RRs and AR should be given complete spelling when it first appeared at page 6.

5. In the part of "Statistical methods" at page 6, 13 age groups were defined. However, the incidence rates of breast cancer for Chinese females aged either younger than 30 years old or elder than 85 years are very low. Therefore, age-specific risk predictions among Chinese females aged either younger than 30 years old or elder than 85 years seem useless.

6. In the part of "Statistical methods" at page 7, the authors should address which packages of R software were used in their studies.

7. At page 8, the author should list the P values for the comparisons between cancers and non-cancers in table 1.

**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.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

No

**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 am able to assess the statistics
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