Reviewer’s report

Title: Application of three statistical models for predicting the risk of diabetes

Version: 0 Date: 04 Mar 2019

Reviewer: Wen-Yuan Lin

Reviewer's report:

Please include all comments for the authors in this box rather than uploading your report as an attachment. Please only upload as attachments annotated versions of manuscripts, graphs, supporting materials or other aspects of your report which cannot be included in a text format.

Please overwrite this text when adding your comments to the authors.
This is a community study to explore the risk factors and conduct prediction models for diabetes in China. The approach is not very novel. Previous studies have provided prediction model for undiagnosed diabetes or incident diabetes. Several critical concerns should be addressed.

1. The rationale for this study was not exposed clearly enough. Many previous studies with large sample reported risk models for screening undiagnosed diabetes in China population, such as Gao, W.G. et. al. (2010), Li, W., et al. (2018), and Wang, A., et al. (2016) etc. The same population (China) was studied in this study, but what gap between previous studies and clinical practice would be bridged wasn't illustrated clearly. This point is crucial and should be documented while considering publication in a medical journal.

2. Only one dataset was used. The validation of modeling is absent.

3. In Abstract section, please identify abbreviations or acronyms should be explained when first used. (such as ROC, OR, TG, TC, AUC, and BP). In results of abstract, the description of risk factors should be TG$\geq$1.17mmol/L and age$\geq$70 years, instead of TG and age.

4. In Background section, please indicate what gap between previous studies and clinical practice would be bridged by this study.

5. The outcome definition is not clear in terms of what kind of diabetes patients was predicted, because the exclusion criteria of study subjects did not mention. In general, the patients with known diabetes should be excluded in a prediction study for screening undiagnosed patients or for incident cases. Providing a flowchart of study subject's recruitment in Objects and Methods section is suggested.
6. Please identify the year of study and the study sample crawled by using probability or non-probability sampling in Objects and Methods section. Furthermore, the definition of dyslipidemia by only using laboratory tests is not appropriate. The dyslipidemia history or dyslipidemia medication should be considered in this study.

7. In the analysis of logistic regression model, all lipid markers (TG, TC, LDL-C, HDL-C) were considered synchronously in one mode. Collinearity of these markers may happen because of high correlation among them.

8. The discussion of risk factors (the second to fifth paragraphs in Discussion section) is not appropriate. Authors used the results of bivariate analysis instead of multivariate analysis. Gender, level of LDL-C, smoking and stroke were not an independent risk factor according to the results of multivariate logistic regression model. It is not correct to report those factors were associated with diabetes.

9. Tables and figures

1) Please identify the first letter of variable name in upper case in all tables.

2) It would be helpful to provide readers with the marginal distribution of all independent variables in Table 1, 2, and 3.

3) In table 3, the cut-off point of TC, TG, LDL-C, and LDL-C should be provided.

4) Figure 1 provided the same information which was shown in the last column of table 7. Please delete figure to avoid redundant expression.

10. Please consult a Native American to let the readers to easy reading. Please do English revision.
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.

Yes

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:

Not suitable for publication unless extensively edited

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