

Reviewer #1:

**Question:**

**Specific Comments to Authors:** This paper titled "Age-specific differences in the association between prediabetes and cardiovascular diseases in China: A national cross-sectional study" is interesting and important for CVD prevention through prediabetes. It analyzed the data of epidemiological survey from June 2007 to May 2008 of the Chinese people, including cigarette smoking, elevated systolic blood pressure, low-density lipoprotein cholesterol, and living in rural areas and concluded that "the susceptibility to CVD is age-specific in newly-diagnosed prediabetes." It is of great significance for preventing and controlling the incidence of CVD. It is interesting that moderate drink of alcohol has advantageous impact. If possible, would you please give more explanations or possible reasons? The authors should check their manuscript in detail avoiding some mistakes before publish, such as punctuation marks in the abbreviation list and some data in table 1 (data of women seem lost?).

**Answer:**

**Dear reviewer,**

Thanks for your suggestions.

First, I found through literature review that for low to moderate alcohol consumption, a lower CVD risk is observed compared to abstaining and excessive drinking in some observation studies. Nevertheless, as most of the protective evidence of low to moderate alcohol consumption on CVD is from observational studies, it is uncertain whether this effect is a result of different forms of bias. Due to the lack of RCT, which is the gold standard, the focus in research has now shifted to new analytical methods, such as Mendelian randomization studies. However, none of these studies could truly resolve the pressing question of whether alcohol is the protective factor of CVD. Therefore, there is still remaining the controversy regarding the effects of moderate alcohol consumption on CVD.

Second, I have checked the data punctuation marks in the abbreviation list and the data in the table.

Sincerely,

Bo Zhang

Reviewer #2:

**Question:Specific Comments to Authors:** I am really grateful to review this manuscript. In my opinion, this manuscript can be published once some revision is done successfully. I made one suggestion and I would like to ask your kind understanding. This study used numeric data from 46,239 prediabetes participants, applied logistic regression and identified major predictors varying across different age groups for the prediction of their cardiovascular disease (i.e., younger than vs. older than 40 years). I would argue that this is a rare achievement. However, it can be noted that the random forest outperforms logistic regression in terms of prediction and random forest variable importance and Shapley Additive Explanations (SHAP) summary plot are very effective to identify the strength and direction of association between cardiovascular disease and its major predictor. In this context, I would like to ask the authors to derive random forest variable importance and SHAP summary plot.

Answer:

Dear reviewer,

Thanks for your suggestions.

Random forest outperforms logistic regression in terms of prediction, therefore, we derive random forest variable importance and SHAP summary plot for both age groups in the revised manuscript. We found that in the younger age group, 10-year risk of CVD was found to be more closely linked to family history of CVD rather than lifestyle, whereas in the older age group, resident status was more closely linked. Thus it is necessary to develop targeted approaches for the prevention and management of CVD in adults across various age brackets.

Sincerely,

Bo Zhang