

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prevalence and risk factors for hyperhomocysteinemia: a population based cross-sectional study from Hunan, China
AUTHORS	Yang, Yide; Zeng, Yuan; Li, Feifei; Zheng, Chanjuan; Dong, Yanhui; Li, Jian; He, Quanyuan; Ye, Xiangli; Lv, Yuan; Hocher, Carl-Friedrich; Kraemer, Bernhard; Hong, Xiuqin; Hocher, Berthold

VERSION 1 – REVIEW

REVIEWER	Oliveira, Isabel Federal University of Pelotas (UFPel), Postgraduate Programme in Epidemiology
REVIEW RETURNED	12-Mar-2021

GENERAL COMMENTS	<p>The manuscript should be edited for occasional grammatical mistakes and English writing.</p> <p>Abstract</p> <p>Line.28</p> <p>Have the negative results from alcohol and physical activity associations with hyperhomocysteinemia influenced the aim of the study? The study title, methods and results include other risk factors.</p> <p>Introduction</p> <p>Line. 66</p> <p>Symbols for genes should be italicized.</p> <p>Lines.70-71</p> <p>Even though scarce literature, the authors should include references about the affirmations at Lines. 70-71 (“There are few studies looking at underlying causes of elevated homocysteine concentrations in the blood in humans”) and at Lines. 78-80 (“Diet as a risk factor for hyperhomocysteinemia has been insufficiently studied. Possible effects of smoking on the risk of developing hyperhomocysteinemia are also lacking.”)</p> <p>Methods</p> <p>Line.96</p>
-------------------------	---

	<p>Have the authors considered the differences known about hyperhomocysteinemia in urban and rural Chinese communities at sampling?</p> <p>Data collection and questionnaire survey</p> <p>Was physical activity measurement based on The International Physical Activity Questionnaire (IPAQ) (Short)?</p> <p>The references used to describe the methods for alcohol drinking and physical activity measurements should be reviewed based on the following observations:</p> <p>There is no explanation about alcohol consumption in the reference Hong et al.,2016.</p> <p>Lu et al. (2018) did not assess diet and physical activity in their study.</p> <p>It is better to standardize the name of variable: “Alcohol drinking”, “drinking”, “alcohol drinking situation”</p> <p>Line.124</p> <p>The number of alcohol drinking categories studied should be reviewed. In the text, it is mentioned two groups but four groups were described in the tables.</p> <p>Line.139-142</p> <p>Is the affirmative “at least for ten minutes in the past week” correct?</p> <p>Line.150</p> <p>“values where the first and fifth Korotkoff sounds appeared.”</p> <p>As known, the phase I (appearance of a tone) and the phase V (disappearance) of Korotkoff sounds correspond to systolic and diastolic blood pressure, respectively.</p> <p>Lines. 161-162</p> <p>According to one reference mentioned (Veeranna et al., 2011), hyperhomocysteinemia was defined as homocysteine $\geq 15 \mu\text{mol/l}$ but not in the other reference cited (Liu et al., 2015).</p> <p>Lines.167-169</p> <p>The data about risk of hypertension according to homocysteine categories should be presented in the results section with the description of models or only showed at supplementary material.</p> <p>Statistical Analyses</p> <p>It is important to mention that homocysteine is a variable without normal distribution. Also, to justify/explain the use of Z score in the statistical section.</p>
--	--

	<p>Was sex interaction tested? Have the authors explored the data performing analyses stratified by sex as presented in the Table.1? Differences in blood homocysteine concentration and lifestyle between sexes should be taken into account.</p> <p>To avoid problems with estimation, were the independent variables tested to know correlation between them?</p> <p>Although both terms often used interchangeably in the literature, multivariate analysis seems not to be equal to multivariable analysis. It is necessary to better clarify the analyses performed.</p> <p>Why variables that not showed to influence significantly the risk of hyperhomocysteinemia in the univariate analysis were maintained in the multivariable analysis? Have the authors suspected of a specific interaction with other variables? It is important to clarify in methods.</p> <p>Results</p> <p>Line.207</p> <p>The sentence at line 207 seems to be unnecessary.</p> <p>Lines.229-230</p> <p>“With multivariate linear regression, we found older age, male, alcohol drinking (0=non-drinker, 1=drinker), intake of fish or seafood is associated with higher level of homocysteine (P <0.05, Table S2)”</p> <p>This affirmative is in conflict with the data presented in the Table S2. Some associations are negative but others, positive such as for Drinking -0.71 (-1.33, -0.09) p=0.024 and for fish/seafood 0.61(0.11, 1.10) p=0.017.</p> <p>Table 1</p> <p>The variables should be showed by groups: sociodemographic, anthropometric, behavioral and diseases variables.</p> <p>Annual household income (¥) variable was not mentioned previously in the methods.</p> <p>The binary variable (fish/seafood) was not included in the legend of Table.1. Also, in the legend of Table.2 and Table.3.</p> <p>Table.3</p> <p>Although it was written multivariable in the title of Table.3 (pg.16) it was described as multivariate at line 215.</p> <p>Table.S2</p> <p>It was written multivariable in the title of Table.S2 (pg.32) but multivariate at line 230.</p> <p>The legend of Table.S2 should include the categorization of all</p>
--	---

	<p>variables.</p> <p>Discussion</p> <p>The discussion could be improved.</p> <p>It is suggested to begin with a summary of findings. Also, strengths and limitations of the study could be showed together.</p> <p>Could bringing the rural and urban communities together have influenced the results?</p> <p>Although negative results for alcohol consumption and physical activity associations were found, it is important to mention and to compare with data from the literature.</p> <p>Lines.245-246</p> <p>The references mentioned only reported data from studies that found high prevalence of hiperhomocysteinemia. It is suggested review the references.</p>
--	--

REVIEWER	Challa, Feyissa Ethiopian Public Health Institute, National References Laboratory for Clinical Chemistry
REVIEW RETURNED	13-Mar-2021

GENERAL COMMENTS	<p>The authors estimated the prevalence of hyperhomocysteinemia in China Population 4012. They observed a homocysteine concentration greater than 15 $\mu\text{mol} / \text{L}$ in 35.4 % of cases and an association with age,BMI and negative effect fruit and vegetable consumption. The work is based on a large population but lacks essential data, including creatinine concentration, folate, and vitamin B12 status. As a result, the public health impact of the data presented is limited.</p> <p>specific comments</p> <p>I suggest the title to be <<Prevalence and risk factors for hyperhomocysteinemia: a population-based cross-sectional study from Hunan, China>></p> <p>abstract page #3, line 33.34, It says a total of 4012 males 1641, and female 2648...please correct it</p> <p>Result in abstract page #3, line 39 The prevalence of hyperhomocysteinemia and is 35.4% (45.4% vs 28.5% for men...please delete "and"</p> <p>abstract part..include the methodology used to measure Hcy.</p> <p>Introduction</p> <p>Not only MTHFR involved in Hcy but also Folate, Vitamin B12, Vitamin D, renal diseases, hypothyroidism ..etc</p> <p>data analysis</p> <p>Did the author checked for normal distribution data</p> <p>Results</p> <p>How to explain that there are 59% of participants who are women?</p> <p>Table 2</p> <p>is better to put (homocysteine<15$\mu\text{mol}/\text{L}$) or (homocysteine>15$\mu\text{mol}/\text{L}$)...case and control don't define by HCY value only.</p> <p>suggestion..</p> <p>It will be better if we put both Univariate logistics and Multivariable logistic regression in one table</p>
-------------------------	--

	<p>In both Univariate logistics and Multivariable logistic regression, it will better to use reference those decreases HCy value..for example...those who consumed Fruit and vegetable as references, high Physical activity as references, etc table ...check all table value - page 18 line 234...Discussion: Despite what is indicated in the discussion and conclusion, this work is not the first to study the prevalence of hyperhomocysteinemia in China. Most of Hcy study was from China</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer# 1

Dr. Isabel Oliveira, Federal University of Pelotas (UFPEL)

Comments to the Author:

Comments

The manuscript should be edited for occasional grammatical mistakes and English writing.

Response: The authors thank the reviewer for his valuable and helpful comment. We have edited the whole manuscript for the grammatical mistakes and inappropriate English writing. And all the changes were marked with track changes.

Abstract

Line.28

Have the negative results from alcohol and physical activity associations with hyperhomocysteinemia influenced the aim of the study? The study title, methods and results include other risk factors.

Response: The authors thank the reviewer for his valuable and helpful comment. The aim of the study was predefined as outlined in our manuscript. We defined outcomes a priori before doing any analysis. Although, no significant association between alcohol consumption and physical activity and risk of hyperhomocysteinemia was found in the present study, for physical activity, the P value for this association is quite borderline ($P=0.050$), and the direction of association estimates is consistent with previous studies^{51,52} showing that higher level of physical activity is protective for high homocysteine. Similarly, the direction of association between alcohol drinking and risk of hyperhomocysteinemia is consistent with previous studies, but not reach the statistical significance^{53,54}. And we think the non-significant result for alcohol and physical activity wouldn't influence the aim of the present study, since in the multivariable logistic regression models, the significant associations between smoking, fruit and vegetable consumption and risk of hyperhomocysteinemia were independent of alcohol drinking and physical activity.

51. e Silva Ade S, da Mota MP. Effects of physical activity and training programs on plasma homocysteine levels: a systematic review. *Amino Acids* 2014;46:1795-804.

52. Buckner SL, Loenneke JP, Loprinzi PD. Single and combined associations of accelerometer-assessed physical activity and muscle-strengthening activities on plasma homocysteine in a national sample. *Clin Physiol Funct Imaging* 2017;37:669-74.

53. Coppola M, Mondola R. Correlation between plasma homocysteine levels and craving in alcohol dependent stabilized patients. *Clin Nutr* 2018;37:1061-65.

54. Choi SH, Choi-Kwon S, Kim MS, et al. Poor nutrition and alcohol consumption are related to high serum homocysteine level at post-stroke. *Nutr Res Pract* 2015;9:503-10.

Introduction

Line. 66

Symbols for genes should be italicized.

Response: Done. We have revised the symbols for genes according to the comment.

Lines.70-71

Even though scarce literature, the authors should include references about the affirmations at Lines. 70-71 ("There are few studies looking at underlying causes of elevated homocysteine concentrations in the blood in humans") and at Lines. 78-80 ("Diet as a risk factor for hyperhomocysteinemia has been insufficiently studied. Possible effects of smoking on the risk of developing hyperhomocysteinemia are also lacking.")

Response: We have added relative references for the affirmations listed above as follows.

Line 76-77

There are few studies looking at underlying causes of elevated homocysteine concentrations in the blood in humans^{11 12}.

Line 84-86

Diet as a risk factor for hyperhomocysteinemia has been insufficiently studied²². Possible effects of smoking on the risk of developing hyperhomocysteinemia are also lacking²³.

11. Elsherbiny NM, Sharma I, Kira D, et al. Homocysteine Induces Inflammation in Retina and Brain. *Biomolecules* 2020;10

12. Chen Y, Li J, Li T, et al. Association between homocysteine, vitamin B 12 , folic acid and erectile dysfunction: a cross-sectional study in China. *BMJ Open* 2019;9:e023003.

22. Zhang W, Li Y, Wang TD, et al. Nutritional status of the elderly in rural North China: a cross-sectional study. *J Nutr Health Aging* 2014;18:730-6.

23. Bazzano LA, He J, Muntner P, et al. Relationship between cigarette smoking and novel risk factors for cardiovascular disease in the United States. *Ann Intern Med* 2003;138:891-7.

Methods

Line.96

Have the authors considered the differences known about hyperhomocysteinemia in urban and rural Chinese communities at sampling?

Response: According to a systematic review and meta-analysis about prevalence of hyperhomocysteinemia in China, there is no significant difference in the prevalence of hyperhomocysteinemia between urban areas and rural areas (28.1% vs 26.5%, for rural and urban areas respectively) [1]. Therefore, for the sampling, the differences about hyperhomocysteinemia in urban and rural Chinese communities were not considered at sampling.

[1]Yang B, Fan S, Zhi X, Wang Y, Wang Y, Zheng Q et al. Prevalence of hyperhomocysteinemia in China: a systematic review and meta-analysis. *Nutrients*. 2014; 7(1):74-90.

<https://doi.org/10.3390/nu7010074>.

Data collection and questionnaire survey

Was physical activity measurement based on The International Physical Activity Questionnaire (IPAQ) (Short)?

Response: Physical activity measurement was not based on The International Physical Activity Questionnaire (IPAQ) (Short). For investigation of physical activity, we referred the questionnaire of

the China Health and Retirement Longitudinal Study (CHARLS) [25]. We added more description of the questionnaire survey in the Method section (Line 113-117) as follows.

Data of subjects were collected with a standard and reliable questionnaire by face-to-face interviews¹⁷. The questionnaire contains demographic information, cigarette smoking status, alcohol drinking situation, dietary factors, walking and physical activity level¹⁷. For the demographic factors, physical activity, alcohol drinking and smoking behavior, we referred the questionnaire of the China Health and Retirement Longitudinal Study (CHARLS)²⁵.

25. Wang Z, Zou Z, Yang Z, et al. Association between exposure to the Chinese famine during infancy and the risk of self-reported chronic lung diseases in adulthood: a cross-sectional study. *BMJ Open* 2017;7:e015476.

The references used to describe the methods for alcohol drinking and physical activity measurements should be reviewed based on the following observations:

There is no explanation about alcohol consumption in the reference Hong et al.,2016.

Lu et al. (2018) did not assess diet and physical activity in their study.

Response: We are sorry for this mistake, and have deleted the unrelated references.

For physical activity, we referred the questionnaire of the China Health and Retirement Longitudinal Study (CHARLS)²⁵.

For alcohol drinking, we referred to Lu et al's study³⁰. Alcohol drinking is categorized into four groups: Abstainer (Non-drinker), light or moderate drinker (0-2 drinks/day), heavy drinker (≥ 2 drinks/day) and drinker who did not report the specific drinks³⁰.

25. Wang Z, Zou Z, Yang Z, et al. Association between exposure to the Chinese famine during infancy and the risk of self-reported chronic lung diseases in adulthood: a cross-sectional study. *BMJ Open* 2017;7:e015476.

30. Lu W, Xu J, Taylor AW, et al. Analysis of the alcohol drinking behavior and influencing factors among emerging adults and young adults: a cross-sectional study in Wuhan, China. *BMC Public Health* 2019;19:458.

It is better to standardize the name of variable: "Alcohol drinking", "drinking", "alcohol drinking situation"

Response: According to the suggestion, we have revised the whole paper to standardize the name of the variable.

Line.124

The number of alcohol drinking categories studied should be reviewed. In the text, it is mentioned two groups but four groups were described in the tables.

Response: Sorry for this mistake, we have revised it as four groups: group 1: Abstainer (Non-drinker), group 2: light or moderate drinker (0-2 drinks/day), group 3: heavy drinker (≥ 2 drinks/day) and group 4: drinker who did not report the specific drinks³⁰.

30. Lu W, Xu J, Taylor AW, et al. Analysis of the alcohol drinking behavior and influencing factors among emerging adults and young adults: a cross-sectional study in Wuhan, China. *BMC Public Health* 2019;19:458.

Line.139-142 Is the affirmative "at least for ten minutes in the past week" correct?

Response: Sorry for that we did not describe this appropriately. Here we mean that if the duration of the participants' physical activity is less than 10 continuous minutes every time, and we categorized this as invalid physical activity, the physical activity lasting for at least 10 continuous minutes was categorized as valid physical activity [31]. And then if the participants didn't achieve any valid physical activity in the last week, then we considered this participant was physically inactive (no valid physical activity). We have revised the description accordingly as follows (Line 144-148).

For physical activity, subjects were categorized as two groups (physically inactive or not physically inactive). Physical activity lasting for at least 10 continuous minutes was categorized as valid physical activity, less than 10 continuous minutes as invalid physical activity³¹. Subject who had no valid physical activity in the past week was considered as physically inactive, otherwise not physical inactive.

[31]Craig R, Mindell J, Hirani V. Health survey for England 2008. Physical activity and fitness. summary of key findings. 2009.

Line.150

“values where the first and fifth Korotkoff sounds appeared.”

As known, the phase I (appearance of a tone) and the phase V (disappearance) of Korotkoff sounds correspond to systolic and diastolic blood pressure, respectively.

Response: The authors thank the reviewer for his valuable comment. We are sorry for this mistake and revised it accordingly as follows (Line 156).

Systolic BP (SBP) and diastolic BP (DBP) were recorded as the values where the first Korotkoff sound appeared and the fifth Korotkoff sound disappeared.

Lines. 161-162

According to one reference mentioned (Veeranna et al., 2011), hyperhomocysteinemia was defined as homocysteine $\geq 15 \mu\text{mol/l}$ but not in the other reference cited (Liu et al., 2015).

Response: In Liu et al's study high homocysteine is defined as $>15\text{mmol/l}$, slightly different from definition of $\geq 15 \mu\text{mol/l}$. We have revised the reference accordingly (Line 170).

32. Ying Y, Lin S, Kong F, et al. Ideal Cardiovascular Health Metrics and Incidence of Ischemic Stroke Among Hypertensive Patients: A Prospective Cohort Study. *Front Cardiovasc Med* 2020;7:590809.

Lines.167-169

The data about risk of hypertension according to homocysteine categories should be presented in the results section with the description of models or only showed at supplementary material.

Response: We have moved this part into the result section with the description of the models as follows (Line 212-219).

For the definition of hyperhomocysteinemia, we used the cut-off point of $15 \mu\text{mol/L}$. We categorized the homocysteine level into three groups ($<10 \mu\text{mol/l}$; $10\text{-}14.9 \mu\text{mol/l}$; and $\geq 15 \mu\text{mol/L}$), and then used the new categorical homocysteine variable to predict the risk of hypertension in different models with adjustment of potential covariates, we found in the full adjusted model only for participants with homocysteine $\geq 15 \mu\text{mol/L}$ the risk of hypertension increased significantly (OR=3.74, 95%CI: 3.06-4.56) compared with participants with homocysteine $<10 \mu\text{mol/L}$, but not in the second group ($10\text{-}14.9 \mu\text{mol/l}$) with OR=1.114 (95%CI: 0.91-1.36) (Table S1).

Statistical Analyses

It is important to mention that homocysteine is a variable without normal distribution. Also, to justify/explain the use of Z score in the statistical section.

Response: We checked the normal distribution of the continuous variables, and they are all without normal distribution, so we revised the method of description of these variables with median and inter quartile range (IQR). The corresponding content in the Method section were also revised accordingly as follows (Line 180). For the use of Z score of homocysteine in the result section, Z score of homocysteine were calculated for better comparison of associations between different variables and homocysteine. We added the reason and formula of Z score in the statistical section as follows.

Line 172-173

Medians and inter-quartile range (IQR) were calculated for numerical variables without normal distribution.

Line 177-180

Z score of homocysteine were calculated for better comparison of associations between different variables and homocysteine. Z score of homocysteine = (original value of homocysteine – mean value)/standard deviation. So, the variable (homocysteine) was standardized to a mean of 0 and an SD of 1 before analysis.

Was sex interaction tested? Have the authors explored the data performing analyses stratified by sex as presented in the Table.1? Differences in blood homocysteine concentration and lifestyle between sexes should be taken into account.

Response: We further analyzed the sex interaction with other risk factors on the risk of hyperhomocysteinemia. Firstly, we did the stratified analysis by sex, then we included the sex interaction item in the logistic regression models to test the significance of the interaction. And the results were presented in the following table (Table 3). We found there is a significant interaction between sex and education level or alcohol drinking on the risk of hyperhomocysteinemia. We have added the results of sex interaction in the result section, and also relative description and discussion were added in the manuscript.

Line 186-192 (In the Method section)

Stratified analyses of sex of association between risk factors and hyperhomocysteinemia were conducted, then the sex interactions were tested with multivariable logistic regression models including the interaction term with age, BMI, sex, marriage status, education level, annual household income, smoking, alcohol drinking, walking, physical activity, fruit and vegetable, sweet food, pickled vegetables as covariates. In the stratified analyses, sex interaction was defined as different effect sizes in female and male group, and the P value for interaction was statistically significant ($P_{interaction} < 0.05$)³³.

33. Dempfle A, Scherag A, Hein R, et al. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. *Eur J Hum Genet* 2008;16:1164-72.

Line 259-273 (In the Result section)

For the interaction between sex and other factors on the risk of hyperhomocysteinemia, we did stratified analysis by sex, and also did the interaction analysis (Table 3). In the female group, participants with college and above education have a significantly higher risk of hyperhomocysteinemia (OR=1.70, 95%CI: 1.18-2.46) compared with participants with under primary school education, while in the male group, participants with college and above education had a lower risk of hyperhomocysteinemia (OR=0.60, 95%CI: 0.41-0.87). When further interaction analysis was conducted, we identified interaction effects between sex and education level on risk of

hyperhomocysteinemia (Pinteraction=0.031).

Similar interaction effect between sex and alcohol drinking on risk of hyperhomocysteinemia was also found (Pinteraction <0.001). In the female group, participants who were light or moderate alcohol drinker had a significantly lower risk of hyperhomocysteinemia (OR=0.47, 95%CI: 0.30-0.74) compared with participants who were not alcohol drinker, and female heavy alcohol drinker had a more than triple risk of risk of hyperhomocysteinemia (OR=3.74, 95%CI: 1.41-9.88). While in the male group, compared with non-alcohol drinker, light or moderate male drinker have a significantly higher risk of hyperhomocysteinemia (OR=1.69, 95%CI:1.28-2.23), heavy drinker also tended to have a higher risk of hyperhomocysteinemia but the association did not reach statistical significance (P>0.05).

Line 333-348 (In the Discussion section)

Interaction between sex and education level, or alcohol drinking

Interestingly, we also found there are significant interactions between sex and education level, sex and alcohol drinking on the risk of hyperhomocysteinemia. In our study, the association between education level or alcohol drinking with risk of hyperhomocysteinemia is quite different in males and females. Higher education level (college and above) is a risk factor of hyperhomocysteinemia for females but seems to be a protective factor for male. Also, the association between heavy alcohol drinker and risk of hyperhomocysteinemia is much more pronounced in females than in males. And light alcohol drinker was a protective factor in women but a risk factor in men. These sex-different associations are independent of other potential covariates, including age, smoking, alcohol drinking, and BMI et al. Actually, many studies have demonstrated the association between alcohol drinking and homocysteine, but inconsistent result were found. Some studies showed that alcohol drinking is associated with higher level of homocysteine^{55 56}, while some other studies reported a U-shaper relationship and found that mild to moderate alcohol drinking was associated with lower homocysteine^{57 58}. The association between alcohol drinking should be examined in future longitudinal or interventions studies. These sex interactions were firstly examined in our study, not reported in previous works. But sex-smoking interaction, and gene-environment interaction were examined in previous studies^{59 60}.

55. van der Gaag MS, Ubbink JB, Sillanaukee P, et al. Effect of consumption of red wine, spirits, and beer on serum homocysteine. *Lancet* 2000;355:1522.

56. Cravo ML, Gloria LM, Selhub J, et al. Hyperhomocysteinemia in chronic alcoholism: correlation with folate, vitamin B-12, and vitamin B-6 status. *Am J Clin Nutr* 1996;63:220-4.

57. Dixon JB, Dixon ME, O'Brien PE. Reduced plasma homocysteine in obese red wine consumers: a potential contributor to reduced cardiovascular risk status. *Eur J Clin Nutr* 2002;56:608-14.

58. Refsum H, Nurk E, Smith AD, et al. The Hordaland Homocysteine Study: a community-based study of homocysteine, its determinants, and associations with disease. *J Nutr* 2006;136:1731S-40S.

59. Mu L, Lin Y, Huang X, et al. Sex differences in the prevalence and clinical correlates of hyperhomocysteinemia in patients with bipolar disorder. *Hum Psychopharmacol* 2020;35:e2724.

60. Kim HJ, Kim MK, Kim JU, et al. Major determinants of serum homocysteine concentrations in a Korean population. *J Korean Med Sci* 2010;25:509-16.

To avoid problems with estimation, were the independent variables tested to know correlation between them?

Response: We have quantified the possible multicollinearity between the independent variables using the variance inflation factor (VIF) of collinearity diagnostics. A VIF larger than 4.0 was considered an indication of severe multicollinearity in the regression model. VIFs were all less than 4.0 indicated that there was no significant collinearity between the independent variables (Table S3). The corresponding results and description of collinearity were added in the manuscript as follows.

Line 192-195

The possible multicollinearity between the independent variables was quantified using the variance inflation factor (VIF) of collinearity diagnostics. A VIF larger than 4.0 was considered an indication of severe multicollinearity in the regression model³⁴.

Line 237-239

With the collinearity diagnosis between the independent variables, VIFs were all less than 2.0 indicated that there was no significant collinearity between the independent variables (Data not shown).

[34]Yang X, Luo W, Han S, Zha L, Zhang J, Li X et al. Prevalence of high-risk coronary plaques in patients with and without metabolic syndrome and the relationship with prognosis. *BMC Cardiovasc Disord.* 2020; 20(1):73. <https://doi.org/10.1186/s12872-020-01358-8>.

Although both terms often used interchangeably in the literature, multivariate analysis seems not to be equal to multivariable analysis. It is necessary to better clarify the analyses performed.

Response: We have revised it accordingly and clarify the statistical analysis performed in our study as follows.

Line 181-186

Crude odd ratio (OR) and adjusted OR with 95% confidence interval (95%CI) of having hyperhomocysteinemia were determined by using univariable or multivariable logistic regression models. For the multivariable logistic regression model, we included all the potential variables (age, sex, BMI, marriage status, education level, annual household income, smoking, alcohol drinking, walking, physical activity, fruit and vegetable, sweet food, pickled vegetables) in the models to adjusted ORs for risk factors.

Why variables that not showed to influence significantly the risk of hyperhomocysteinemia in the univariate analysis were maintained in the multivariable analysis? Have the authors suspected of a specific interaction with other variables? It is important to clarify in methods.

Response: The authors thank the reviewer for his valuable and constructive comment. We also conducted the multivariable logistic regression analysis only including the statistically significant influencing factors of hyperhomocysteinemia in the model, similar result was found as follows (Table S2). For the potential interaction with other variables, we thank the reviewer for this constructive comment. We further made stratified analysis of sex and found significant sex interaction with education level and alcohol drinking on the risk of hyperhomocysteinemia (Table 3). We have added the results of sex interaction in the result section, and also relative description and discussion were added in the manuscript.

We added this additional analysis in the supplementary material and relative description was also added in the statistical analysis part or in the result section as follows (Line 239-241).

For the sensitivity analysis of the multivariable logistic regression, we also calculated a model with only including significant risk factors in the univariable analysis as covariates, similar results were found (Table S2).

Results

Line.207

The sentence at line 207 seems to be unnecessary.

Response: We have deleted this sentence according to the Reviewer's comment.

Lines.229-230

“With multivariate linear regression, we found older age, male, alcohol drinking (0=non-drinker, 1=drinker), intake of fish or seafood is associated with higher level of homocysteine (P <0.05, Table S2)”

This affirmative is in conflict with the data presented in the Table S2. Some associations are negative but others, positive such as for Drinking -0.71 (-1.33, -0.09) p=0.024 and for fish/seafood 0.61(0.11, 1.10) p=0.017.

Response: According to the reviewer’s suggestion, we did stratified analysis of sex and found significant sex interactions with alcohol drinking on the risk of hyperhomocysteinemia. And we found especially in female participants, the association between alcohol drinking and homocysteine might not be linear but U-shaped, since we found that light to moderate alcohol drinker is protective for hyperhomocysteinemia in females, but heavy drinker is a significant risk factor in females (Table 3) as follows. We think the linear regression analysis is not appropriate so we deleted this result table.

Table 1

The variables should be showed by groups: sociodemographic, anthropometric, behavioral and diseases variables.

Annual household income (¥) variable was not mentioned previously in the methods.

The binary variable (fish/seafood) was not included in the legend of Table.1. Also, in the legend of Table.2 and Table.3.

Response: We categorized the variables into sociodemographic, anthropometric, behavioral and diseases variables in Table 1. We added the relative description for fish/sea food in the legend of Table 1, Table 2 and Table 3. For fish or seafood: 1=consumed fish or seafood once every week, 0=didn’t consumed fish or seafood once every week.

The annual household income (¥) variable was added in the Method section as follows (Line 125-126).

The income of the family every year (annual household income, ¥) was asked for every participant.

Table.3

Although it was written multivariable in the title of Table.3 (pg.16) it was described as multivariate at line 215.

Response: The authors thank the reviewer for his helpful comment. We have revised it all as multivariable in the whole manuscript.

Table.S2

It was written multivariable in the title of Table.S2 (pg.32) but multivariate at line 230.

The legend of Table.S2 should include the categorization of all variables.

Response: We have revised it all as multivariable in the whole manuscript. And the categorization of all variables was all included the legend of all Tables.

Discussion

The discussion could be improved.

It is suggested to begin with a summary of findings. Also, strengths and limitations of the study could be showed together.

Response: We have added a summary of findings, the strengths and limitations of our study in the

Discussion section as follows (Line 276-285).

To the best of our knowledge, our study represents the only analyses of high homocysteine and its related risk factors in a large, population-based central southern Chinese population to date. The results showed that the prevalence of hyperhomocysteinemia is relatively high (total, 35.4%, and 45.4% vs 28.5% for men and women). Interestingly, we found that current smoking is a risk factor for high homocysteine, but not for smoking quitting. Notably, fruit and vegetable intake may decrease the risk of hyperhomocysteinemia with adjustment of potential covariates. Male sex, older age and higher BMI are risk factors for hyperhomocysteinemia. In addition, we examined significant interactions between sex and educational level, sex and alcohol drinking on the risk of hyperhomocysteinemia. But only associations, not causal relationships could be concluded in this cross-sectional study, future cohort studies or intervention studies should be done to verify these findings.

Could bringing the rural and urban communities together have influenced the results?

Response: For the influence of community's location (rural/urban) on the association between risk factors and hyperhomocysteinemia, we made further analysis to explore the potential impact of area on the associations, and the results keep consistent. We made a multivariable logistic regression model with including the area of the communities as a covariate, and quite similar results were found (Table S4).

Although negative results for alcohol consumption and physical activity associations were found, it is important to mention and to compare with data from the literature.

Response: We reviewed more papers about hyperhomocysteinemia and made a deeper discussion (Line 327-332).

Although, no significant association between alcohol consumption and physical activity and risk of hyperhomocysteinemia was found in the present study, for physical activity, the P value for this association is quite borderline ($P=0.050$), and the direction of association estimates is consistent with previous studies^{51 52}, showing that higher level of physical activity is protective for high homocysteine. Similarly, the direction of association between alcohol drinking and risk of hyperhomocysteinemia is consistent with previous studies, but not reach the statistical significance^{53 54}.

51. e Silva Ade S, da Mota MP. Effects of physical activity and training programs on plasma homocysteine levels: a systematic review. *Amino Acids* 2014;46:1795-804.

52. Buckner SL, Loenneke JP, Loprinzi PD. Single and combined associations of accelerometer-assessed physical activity and muscle-strengthening activities on plasma homocysteine in a national sample. *Clin Physiol Funct Imaging* 2017; 37:669-74.

53. Coppola M, Mondola R. Correlation between plasma homocysteine levels and craving in alcohol dependent stabilized patients. *Clin Nutr* 2018; 37:1061-65.

54. Choi SH, Choi-Kwon S, Kim MS, et al. Poor nutrition and alcohol consumption are related to high serum homocysteine level at post-stroke. *Nutr Res Pract* 2015;9: 503-10.

Lines.245-246

The references mentioned only reported data from studies that found high prevalence of hyperhomocysteinemia. It is suggested review the references.

Response: Sorry for this mistake. We have revised it accordingly.

Reviewer# 2

Dr. Feyissa Challa, Ethiopian Public Health Institute, Ethiopian Public Health Institute

Comments to the Author:

The authors estimated the prevalence of hyperhomocysteinemia in China Population 4012. They observed a homocysteine concentration greater than 15 $\mu\text{mol} / \text{L}$ in 35.4 % of cases and an association with age, BMI and negative effect fruit and vegetable consumption. The work is based on a large population but lacks essential data, including creatinine concentration, folate, and vitamin B12 status. As a result, the public health impact of the data presented is limited.

specific comments

I suggest the title to be <>

Response: Done. We have revised the title accordingly as follows (Line 2-3).

Title: Prevalence and risk factors for hyperhomocysteinemia: a population based cross-sectional study from Hunan, China

abstract page #3, line 33.34, It says a total of 4012 males 1641, and female 2648...please correct it

Response: Done. We have revised it as follows (Line 34-35):

The median age is 55 (inter-quartile range:45-63) years, with 1644 males (41%) and 2368 females (59%).

Result in abstract page #3, line 39 The prevalence of hyperhomocysteinemia and is 35.4% (45.4% vs 28.5% for men...please delete "and"

Response: Thanks for this comment and sorry for this mistake. We have revised it accordingly.

abstract part..include the methodology used to measure Hcy.

Response: Done. We have revised it accordingly as follows (Line 36-37).

Homocysteine level were measured by the microplate enzyme immunoassay method.

Introduction

Not only MTHFR involved in Hcy but also Folate, Vitamin B12, Vitamin D, renal diseases, hypothyroidism ..etc

Response: Done. We have added more description about potential correlates of Hcy in the Discussion section as follows.

Line 74-75

Also, dietary nutrients intake (folate, Vitamin B12, and Vitamin D), or the presence of some disease (renal disease or hypothyroidism) are reportedly associated with hyperhomocysteinemia⁸⁻¹⁰.

8. Dhonukshe-Rutten RA, de Vries JH, de Bree A, et al. Dietary intake and status of folate and vitamin B12 and their association with homocysteine and cardiovascular disease in European populations. *Eur J Clin Nutr* 2009;63:18-30.

9. Ni J, Zhang L, Zhou T, et al. Association between the MTHFR C677T polymorphism, blood folate and vitamin B12 deficiency, and elevated serum total homocysteine in healthy individuals in Yunnan Province, China. *J Chin Med Assoc* 2017;80:147-53.

10. Jones P, Lucock M, Martin C, et al. Independent and Interactive Influences of Environmental

UVR, Vitamin D Levels, and Folate Variant MTHFD1-rs2236225 on Homocysteine Levels. *Nutrients* 2020;12
data analysis

Did the author checked for normal distribution data

Response: We have checked the normal distribution of our data. All variables were not with normal distribution, so we revised the description of the variables in Table 1 and the description. The general characteristic of participants was showed with medians and IQRs (inter-quartile ranges).

Results

How to explain that there are 59% of participants who are women?

Response: The authors thank the reviewer for his valuable comment. in the present study, 59% of the participants were women might be due to this is a community-based study, and our investigation were conducted in the daytime, and more women tend to be at home to take care of the children or do the housework, willing to take part in the study and men are more likely to be out for work. This is a limitation of our study; we have added this in the Discussion section as follows (Line 353- 355).

Because of the design of the study, more women participated in the present study than men (59% vs 41% for females and males, respectively), this is an inevitable limitation for this study.

Table 2 is better to put (homocysteine<15µmol/L) or (homocysteine>15µmol/L)..case and control don't define by HCY value only.

Response: We have revised the Table 2 and the corresponding description in the result section.

suggestion.

It will be better if we put both Univariate logistics and Multivariable logistic regression in one table In both Univariate logistics and Multivariable logistic regression, it will better to use reference those decreases HCY value..for example...those who consumed Fruit and vegetable as references, high Physical activity as references, etc

Response: We have revised the tables accordingly as follows (Table 1). We combined the univariable and multivariable logistic regression results in one table, for the reference problem, we revised them in all table according to the reviewer's suggestion.

table ...check all table value

Response: We have checked all the table values and revised the errors accordingly.

- page 18 line 234...Discussion: Despite what is indicated in the discussion and conclusion, this work is not the first to study the prevalence of hyperhomocysteinemia in China. Most of Hcy study was from China

Response: The authors thank the reviewer for his valuable comment as follows (Line 286-287).

Our study population is a representative study conducted in the adult population in China using a multistage random sampling method.

VERSION 2 – REVIEW

REVIEWER	Oliveira, Isabel Federal University of Pelotas (UFPel), Postgraduate Programme in Epidemiology
REVIEW RETURNED	08-Jul-2021

GENERAL COMMENTS	<p>I suggest a small review as listed bellow.</p> <p>Lines 75-76 The references 11 and 12 included are about specific association studies and they are not about underlying causes of plasma elevated homocysteine concentration.</p> <p>Lines 144-145 It is necessary define: Are at least 10 continuous minutes per week ? Or per day? Monthly?</p> <p>Line 163 ... homocysteine concentration...insteadconcentration</p> <p>Line 211/212 The sentence is unnecessary because hyperhomocysteinemia was already define at line 166.</p> <p>In Tables 1 to 3, review “didn’t consumed consumed fruit once or vegetable once every day.</p> <p>Line 297 “... 7% in southerner or southern provinces?</p> <p>Line 321 I do not agree with the word however and I suggested “In addition to”.</p>
-------------------------	---

REVIEWER	Challa, Feyissa Ethiopian Public Health Institute, National References Laboratory for Clinical Chemistry
REVIEW RETURNED	30-Jun-2021

GENERAL COMMENTS	The authors have responded to the reviewers' comments.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Dr. Isabel Oliveira, Federal University of Pelotas (UFPel)

Comments to the Author:

Dear author,

I suggest a small review as listed bellow.

Lines 75-76

The references 11 and 12 included are about specific association studies and they are not about underlying causes of plasma elevated homocysteine concentration.

Response: Many thanks for your careful review. We have replaced the references with a more suitable reference.

There are few studies looking at underlying causes of elevated homocysteine concentrations in the blood in humans¹¹.

11. Zaric BL, Obradovic M, Bajic V, et al. Homocysteine and Hyperhomocysteinemia. Curr Med Chem 2019;26:2948-61.

Lines 144-145

It is necessary define: Are at least 10 continuous minutes per week? Or per day? Monthly?

Response: Many thanks for your careful review. It was at least 10 continuous minutes per week. We now wrote (Line 144-146).

Physical activity lasting for at least 10 continuous minutes per week was categorized as valid physical activity, less than 10 continuous minutes per week as invalid physical activity³⁰.

Line 163

... homocysteine concentration...insteadconcentration

Response: We corrected the typo error. Many thanks again.

Line 211/212

The sentence is unnecessary because hyperhomocysteinemia was already define at line 166.

Response: Thanks for your comments. we have deleted this unnecessary sentence.

In Tables 1 to 3, review "didn't consumed consumed fruit once or vegetable once every day.

Response: Many thanks for your careful review. We altered the notes under Table 1-3 accordingly.

Line 297

"... 7% in southerner or southern provinces?"

Response: Many thanks for your careful review. We have revised it accordingly.

Line 321

I do not agree with the word however and I suggested "In addition to".

Response: Many thanks for your comment. We have revised it according to your suggestion.