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)	Coffee consumption and risk of prostate cancer: a systematic review and meta-analysis
AUTHORS	Chen, Xiaonan; Zhao, Yiqiao; Tao, Zijia; Wang, Kefeng

VERSION 1 – REVIEW

REVIEWER	Huan Liu Department of Urology, Shanghai Tenth People's Hospital, School of Medicine in Tongji University, Shanghai 200072, China
REVIEW RETURNED	08-May-2020

GENERAL COMMENTS	In this article titled"Coffee consumption is associated with a lower risk of prostate cancer: a systematic review and meta-analysis", the authors aimed to investigate the influence of coffee intake on prostate cancer risk. Though several meta-analyses have been conducted to explore the relationship, the results are still conflict. The present meta-analysis included several latest cohort-studies and provides new results. This paper is well organized and the tables and figures were presented appropriately to make it understood easily. I think the paper is publishable in the "BMJ Open" after some minor revisions: 1. The UK Biobank cohort which included 46 155 cases and 270 342 controls of White British also evaluated whether coffee intake is associated with: (i) overall risk of being diagnosed with/dying from any cancer; and (ii) risk of individual cancers (PMID: 31412118). It'll be more convincible if the present meta-analysis can include this cohort. 2. In line 111, the authors had contacted the authors of the article (Reference: 15) to request the relevant data, but didn't include this study (line 172), why? 3. It'll be better to help read this paper if the authors add the forest plot of subgroup analysis stratifying by cancer grade.

REVIEWER	Xilin Yang Tianjin Medical University, Tianjin, China
REVIEW RETURNED	20-May-2020

GENERAL COMMENTS	This systematic review and meta-analysis used cohort studies to estimate the association between coffee consumption and the risk of prostate cancer. My remit is the statistical aspects. The analysis needs to be improved or clearer and some concerns need to be addressed. Specifically, Comments: 1) it would be better to provide a search strategy for PubMed or other sources.

	2) given that follow-up time and age might make impacts on the estimation for the risk of prostate cancer occurrence, authors should provide their information for every included study, and potential bias in original articles also need to be presented in Supplementary Table S2. 3) the definition of "highest coffee consumption" and "lowest coffee consumption" might be different in each study, please make the definitions clear in Methods and provide the case numbers for each group in Supplementary Table S2. 4)in Table 1, significant associations were observed in some subgroup analysis with higher heterogeneous, especially in those stratified by adjustment confounders, however, most of analyses with lower heterogeneous did not find significant inverse associations. Authors should be more cautious when making conclusions. Moreover, explanatory notes for superscripts, e.g., 2 and c, need to be added. 5)Where is the plot of Egger's test? Comments: 1) it would be better to provide a search strategy for PubMed or other sources. 2) given that follow-up time might make an influence on the estimation for the risk of prostate cancer occurrence, authors should provide this information for every included study, and potential bias in original articles also need to be presented in Supplementary Table S2. 3) the definition of "highest coffee consumption" and "lowest coffee consumption" might be different in each study, please make the definitions clear in Methods and provide the case numbers for each group in Supplementary Table S2. 4)in Table 1, significant associations were observed in some subgroup analysis with higher heterogeneous, especially in those stratified by adjustment confounders, however, most of analyses with lower heterogeneous did not find significant inverse associations. Authors should be more cautious when making conclusions. Moreover, explanatory notes for superscripts, e.g., 2 and c, need to be added. 5)Where is the plot of Egger's test?

REVIEWER	Zumin Shi Qatar University
REVIEW RETURNED	15-Aug-2020

GENERAL COMMENTS	Overall, the study is well conducted. The statistical analyses are appropriate. There are several issues need to be addressed. 1. Line 151 and Line 33. Please provide the name of the user-written Stata program for the analyses. 2. Please provide the mean coffee consumption in each study. What was the distribution of coffee consumption in the dose-response meta-analysis? 3. One of the limitations is that most of the studies are from the USA and Europe. It is unknown whether the findings are relevant in other regions. The effect size is very small.

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer #1 's Comments:
In this article titled "Coffee consumption is associated with a lower risk of prostate cancer: a systematic review and meta-analysis", the authors aimed to investigate the influence of coffee intake on prostate cancer risk. Though several meta-analyses have been conducted to explore the relationship, the results are still conflict. The present meta-analysis included several latest cohort-studies and provides new results. This paper is well organized and the tables and figures were presented appropriately to make it understood easily. I think the paper is publishable in the "BMJ Open" after some minor revisions:

1. The UK Biobank cohort which included 46 155 cases and 270 342 controls of White British also evaluated whether coffee intake is associated with: (i) overall risk of being diagnosed with/dying from any cancer; and (ii) risk of individual cancers (PMID: 31412118). It'll be more convincible if the present meta-analysis can include this cohort.

Response: Thanks for bringing this to our attention. This is a case-control study nested within the UK biobank cohort. The study reported the odds ratio for the association between per 1 cup/day increase of coffee intake and risk of prostate cancer. As you suggested, we have added this study in the dose-response meta-analysis, and found that the updated result (RR: 0.988; 95% CI: 0.981, 0.995) did not change materially compared with the pervious one (RR: 0.989; 95% CI: 0.982, 0.997). Please see Page 10 and 11 for the corresponding updates on Results section of the revised manuscript.

2. In line 111, the authors had contacted the authors of the article (Reference: 15) to request the relevant data, but didn't include this study (line 172), why?

Response: Thanks for bringing this to our attention. After requesting the unpublished data from the authors, two authors (Russnes and Nilsson) provided us the relevant data for the dose-response meta-analysis. We have included Nisson et al study in the meta-analysis.

However, we did not include the Russnes et al study (Reference 20) in the meta-analysis, although we have obtained unpublished data from the authors. We excluded this study because the study population are the same with another included cohort study (Wilson et al study). These 2 studies both conducted in the US Health Professionals Follow-up Study. But the primary aims of these studies are different. Wilson et al study mainly focused on evaluating the association of coffee consumption and prostate cancer risk, however, Russnes study mainly focused on evaluating the associations of antioxidant intakes and prostate cancer risk. It only reported subgroup results on coffee intake as a major source of antioxidants. Thus, in current meta-analysis, we finally decided to include this Wilson et al study and excluded the Russnes et al study, because 1) the primary aim of the Wilson et al study was consistent with the current meta-analysis, 2) the Wilson et al study is more informative and transparent. For example, the study methods (e.g. coffee consumption measurement and assessment) and results (e.g. person-years, numbers of cases and HRs and 95%CIs for each coffee intake categories) are clearly presented in the Wilson et al study. All the source data needed in the meta-analysis have been published in this study, so that it may facilitate other researchers to better replicate the current meta-analysis.
As you suggested, we have elaborated on this issue in the revised manuscript. (Please see Page 7)

3. It'll be better to help read this paper if the authors add the forest plot of subgroup analysis stratifying by cancer grade.

Response: Thanks for the suggestions. We have added the forest plot with subgroup results by cancer grades in the revised manuscript. (Please see Figure 4)

Response to Reviewer #2 's Comments
This systematic review and meta-analysis used cohort studies to estimate the association between coffee consumption and the risk of prostate cancer. My remit is the statistical aspects. The analysis needs to be improved or clearer and some concerns need to be addressed. Specifically,
1) it would be better to provide a search strategy for PubMed or other sources.

Response: Thanks for the suggestions. We have provided the search strategy as a supplementary document in the revised manuscript. Please see Supplementary Document 3.

2) given that follow-up time and age might make impacts on the estimation for the risk of prostate cancer occurrence, authors should provide their information for every included study, and potential bias in original articles also need to be presented in Supplementary Table S2.

Response: Thanks for the suggestions. We have provided the age of the study population, follow-up information and potential bias in Supplementary Document 4 in the revised manuscript.

3) the definition of "highest coffee consumption" and "lowest coffee consumption" might be different in each study, please make the definitions clear in Methods and provide the case numbers for each group in Supplementary Table S2.

Response: Thanks for the suggestions. In the revised manuscript, we have elaborated on the definitions in the methods and provide the number of cases for each coffee intake categories (when available) in Supplementary Document. Please see Page 6 and supplementary document 4.

4) in Table 1, significant associations were observed in some subgroup analysis with higher heterogeneous, especially in those stratified by adjustment confounders, however, most of analyses with lower heterogeneous did not find significant inverse associations. Authors should be more cautious when making conclusions. Moreover, explanatory notes for superscripts, e.g., 2 and c, need to be added.

Response: Thanks for the constructive suggestions. We agree that the heterogeneous results of certain subgroups may limit the interpretation of the results. As you suggested, we have made conclusion with cautious in the revised manuscript. We have added the explanatory notes accordingly in Table 1. Please see Table 1, Page 3 and 18.

5) Where is the plot of Egger's test?

Response: As you suggested, we have provided plot of Egger's test as supplementary figure in the revised manuscript. Please see Supplementary Document 6.

Response to Reviewer #3 's Comments:
Overall, the study is well conducted. The statistical analyses are appropriate. There are several issues need to be addressed.
1. Line 151 and Line 33. Please provide the name of the user-written Stata program for the analyses.

Response: Thanks for the suggestions. Meta-analysis was performed with the "metan" and "glst" commands in Stata 14.0. As you suggested, we have provided the name of the Stata program in the revised manuscript. Please see Page 2 and 9.

2. Please provide the mean coffee consumption in each study. What was the distribution of coffee consumption in the dose-response meta-analysis?

Response: Thanks for the helpful suggestions. After we checked all the original studies, Sen et al 2019, Hashibe et al. 2015 and Discacciati et al. 2013 reported data on the median/mean value of coffee consumption. When mean coffee intakes per category were not presented in the publications, the mean intake in each intake category would be estimated based on the methods used in previous dose-response meta-analysis 1, 2, i.e. we estimated the mean coffee intake in each category by calculating the midpoint of the upper and lower boundaries. When the upper boundary of the highest intake category was not reported, we assumed that it had the same magnitude of intake as the preceding category.

In the dose-response meta-analysis, we examined a potential nonlinear relation between coffee consumption and prostate cancer risk by modeling coffee consumption using restricted cubic splines for nonlinear trends with 4 knots at fixed percentiles (5%, 35%, 65%, and 95%) of the distribution with the corresponding coffee consumption of 0, 0.5, 2.1 and 6.5 cups/day, suggesting a right skewed distribution of coffee consumption in this meta-analysis. There is no indication of non-linearity for the association (P=0.193 for non-linearity).

In the revised manuscript, we have provided the mean or median, highest category and lowest category of coffee consumption for each study, if the corresponding information were provided in the original studies. Please see Supplementary Document 4.

References:
1. Kennedy OJ, Roderick P, Buchanan R, et al. Coffee, including caffeinated and decaffeinated coffee, and the risk of hepatocellular carcinoma: a systematic review and dose-response meta-analysis. BMJ Open 2017;7(5):e013739.
2. Crippa A, Discacciati A, Larsson SC, et al. Coffee consumption and mortality from all causes, cardiovascular disease, and cancer: a dose-response meta-analysis. Am J Epidemiol 2014;180(8):763-75.

3. One of the limitations is that most of the studies are from the USA and Europe. It is unknown whether the findings are relevant in other regions. The effect size is very small.

Response: We agree that this is a potential limitation of this meta-analysis. Since the coffee intake and incidence of prostate cancer in the US and Europe are relatively high, most of the studies are conducted in these regions. Since the effect size is small, we should be cautious when generalizing the results to other areas, especially where the incidence is relatively low.

As you suggested, we have discussed this limitation and draw conclusions more cautious in the revised manuscript. (Please see Page 18)

VERSION 2 – REVIEW

REVIEWER	Xilin Yang Tianjin Medical University, China
REVIEW RETURNED	05-Oct-2020

GENERAL COMMENTS	No further comments.

REVIEWER	Zumin Shi Qatar University
REVIEW RETURNED	18-Oct-2020

GENERAL COMMENTS	My comments have been addressed adequately.