Peer Review Report

Review Report on Solid fuel use and the progression of multimorbidity in middle-aged Chinese: a prospective cohort study

---

EVALUATION

Q 1 **Please summarize the main findings of the study.**

This study examined the association of solid fuels use for cooking and heating with multimorbidity in a longitudinal study. The manuscript is relevant in the field given that existing studies have linked solid fuels use with single chronic conditions and less is known about associations with multiple chronic conditions.

Q 2 **Please highlight the limitations and strengths.**

The manuscript is relevant in the field given that existing studies have linked solid fuels use with single chronic conditions and less is known about associations with multiple chronic conditions. The manuscript could benefit from a thorough revision on writing. The strength of the justification of the study and the importance in the field can be strengthened in the Introduction and the Discussion. There are also important revisions to be made in the methods and the results. The way exposure to household air pollution was assessed has important limitations that need to be addressed and justified. This limitation might require re-structuring the analysis and the way the results are presented and discussed. The methodology needs to be revised to ensure reproducibility and clarity. The discussion is limited, and it will benefit from updated, additional references and the inclusion of important concepts to analyze (such as local conditions in terms of temperature, as well as other exposures and the cultural context that affects patterns of cooking and heating activities and ventilation).

Q 3 **Please provide your detailed review report to the authors. The editors prefer to receive your review structured in major and minor comments. Please consider in your review the methods (statistical methods valid and correctly applied (e.g. sample size, choice of test), is the study replicable based on the method description?), results, data interpretation and references. If there are any objective errors, or if the conclusions are not supported, you should detail your concerns.**

Abstract conclusions:
Consider editing these conclusions to mention the broader implications and how these results add to existing literature rather than just stating the association found.

Introduction:

The introduction could be enriched to explain the need for the type of assessment done in this work. This section could benefit from mentioning updated references that have looked or attempted to evaluate several outcomes in the context of solid fuel use. Is this really the first study that looks at several morbidity outcomes? There is indeed a need to evaluate multimorbidity in the context of solid fuel use and this is a perspective that will benefit the field. The authors could consider mentioning the work and estimations done by the Global Burden of Disease and how this manuscript and this methodology cold benefit such estimations to help prioritize policy and help inform the estimation of the impact of interventions.

Line 39 and 40: Revise statement for clarity. This last sentence of this paragraph seems out of place and is also not accurate. Each paragraph should focus on arguing one distinct idea and this sentence is jumping towards a topic that has not yet been introduced. Move or remove this sentence form this paragraph and if keeping it, please elaborate or specify that the authors probably mean this really in reference to multimorbidity and probably referring to morbidity form several chronic conditions as opposed to chronic conditions in isolation.

Line 42: revise and replace the generic and unprecise term "heart problem".

Methods:

The methods presented should be explained in concise manner but also should be consistent and have enough detail to be able to reproduce the work of the manuscript. The writing needs to be improved and this will certainly help in the readability and clarity of the methods section, however several concepts also need to be defined better for the methods to be clear and reproducible. The selection of the outcomes should be better justified, and the limitations related to these should be discussed. Ascertaining exposure seems overly simplified and I would expect more justification for the way exposure was classified in the models. Further analysis, that might be worth exploring, was also suggested to the authors to comment on or include.

It Is possible that the exposure can vary depending on the use of the solid fuel (i.e. heating vs cooking), however it might be much more relevant to understand the exposure overall regardless of the reason of use. I am much more concerned on understanding how exposure was determined and how do authors know that if the participants reported use of gas, that they did not use solid fuels at all or with some frequency or that they used some solid fuels in the past. It seems unclear to me why so much emphasis is placed on whether the reason for using the fuel was heating or cooking rather than the frequency and consistency of use of the solid vs clean fuels in the past.

Please provide more relevant local and cultural context that helps explain the use of solid fuels in the regions assessed. The need for the separation for the use of fuels for heating vs cooking needs to be explained better. It seems to me less relevant to separate the use of fuels the way it was explored in the sensitivity analyses described in Lines 103 to 108. Specifically, it seems less relevant to have the following two separate groups: "solid fuels for heating and clean fuels for cooking, solid fuels for cooking and clean fuels for heating". I wonder why the authors did not consider having levels of exposure separated by frequency of use of the solid fuel to get a better sense of overall exposure: For example something like: Solid fuel use very frequent: like used every day,…, to something like: Rarely uses solid fuels less than once a month, or only in the winter.

Please explain if fuel use and exposure information was also collected in every visit and how was it collected or verified.

Please make sure to explain and specify how exposure was classified and updated in the follow up visits and if past fuel use and past exposure behavior was assessed. This should be addressed if possible and discussed as a mayor limitation if not.

The scale of conditions could be relevant for a sensitivity analysis using binomial models and the scale of 9 possible combinations of number of conditions as presented in Table 1. Having more resolution on the outcome in a sensitivity analysis can strengthen the findings. Did the authors consider this approach at all?

Please define the outcome of memory problems and explain why was it included? I wonder if there are other cognitive function outcomes that might be better suited to capture what the authors were capturing here and if this might be a limitation of the data availability that should be discussed.

Please explain how the outcomes overall were selected and the discussion should include how these outcomes might have impacted the association. Several of these outcomes have not yet been associated with household air pollution.

Line 63, consider mentioning the covariates here. This is relevant since missing covariates could introduce biases to the final sample of participants and these types of issues should be included in the methods and discussed.

Figure1: covariate list and list of chronic conditions should also be included in this figure as a caption.

Line 70: Please define marsh gas. This term is less common in this type of literature.

Please make sure to define and clearly explain any terms or expressions the first time they are mentioned. For example, line 77: please define "progression of stable condition group".

The description of the Table 1 in the methods could be improved, as it is now it is hard to follow. It seems irrelevant to explain and lay out all nine possible combinations of situations if at the end only three categories of conditions are defined.

If you decide to keep Table 1 please include howe many participants fall in each category. Please also consider adding to this table the percentage of participants that used each type of fuel.

Make sure to define and justify why the covariates used in the methods are relevant. For example, explain why marital status was selected as one of the covariates; explain what household residence means and how is it defined (categorical?); What are the categories used for smoking status and drinking status.

Please be consistent, some covariates are defined, and some are not. All covariates should be clearly defined and explained.

I wonder how many homes use electric and natural gas and if it might be work doing some additional models were the electric or solar are the reference to see if the effect estimates get even stronger. Did the authors explore any of these further associations or trends depending on specific types of fuel?

Consider including tables (can be supplement tables) the frequency of use each type of fuel and frequencies of the outcomes. It would be interesting to see how the specific diseases vary and if certain diseases are more common than others among homes that use certain types of fuels.

Line 102: It is not clear what is meant by this sentence: "Solid fuels for cooking and heating were mutually adjusted in all models."

Results

Results can be presented more clearly. The tables need to be revised and additional results should be included in the main paper not as supplementary or sensitivity analysis. I suggest the authors to present as main models the combination of no solid fuel use (nether for heating nor cooking) and including the use of solid fuel use for both heating and cooking.

Tables 2 and 3. It is unclear what comparison results of the p-values presented are for. Please include captions and notes on the table to make sure all comparisons are explicit. It doesn't look like mean age is not different by type of fuel in any of the groups however the p-value is <0.01. Each comparison should also explicitly state the test used, use one symbol for t-tests and another one for the categorical comparisons.

Please revise the results presented in Table 2. It doesn't make sense to me that asthma prevalence is significantly different in the cooking fuels category with a p-value <0.01 and prevalence of 3.6% and 5.2% but not in the heating fuels category were the p-value is 0.1 and prevalence of 3.7% and 4.8%, and the number of cases seem clearly lower for clean fuel categories ~40 cases vs ~180 and 200 cases in the solid fuel categories.

Please define the generic diseases in Table 2 like digestive diseases and memory problems.

In Table 2 please include the category of no chronic conditions, even if its 0% of people.

In Table 2 it might also be worth including the categories of any chronic condition too and compare by fuel use group. Alternatively, consider including fuel use categories in Table 3.

Make sure to state the reference category for the comparisons made in Table 3.

Please include in the supplement information the outcomes recorder by visit and how these changed in time during follow-up. Please discuss any relevant economical events that might have affected the employment and mental health status of the participants during the follow up time.

I strongly suggest making an additional table like Table 2 that includes those that did not use any solid fuels for either cooking or heating and those that used it for both cooking and heating, and one additional category that is used solid fuels either for cooking OR heating. This should be revised and considered also based on the relevance of local and cultural patterns of solid fuels use.

I strongly suggest revising the way results are presented and present as main models the combination of no solid fuel use (nether for heating nor cooking) and including the use of solid fuel use for both heating and cooking. The other two additional categories already presented as sensitivity analysis (solid fuels in cooking or heating) can be added here too either separated or joined.

Line 123, 124: This statement is unclear and unexpected. Please make sure to clearly define the expression: "multimorbidity progression".

I strongly suggest: Table S3, last column, should be adapted and edited to be presented as the models in Table 4 are presented with unadjusted results and then fully adjusted for all confounders. This should be presented in the main body of the paper not as supplemental information.

Please make sure to explain the sensitivity analysis presented in Tables S1 and S2 in the methods.

Table S1: It is unclear what follow-up timepoints are included here.

Discussion

The discussion needs to be revised and edited to improve its content and the writing. The importance of addressing multiple comorbidities is essential an idea for this paper that has to be addressed in the discussion further. The analysis and discussion regarding the limitations of the evaluation of household air pollution needs mayor revisions. In addition, several additional important concepts must be incorporated in the discussion such variables that can affect the household air pollution exposure and biases in the way it was assessed; for example, altitude and season variability and how temperature might affect the use of solid fuels and ventilation in homes. Variability in solid fuel use over time must be addressed and discussed thoroughly in the paper as a whole.

The discussion needs to connect the objective introduced regarding the importance of exploring associations with several outcomes instead of one outcome at a time. This is an essential contribution of this paper, and it should be connected in the discussion with the importance of this methodology and how this can help better inform other studies and interventions and to better estimate the burden of this activity in both morbidity and mortality. These types of diseases never occur in isolation and usually individuals that develop a disease due to an exposure usually develop several associated co-morbidities. This should be further discussed and updated references to support the importance of this idea should be included in the discussion. There is a big potential of including strong arguments towards this idea that have not been included.

Please discuss the potential implications for estimations of the global burden of disease of household air pollution and on policy of these findings and how can the results be used to keep advancing in the field. Include updated references of the estimations o the impact of household air pollution on the global burden of disease. What are next steps regarding better estimations of co-morbidities and how to better use this methodology to inform interventions and priorities? How would the next models be better with more information? The discussion of next steps is limited please consider going beyond the idea that pathways need to be better understood.

Again, it is not clear to me what is the importance of looking at cooking and heating separately, what is the benefit of this? This should be clearer from the Introduction and then in the Methods and discussed more clearly in the Discussion. Other important concepts that are overlooked seem much more important to address such as temperature (altitude and latitude), urban vs rural, frequency of the use of solid fuels.

Why was cancer not one of the outcomes included? Please consider: if the selection of outcomes was a convenient one since it was a secondary analysis using a parent study and therefore it should be discussed as part of the limitations and need for future work.

Line 190: fuel use assessed only on baseline: This seems to be the first time this is mentioned. Please make sure to be explicit about this in the methodology. This is a huge limitation, and the authors should attempt to address this in the paper. Why is this ok to do for this population? Why is it ok to assume fuel use will not change in 8 years on both urban and rural populations? The authors should address this with further sensitivity analysis within types of populations and should include information that supports how fuel use might change in time among the types of rural or urban populations assessed. A good clear argument on why this information was collected should also be presented early on in the methods section. What other fuel use variables collected can be used to assess exposure of smoke from solid fuel use that can be included in the analysis?

Without a proper justification of the lack of fuel use data longitudinally the authors should consider limiting the paper to a cross-sectional analysis or a lagged analysis based on how much time the selected outcomes can take to develop. The focus of the paper cannot be a longitudinal focus if the exposures were not collected longitudinally.

Other potential exposures not mentioned should be included in the discussions such as urban traffic related air pollution.

The authors might want to consider a sensitivity analysis within urban and rural subgroups. The lifestyle of these two groups in terms of how other exposures might have affected them should addressed. This is not even considered as a cofounder in the models, but it is a major factor. Job title, proximity to major roads, industries, urban traffic exposures are going to impact urban and rural homes differently. Urban or rural should be considered as a confounder and addressed in the methodology.

Please explain why the authors not included season patterns in the results or the discussion since it is relevant for exposures due to heating. What are seasonal, altitude, and temperature variations in the areas of the study and how does this variability differ between regions. Some regions might have more extreme weather changes than other regions.

Please discuss the relevance of looking at fuel use for cooking and heating separately from an intervention standpoint.

Line 156-157: state that rate of solid fuel use here.

Line 164, 165: Please mention and discuss in the text the location of this other study. Cultural and local differences might also explain these differences.

Line 167, mention that rate here, how much higher weas this rate compared to this study?

Lines 168 to 173, these are valid and important points, hoverer the discussion has to be developed better. It is not clear how these factors might explain the differences with the study referenced in the sentence above. Please revise.

Additional comments:

Some further editing and revision of the writing is required. A few select examples of sentences that require editing and revising are listed below:

Overall: consider revising the term "solid fuels use" to either "solid fuel use" or "use of solid fuels".

Abstract. Results: revise "compared with the clean fuels use". Consider either "compared with clean fuel use" or "compared with the use of clean fuels". Revise this type of issue throughout the manuscript consistently.

Contribution to the field:
Please revise: "used large population samples." Corroborate if the use of the plural form of "sample" is necessary.
"Findings from our study added the evidence for reduction of solid fuels use and …". Please revise the expression "added the evidence for reduction of solid fuel use".

Introduction line 31: "related to the risen burden". Revise "risen", consider the use of "rise".

Line 37: "are two different patterns" Revise "are" consider "have".

Line 38: consider replacing the word "being" with something like "taking place" or "occurring".

Line 44: specify if this multimorbidity percentage refers to prevalence.

Line 46: revise: "Challenging the public health system". This is generic and unprecise, please edit to be clear on what the authors mean to say here. These types of writing issues are very common throughout the manuscript. Please revise in general for these types of issues.

Line 95, please state explicitly here what variables were compared among groups, all covariates?

Lines 177, 187 and 179. Include references for this sentence and mention what conditions have been found to be associated with solid fuel use.

Line 180: This is inaccurate since this study did not address pathways, please revise and be more precise.

PLEASE COMMENT

Q 4   **Is the title appropriate, concise, attractive?**

The way exposure to household air pollution was assessed might require re-structuring the analysis and the title.

Q 5   **Are the keywords appropriate?**

yes, but they magic need to be revised based on revised methods suggested.

Q 6   **Is the English language of sufficient quality?**

No. The writing needs to be revised.

Q 7 **Is the quality of the figures and tables satisfactory?**

No.

Q 8 **Does the reference list cover the relevant literature adequately and in an unbiased manner?)**

It needs to be revised and updated.

QUALITY ASSESSMENT

Q 9 **Originality** — 5/5

Q 10 **Rigor** — 2/5

Q 11 **Significance to the field** — 5/5

Q 12 **Interest to a general audience** — 4/5

Q 13 **Quality of the writing** — 1/5

Q 14 **Overall scientific quality of the study** — 1/5

REVISION LEVEL

Q 15 **Please make a recommendation based on your comments:**

Major revisions.