Peer Review Report

Review Report on Sex associations between air pollution and estimated atherosclerotic cardiovascular disease risk determination

Original Article, Int J Public Health

Reviewer: Pablo Knobel
Submitted on: 22 Jul 2023
Article DOI: 10.3389/ijph.2023.1606328

EVALUATION

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

The authors found that PM2.5 and air pollution core were associated with ASCVD and that there are differences between sexes.

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

Strengths:
+ Very large sample size with comprehensive health and behavior data.

Limitations:
+ Comboluted and outdated (in my opinion) methods.
+ The writing needs substantial improvement.

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.**

---Main comments:
+ The writing is quite difficult to follow which makes some of the methods difficult to evaluate. Additionally, some of the methods used are not properly justified or even presented in the methods section.
+ The authors focus the results on an effect modification by sex (which can be interesting), but need to introduce it better and still present the overall results.
+ I have major concerns about how the exposure was calculated and why all exposures were included in the model at the same time despite the high collinearity.
---Major Comments:
+Introduction:
Line 65: the second and third sentences of this paragraph are a little bit convoluted.
Line 75: I'm not sure if it's still overlooked. If the author wants to maintain this statement I would recommend using a more recent reference.
The introduction is a bit all over the place. Topics are introduced that might not be used later in the manuscript, like indoor air pollution.
+Methods:
Line 98: I'm not sure I'm following the numbers here. The UK Biobank enrolled 502,478 individuals between 2006 and 2010. This can't be 5.5% of the total cohort. Then where do those 90 million individuals come from? That's more than the UK population if I remember correctly. What is the follow-up time?
Line 101: if possible, would it be possible to know what where the reasons that precluded linking in some cases? What was the link to the national health registries used for?
Line 118: again, the number are not adding up. 399,067 participants without missing or CV. Then 114,022 were excluded due to missingness.

Authors should separate between excluded for missingness and for previous CVD. If the missings are not random I would worry about selection bias – the authors should consider sensitivity analysis or, at least, mention it in the limitations section.
Line 132: some of these parameters can be associated to air pollution exposure.
Line 138: The exposure measure and estimation need to be explained more clearly. The measures were made in 2010 and are estimated for the rest of the years.
Line 144: baseline visit address can be problematic if participants move through the follow-up. This should be mentioned in the limitations.
Line 152: This needs to be better explained and probably moved to statistical analysis. For example, if the authors use b estimates, we need to know the model used and the covariates included.
Line 157: these numbers belong in the results.
Line 176: could the education levels be clarified into international references? For example, "high school".
Line 209: why are the authors using both continuous and quantiles of exposures?
Line 214: the source of all covariates needs to be described in the manuscript.
Line 215: what are these sex models? I understand that the authors want to investigate potential effect modification by sex, but this needs to be introduced better.
Line 218: why were the pollutants considered together in the model and not using single pollutant models?
Line 229: Authors should report the overall sample characteristics and then, if interested, point out between-group discrepancies. This section, in general, needs to be heavily reworked for consistency and clarity. A few things (1) Reporting only differences and not overall characteristics [and not introducing the differences between males and females beforehand]; (2) I would personally move away from reporting all the p-values; (3) some cases – like education level – authors give the actual values and p-values and other –like PM2.5 – only the p-value; (4) I'm still unsure how the air pollution score was calculated, but I'm surprised it was presented in table 1; (5) Table 2 is out of place, underexplained, and not introduced in methods.
Line 243: Authors report different models with different adjustment sets but the results remain fairly similar. I would recommend keeping only one adjustment set and reporting it.
Line 247: these exposures are highly correlated, and I do not think this can be used in the same model. Authors should report single pollutant models and, if possible, use mixture methods (like BKMR, WQS, or qGcomp) – I would recommend WQS for this case.
Also: Air pollution studies usually measure exposures in 1 IQR change or 10μg/m^3 to obtain more interpretable betas. Authors should consider changing it.
Line 248: how was the air pollution score calculated? Is this a regression after a regression? Why did the authors decide to use this?
Line 276: the comparison between sexes and urban/rural is engaging. Other covariates could have also been explored: such as income or race. I think the overall associations should be reported before getting into the effect modification.
Line 294 to 321: while interesting, these studies are not being linked to the study results. If they can't be directly linked I would remove it.
Line 337: this could be partially introduced to methods to introduce the air pollution risk score better.

PLEASE COMMENT

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

No.

The title mentions the middle-aged population but goes far into the senior population.

Q 5 **Are the keywords appropriate?**

Yes

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

No

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

Yes.

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

Yes

QUALITY ASSESSMENT

Q 9 Originality

Q 10 Rigor

Q 11 Significance to the field

Q 12 Interest to a general audience

Q 13 Quality of the writing

Q 14 Overall scientific quality of the study

REVISION LEVEL

Q 15 Please make a recommendation based on your comments:

Major revisions.