Reviewer's report

Title: Using three statistical methods to analyze the association between exposure to 9 compounds and obesity in children and adolescents: National Health and Nutrition Examination Survey 2005-2010

Version: 0 Date: 16 Mar 2020

Reviewer: Arce Domingo-Relloso

Reviewer's report:

This study evaluated the association of 13 environmental exposures with BMI and obesity. The authors found associations between 2,5-dichlorophenol, monoethyl phthalate and mono-isobutyl phthalate exposure and obesity. This is a relevant topic given the increasing burden of obesity worldwide and the prevalence among children and adolescents. This work is also relevant given that environmental chemicals such as phthalates or pesticides have traditionally been studied individually. It includes novel statistical methods that are becoming popular in the environmental epidemiology field to address chemical mixtures rather than individual chemicals and explore the joint effects. The manuscript is rich in content and includes state-of-the-art methods, as well as results that could be relevant for public health. The authors also provide the data used for the study, which is helpful for replication. However, from my perspective, this work needs to address important issues in order to be suitable for publication. For instance, the manuscript could be better organized and needs further details in some areas. In addition, I am particularly concerned about the model specification for BKMR. Please see detailed comments below.

Major comments:

1. Introduction: Please make clear which environmental exposures you are studying in the introduction. E.g: "We studied 13 chemical exposures including phenols (BPA, BP-3), parabens (BuP, EtP, MeP, PrP), pesticides (2,5-DCP, 2,4-DCP, 2,4,5-TCP, 2,3,6-TCP) and phthalate metabolites (MBzP, MEP, MiBP)."

2. Methods, page 6, line 15: Does NHANES have other adiposity measures that could be compared to BMI, for instance, body fat?

3. Methods, pages 7 and 8: Please include references that develop in a more detailed way the Weighted Quantile Sum regression and the BKMR (both Gaussian and probit) methods.

4. Methods, page 9, line 4: 1000 iterations seems like a small number for probit BKMR, which tends to converge much slower. You don't talk about model diagnosis in the whole text, did you look at the traceplots? The traceplots should be showing a more or less homogeneously covered space to ensure the convergence is good. You can find how to get traceplots here: https://jenfb.github.io/bkmr/overview.html.
5. Results, page 10, line 8: In the correlations figure (figure 1), please put the numbers (the correlations) in each square for each pair of chemicals. I would consider \( r = -0.06 \) as no correlation rather than a negative correlation. In addition, I would consider 0.81 a high correlation rather than a moderate one. Please use the word "correlation" rather than "association" consistently in that paragraph.

6. Results, page 11, line 7: Why are you considering BuP, EtP, 2,4,5-TCP and 2,4,6-DCP in a different way in Supplementary Tables S1 and S2? Why do you have two references (quartile 1 and quartile 2) for those chemicals? Unless there is a specific reason for doing this (which you should explain both in the text and in the Tables' footnotes), please consider all chemicals to have the same reference (quartile 1).

7. Results, page 13, line 52: "In the phthalate metabolites group, MEP drove the main effect of the whole group (CondPIP: 0.403)." I would not consider that as driving the effect of the whole group as the other two chemicals have PIPs 0.312 and 0.285, which are very similar.

8. Results, page 13, line 59: In Figure 3 a and b, the credibility intervals are huge. Again, I am concerned about the model convergence. BKMR is a very sensitive model to initial values and number of iterations, especially in the probit version. It seems like your model did not converge properly. If you run the model with more iterations or change the tuning parameters, you might get tighter credibility intervals and more reliable PIPs. Also, you should check the acceptance rates of the tuning parameters (you can see that by setting verbose=TRUE in the function call). The r.jump2 parameter is particularly important. Ideally, \( r/delta2 \) should be between 20-45 \%. If the values are too high, you should try higher values of r.jump2 in the function call to see if it moves closer to the target window. This will (generally) lead to a faster convergence.

9. Results, page 13, line 59: I don't think we can talk about a positive tendency having such wide credibility intervals.

10. Results, page 14, line 2: In the pesticides (for obesity) and phthalates (for BMI) groups, it seems like one of the chemicals drives the whole effect. It would be helpful as a sensitivity analysis to introduce those chemicals as single exposures (disregarding the others from the phthalates and pesticides groups) in common with the group of phenols and the group of parabens in a model and see what happens with the association of the mixture. If the credibility intervals tighten, it might mean that the other exposures in the pesticides and phthalates groups are not relevant for the outcome.

Minor comments:

11. Line numbering: Please next time make sure you enable continuous line numbering through the whole manuscript (without starting over in every page). It is easier for reviewers and editor.

12. Abstract, page 1, line 39: "too much environmental endocrine disruption may increase the occurrence of obesity". Delete "too much".

14. Abstract, page 2, line 18: You are missing a $P = (P = 0.001)$.

15. Abstract, page 3, line 1: "2,5-DCP, MEP, and MiBP may play an important role in the 13 exposures." They play an important role in obesity / BMI, not in the 13 exposures.

16. Introduction, page 3, line 22: "Obesity is a pandemic disease that cannot be ignored". I would delete this sentence. I don't think we can say that obesity is nowadays ignored by physicians and scientists (as could be other diseases or health disorders).

17. Introduction, page 3, line 37: "play an essential part". I would rather say "an essential role".

18. Introduction: Please justify the reason why you chose those three statistical methods. There are many more that could be used for this purpose.

19. Methods, page 4, line 59: Please insert a reference about NHANES, especially about the multi-stage stratification probability sampling design, which people might not be familiar with.

20. Methods, page 7, line 11: Please use the word "multivariable" instead of "multivariate" through the text. While multivariable means you have multiple predictors, multivariate means you have multiple outcomes (See Hidalgo et al 2013: Multivariate or Multivariable Regression? (American Journal of Public Health)).

21. Methods, page 8, line 34: "healthy outcomes". Delete the "y".

22. Methods, page 8, line 56: "We grouped the chemical exposures into three groups, according to the resource and correlation with each other". Change the word "resource" for "source". Please describe the groups and the correlations between the chemicals in each group (you list the chemicals in each group in the tables but not in the main text).

23. Results, page 9, line 27: Add the % of obese people to Table 1 (in addition to the number).

24. Results, page 9, line 40: You don't need descriptives of the z-score of the BMI, it doesn't add any valuable information beyond the untransformed BMI descriptives.

25. Results, page 10, line 1: "geometric transformation". Do you mean "geometric mean"?

26. Results, page 10, line 26: I would change the heading "Total" in table 3 for "Continuous" or explain that total means continuous chemical variable in the table footnote.

27. Results, page 10, line 31: "...for children and adolescents are shown in Table 3 and 4, respectively". This looks like you have done separate analyses for children and adolescents. Please make clear that the distinction is between BMI and obesity.
28. Results, page 11, line 16: Change the word "variate" for "variable".

29. Results, page 12, line 1: Given that changes can be observed from model 1 to model 2 (especially for BMI z-score), you might want to fit three models instead of two: model 1 as you are considering it, model 2 additionally adjusting for cotinine and caloric intake, and model 3 additionally adjusting for socioeconomic status (SES) variables: education levels and family income-to-poverty ratio, to see if smoking / caloric intake are the ones that make a difference or, conversely, SES is more related to obesity. At the very least, it would be important to mention in the discussion what variable is mainly responsible for the change in the associations.

30. Results, page 12, line 15: "1.61 (1.25 ~ 2.07)", change the ~ sign for ",".

31. Results, page 12, line 21: Figures should be self-explanatory. Please put a title in both figures 2a) and 2b) and an explanatory footnote.

32. Results, page 14, line 6: "EtP, MeP, PrP, 2,4-DCP and MBzP demonstrate negative association with obesity": the curves for EtP, PrP and 2,4-DCP are flat, I would not consider that a statistically significant association. Also, it is more correct to say "inverse" association rather than "negative".

33. Results, page 14, line 27: "They revealed a positive association of the mixed exposures with the BMI z-score". You need to make clear that there is no statistical evidence (the credibility intervals are too wide).

34. Discussion, page 14, line 44: "Due to the interactions between chemicals…" Rather say correlations, we don't know about interactions, in fact, you don't find any interactions in your study. The thing we are sure about is that some of them are highly correlated, not about interactions.

35. Discussion, page 14, line 48: "However, we still need to comprehensively consider the results of these three methods". I would delete this sentence. It adds nothing.

Level of interest
Please indicate how interesting you found the manuscript:

An article of importance in its field

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:
1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal