Reviewer’s report

Title: In-utero exposure to phenols and phthalates and the intelligence quotient of boys at 5 years

Version: 0 Date: 22 Sep 2017

Reviewer: Joseph Braun

Reviewer's report:

General comments

The authors present a prospective study that investigate the association of maternal urinary phenol and phthalate metabolite concentrations with IQ among boys ages 5-6 years. The results did not suggest an inverse association of in-utero phenols and phthalates exposure with children's IQ. This study has a relatively large sample size. This manuscript is well written and organized, and have contributed to the literature in this research area. However, there are two major concerns. One concern is that one single urine measurement was used to assess exposure, which may introduce measurement errors. The other concern is that the outcome (IQ scores) was not age-standardized, and full-scale IQ was not calculated. In addition, some other minor-moderate issues should be addressed before the manuscript is acceptable for publication.

Specific comments

Major comments:

1. Page 3-4, the background section is missing some important and more recent citations. For example, Page 3, line 98-99, please cite recent work of Stacy et al., Environ Int, 2017 and Braun et al., EHP, 2017. The introduction is also too long.

2. Page 6, line 165-172, 1) Was FSIQ calculated? I believe that it should be and the results presented. 2) Are these age-standardized scores? If not, why not using age-standardized scores? Even in the narrow age range of 5-6 years, the level of cognitive ability represented by the raw scores can be confounded by age. In examining the scoring manuals, it appears that as much as a 2 point difference in scaled component scores could be present. 3) Were these sample-specific ranges of raw scores? If not, then these are not the correct ranges. Vocabulary scores range from 0-43, and block design ranges from 0-40.

3. Page 8, line 229-231, Related to the comments on the IQ test. Why not just use the Full-scale IQ, performance IQ, and verbal IQ? It seems like a lot of effort went into constructing the PIQ and VIQ whereas the scores can easily be calculated from the WPPSI manual.
4. Page 8, line 246-251, while I appreciate the effort to reduce comparisons, but I strongly recommend that you show the traditional analysis, not just the 'fancy' method. If you sum DEHP and the 4 parabens, which are highly correlated, then you only have 14 exposures and 3 outcomes, which is 52 comparisons.

5. Maternal urine was collected between 22 and 29 weeks of pregnancy, which may not be representative of the level of phenol and phthalate metabolite concentrations at other time point of pregnancy. Page 17, line 488, one single urine measurement was used to assess exposure, therefore specific windows of vulnerability cannot be assessed. While the authors did a commendable job at addressing the potential for exposure misclassification, it would be helpful to discuss the assumptions of their disattenuation method.

6. Child's race/ethnicity is associated with IQ, and may be associated with the exposure. The authors should consider adjusting child's race/ethnicity in the model.

7. Phenol and phthalate metabolite concentrations may need to reach a certain level to affect cognitive function. The authors should also consider examining the associations of maternal urinary phenol and phthalate metabolite concentrations with IQ using splines, or quantiles.

Minor comments:

1. Page 2, line 66, the authors should avoid the "our study is the first…” statements (also Page 17, line 501-507).

2. Page 3, line 75, the authors should drop the word "industrial".

3. Page 3, line 100, "focused on behavior", the authors should add a citation for this statement. There are plenty of review articles of this literature.

4. Page 3, line 86, the authors should cite some seminal work examining phthalates and androgens in rodents (i.e., Howdeshell et al.).

5. Page 3, line 87, Citation #11 only deals with thyroid hormones, but not gonadal hormones. I would suggest citing Cohen-Bendahan et al. Neurosci Biobehav Rev 2005.

6. Page 3, line 91, "focused on intelligence" was not appropriate. The domains assessed by Verbal IQ and Performance IQ encompass many of the domains cited in the prior sentence.

8. Page 4, line 117, please change "BSDI" to "BSID".

9. Page 5, line 137, I suggest changing the word "population" to "participants". Samples of participants are drawn from populations.

10. Page 5, line 162-164, it does not appear that this is an official version of the WPPSI-III. Can you please provide more details about this test? How was the translation done? Was back-translation performed? Was there other validation done on the French version?

11. Page 6, line 175-176, I suggest rewording to say that you measured concentrations of . . . in maternal urine samples . . .

12. Page 7, line 199, "confounders were identified a priori from literature" What was the standard? If by Directed Acyclic Graph (DAG), please show the DAG.

13. Page 8, line 231-233, were these informed by an exploratory analysis or set a priori?

14. Page 8, line 249-251, Why? The outcomes and exposures are not independent of each other. See Rothman et al., No Need to Adjust for Multiple Comparisons, Epidemiology, 1991.

15. Page 9, line 259-266, how were the 95% CIs calculated?

16. Page 10, line 290, how did those followed-up compare to the original cohort in characteristics?

17. Page 14, line 418, lowest dose in Xu et al. study was 50 ug/kg/d and this is nowhere close to estimated human exposure of 10-100s ng/kg/d (see work of Lakind and other for these estimates).

18. Page 15, line 443-447, The order seems misleading. The authors should put the association among all children first, and then talk about the results in boys, and compare your results to the association in boys.

**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.