Author’s response to reviews

Title: A Repeated Measures Study of Phenol, Paraben and Triclocarban Urinary Biomarkers and Circulating Maternal Hormones during Gestation in the Puerto Rico PROTECT Cohort

Authors:

Amira Aker (aaker@umich.edu)
Kelly Ferguson (kelly.ferguson2@nih.gov)
Zaira Rosario (zaira.rosario@upr.edu)
Bhramar Mukherjee (bhramar@umich.edu)
Akram Alshawabkeh (a.alshawabkeh@neu.edu)
Antonia Calafat (aic7@cdc.gov)
Jose Cordero (jcordero@uga.edu)
John Meeker (meekerj@umich.edu)

Version: 1 Date: 27 Dec 2018

Author’s response to reviews:

Addressing reviewer comments to the paper entitled “A Repeated Measures Study of Phenol, Paraben and Triclocarban Urinary Biomarkers and Circulating Maternal Hormones during Gestation in the Puerto Rico PROTECT Cohort” by Aker et al., 2018

Submitted to: Environmental Health

Reviewer #1: Summary: This study seeks to evaluate the association of pregnancy urinary phenol concentrations in relation to maternal serum hormone concentrations at 2 times points (16-20 and 20-24 weeks gestation) in the Puerto Rico Birth PROTECT Cohort. This study is an update on a previously published manuscript from the same authors in 2016. The present work uses a more complete cohort and includes additional hormone (estriol, testosterone, and total T3 and T4). Although this study is quite comprehensive, it also reads like a list of exposures and outcomes with little cohesiveness or structure. There are several inconsistencies as described in the discussion and it is unclear how to interpret the present results in relation to the previous work. The results and discussion are also unnecessarily lengthy and need to be streamlined and better synthesized. I would also recommend a table in the appendix comparing the first paper...
results with the second and a column with interpretation notes. Overall, I found this paper difficult to follow. I think it requires a more synthesized results and discussion section, focusing on the most pertinent results rather than listing each individual finding one by one. It is also unclear which among the results are most interesting/relevant and why.

Response: Thank you for your comments. We agree that the article was difficult to follow, and have updated the Results and Discussion section to make it easier to follow. Our original aim was to provide as much background as possible to help put the results in context, but have now changed the focus to the most interesting results, as per your suggestion. We hope this new streamlined version sufficiently addresses your concerns. We also included a new table in the Appendix that includes a summary of the new and old results for the common exposure biomarkers and hormones in the two papers.

Other points to consider:

- I don't readily see the need to include both the copious results in the tables as well as the figures. I would recommend adding the tables to the appendix and using the figures to display the main results.

Response: We agree that the tables and figures are unnecessary, and have moved the tables to the Appendix.

- I cannot easily discern why you need to conduct two separate analyses LMM and MLR. Why not focus your analysis and results on one particular method that is clearly justified based on your research question? What is the added benefit of reporting both? Are they answering different questions? What question do the results from LMM "Regressing hormones vs. exposure biomarkers" answer? Perhaps I am missing something but I think it makes for an unnecessarily exhaustive manuscript that is difficult to interpret. Please also indicate in the footnotes of the tables/figures what covariates were included in these models. Are these the results of the adjusted analysis? Please indicate this as well in the title.

Response: The linear mixed models are stronger models because they account for the intra-individual correlation of serial hormone measurements collected over the two study visits, and
provide more power to detect smaller effects. Therefore, they help us find associations that may be occurring over the entirety of pregnancy that wouldn’t be observed otherwise in cross-sectional studies. The multiple linear regression models, on other hand, allow us to look for windows of susceptibility, albeit with less power than that provided by the mixed models.

While we realize this may make the manuscript a little more difficult to interpret with the extra models, we believe that this analysis adds to the growing literature on the effects of these exposure biomarkers, and is nuanced in that it includes stronger models that help increase statistical power to capture smaller effects, as well as an analysis that looks specifically at each of the two time points.

We included the covariates in the footnote of the tables and changed the titles to “Adjusted models/results”, as suggested.

- Inclusion of covariates in your models are based on >10% change. However, it is preferable to use an a priori selection based on knowledges of exposure - outcome confounders known in the literature. Please consider basing your covariate selection to be more sensitive to potential collider stratification. Furthermore, it is unclear why you add specific gravity to the models when the concentrations have already been adjusted for urinary SG? Please clarify the discrepancy in described methods in the paper (e.g., lines 22-29 p. 8 vs. lines 49-51 p. 6) and justify this decision. Lastly, nowhere in the paper have you described your covariates and how they were measured. Typically, manuscripts should have a section for exposure and outcome assessment as well as 'covariate assessment' but this has been omitted here. Please describe how data on covariates were collected and operationalized.

Response: We agree that an a priori method is a preferred method, and this was implemented before adding the covariates to the models. Select covariates (included in Table 1) were considered to be potential confounders from the existing literature, and the 10% method was used for those selected potential confounders in order to strengthen our models. This is now described in greater detail in the Methods section.

The only covariate that could potentially be a collider was BMI; however this covariate was not included in our final models.

With regards to the covariates, we appreciate the suggestion and the revised Methods section describes that the covariate data was collected during study visit 1 via questionnaire. Table 1 breaks down the demographics of the study population, and provides percentages. We also updated the table to include the differences between the study population and the excluded
population. Additionally, the first paragraph of the revised Results section now describes these results.

The exposure biomarker concentrations in the primary models were not corrected for specific gravity. They were only corrected for specific gravity for descriptive data analyses. This is now clarified further in the Methods section to fix the confusion.

- I would de-emphasize the focus of the introduction on personal care products (PCP). Some of the chemicals examined can have multiple sources and PCP are only one of many others. Please consider reframing and strengthening the introduction. At the moment is seems quite cursory and sophomoric. Please be more targeted in these chemicals in relation to hormones in pregnancy? What do we know already? What do we NOT know? And, what gap are you trying to address and how? State your primary and secondary objectives.

Response: Thank you for your comment. As suggested, we re-wrote the Background section to be more targeted and hypothesis-driven.


This manuscript examines gestational phenol and paraben concentrations in relation to reproductive and thyroid hormones in 602 pregnant women in the PROTECT cohort. The authors observed both positive and negative associations between many phenols and paraben with hormone levels measured between the 2nd and 3rd trimesters. The authors implemented a linear-mixed model to account for intra-individual correlation of hormone measurements, and also used a linear-mixed model with biomarker concentrations as a time-varying variable to assess critical periods of paraben and phenol exposures. Although there is evidence of large day-to-day variability, methods implemented to reduce exposure misclassification bias make this a strong addition to the literature on the potential effects of paraben and phenol exposure on maternal hormone levels during pregnancy. I have a mostly minor comments below that I hope will help improve further iterations of the manuscript.
Methods

2.1 Study participants

- Which samples were used for analyses should be clearer; i.e. although there were 3 study visits, you only used information from the 1st and 3rd (the assumption is because these were the visits that blood was collected as well)? Or were urinary biomarkers from the 2nd visit used also? Page 5: lines 44-53

Response: Thank you for your comment, as per your assumption, we only used information from the 1st and 3rd visits because we did not have serum to analyze for hormones during the 2nd visit. To reduce confusion, we removed references to the visits, and instead referred to the gestational weeks 16-20 and 24-28.

- the number of participants from each visit should be stated. Page 4: line 49

Response: The number of participants is included in Table 2. We hope this is sufficient in addressing your comment.

- what is the total number of participants in the cohort? (not just the total included in analyses). Page 4: line 56

Response: We included the total number of participants in the Methods section as well as Table 1, as suggested.

- where the differences due to loss to follow-up or missing measurements/covariates?

Response: It was a combination of the two. In some cases, women were lost to follow up, and in other cases, the urine and/or blood samples were not yet analyzed at the time of analysis and write up.
2.4 Statistical Analyses

- demographic characteristics should be listed and how they were included in analyses (continuous vs. categorical). Page 6: line 41

Response: As suggested, we included how the covariates were included in the models in the Methods section (page 9).

- Although the authors explained why they used a different socio-economic index for certain models, did they perform sensitivity analyses using both? It is hard to justify using different covariates in regression models and then comparing the results. Page 7: lines 36-48.

Response: Due to collinearity issues, we could not include both covariates in the models. We agree that using a different socio-economic index for two of the hormones is a little unusual; however, we were interested in strengthening our models to better capture the “real” association between the exposure biomarker and hormone, and felt that a small change between the models is justified given that estriol and progesterone were correlated with maternal education, and less so with insurance type.

Tables

Table 1:

- table should be comparing the demographic characteristics of those included and excluded (full cohort) in analyses.

Response: Table 1 was updated as suggested.
- demographics should better reflect regression analyses by having the characteristics for each visit in addition to the full analyses cohort (see above comment).

Response: As mentioned in the Methods section, the demographic details were collected via questionnaire at the first study visit; therefore, we did not have different demographics per study visit.

- title should be reworded; it is not the demographics of the study population but of population included in analyses.

Response: This was updated.

- what are the concentrations of the biomarkers/hormones based on these demographics? Seems they should be included in this table as well.

Response: We did not include the concentrations of the biomarkers and hormones by the demographics because of the number of biomarkers and hormones included in the manuscript. The demographic variables have up to four categories each, totaling 22 categories. The table that broke down the biomarkers and hormones by the demographics had a total of 484 cells, adding to an already results-rich manuscript. Therefore, we opted to exclude this table since we did not feel it added to the interpretation of the results, and focused our attention on the strongest/most interesting results.

Table 2:

- this table could also be reworded to include the term "distribution" and "difference" since the distribution is shown and the p-values for differences between visits are reported.

Response: This was updated.
- Again, if urinary biomarkers were used from visit 2, they should be included in this table as well.

Response: They were not included, but as described above, we substituted the visit numbers with the gestational ages for easier interpretation.

- Total n should be included for each biomarker/hormone instead of the total n for each visit.

Response: We included more details on the number of samples for the biomarkers/hormones in the footnote of the table.

Tables 4 and 5:

- aren't these just tables of results shown in figure 2? Seems that these tables should be moved to supplemental instead being in main manuscript.

Response: These were moved to the Appendix, as suggested.

- It would be helpful to include unadjusted models, so that the reader can understand the overall direction of confounding. This would aid the reader in deciding whether further adjustment would tend to strengthen or attenuate the results.

Response: Thank you for your comment. Due to the large number of independent and dependent variables, instead of including exhaustive tables of unadjusted results, we have now added text describing the most notable differences between crude and adjusted models in the Results section (page 11). There were few differences in interpreting our results when we compared our adjusted and unadjusted models. Some of the associations we observed were slightly strengthened, and some of our suggestive associations (p values between 0.05 and 0.10) became weakened. There were only two major differences. First, BP-3 was associated with a decrease in T4 and TSH in our unadjusted models, but not in our adjusted models. Second, only MPB and PPB were
associated with CRH in our unadjusted models, but in the adjusted models, these associations disappeared, and CRH was associated with BPS and TCS. A further analysis of CRH concentrations across the covariate levels did not reveal any large differences to report.

Figures

- Are the results based on LOD substituted concentrations as well? Seems worth adding this to footnote.

Response: This note was added to Figure 1 as suggested.

- Figure 1: are these correlations of log- or un-log-transformed concentrations? Should be stated in the figure and in results section.

Response: This was added to the Figure and Methods section.

- Figure 2 is incorrectly labeled as Figure 1.

Response: Thank you for pointing this out, it was due to cross-referencing not updating itself. The problem should be fixed now.

Results

- The term "suggestive" was inconsistently used throughout the results section. In some areas "suggestive" is used and in others, "no association" or "not statistically significant" was used. I assume all of the above is referencing "no statistical significance"; if so, it would be best to just report the direction of the effect estimate and 95% CI and state that it the null is included or that the effect estimate did not reach statistical significance. I would make sure all statements are parallel.
Response: Thank you for your comment. We use the term suggestive to indicate p values between 0.05 and 0.10, whereas “no association” refers to associations with higher p values. We streamlined the entire Discussion section to include only the most interesting and significant results, so we hope this eliminates the confusion.

- "data not shown" was stated on Page 10: line 27, it would be best to either include the data in the tables/figures or not include these results in the manuscript.

Response: This was updated.

Discussion

- Could you say something about the overall sizes of the estimates? How large or small are they?

Response: These were added as suggested.

- Discussion of different findings across studies: to what extent can exposure differences explain these findings?

Response: Thank you for your comment. The exposure differences can explain some of the differences, particularly in the case of triclocarban. A further discussion of this was included as suggested.

- Please clarify the term "vulnerable" on page 12: line 36.

Response: This was updated.
- Please include citations for this statement: "...due to the unique role CRH plays in human pregnancies..." Page 14: lines 36-37.

Response: This has been added.

- There is a human study that examined triclosan during pregnancy as a potential vulnerable period as well. Page 15: lines 32-36. Please discuss this human study in addition to the rat study. Jackson-Browne et al 2018 published in Environmental Health Perspectives: "Identifying vulnerable periods of neurotoxicity to triclosan in children."

Response: We thank the reviewer for drawing our attention to this paper, this has been added.

- There is also an additional study examining triclosan and thyroid hormones by Braun et. al. 2017 published in Hormones and Behavior: "Associations of early life urinary triclosan concentrations with maternal, neonatal, and child thyroid hormone levels." Please discuss your studies results compared to theirs. Page 15: lines 39-53.

Response: This was added.

- The Wang et al paper cited [80] also looked at triclosan and fetal testosterone from cord blood. Please include results in discussion. Page 16: line 4

Response: Thank you for your comment. We chose to only discuss their results pertaining to maternal hormones since the association between triclosan and fetal hormones may not necessarily directly reflect the associations between triclosan and maternal hormones.

- In your limitations/strengths sections, you did not discuss whether your results generalize to other populations. You discussed how your associations differed or agreed with results of similar studies but the addition of how the concentrations of biomarker and hormones in this cohort relate to other cohorts would help address the issue of external validity. In addition, adding a
column to Table 1 to help identify how the 602 participants in this study are different from the overall cohort help address internal validity. i.e. are those included in the analyses cohort more educated, higher income, etc.

Response: Thank you for your comment. Our study population was based in a population in Puerto Rico of lower income who also had higher urinary concentrations of some of the exposure biomarkers; therefore, the results may not be fully generalizable to other populations. A discussion of this was added to the paragraph discussing limitations (page 26).