Author's response to reviews

Title: Evaluation of Maternal PCB Exposure related to Time to Pregnancy in Daughters

Authors:

Chris Gennings (gennings@vcu.edu)
Caroline C Carrico (ckcarr2487@gmail.com)
Pam Factor-Litvak (prf1@mail.cumc.columbia.edu)
Nickilou Krigbaum (nkrigbaum@chdstudies.org)
Piera M Cirillo (pccirillo@chdstudies.org)
Barbara A Cohn (bcohn@chdstudies.org)

Version: 2 Date: 22 June 2013

Author's response to reviews: see over
RESPONSE to Reviewer #1 Comments (italics):

The work proposes a novel statistical method to deal with exposure to pollutant mixture in association studies without any a priori knowledge on action mechanisms of these diverse pollutants. In this sense, this work appears to be important since such statistical method is still lacking in environmental epidemiology. The authors applied their method on previous data, already published, studying the relation between maternal serum PCBs levels and the daughter’s time to pregnancy several decades later.

RESPONSE: We agree that our approach is lacking in environmental epidemiology and appreciate the reviewer’s comment.

General comments:
(M) - Applying a statistical method on data that provides converging results does not mean that we reach the « truth ». For instance, we can calculate a beta coefficient with linear regression even if the relationship is not linear. I would expect in this kind of work additional elements that might help us to believe in this novel method: it might be a statistical validation (ex: with simulated data) and a clear discussion on the assumptions of the method and their influence on the result interpretation. Then, discussion comparing between previous results of Cohn et al. and novel results should be extended to understand what we are missing with standard regressions (as in the previous results) and what this novel method might provide, or/and inversely. Similarly, discussion comparing this novel method with other statistical methods proposed for dataset with correlated exposures might be useful.

RESPONSE:
We agree with this thought and have added a discussion paragraph (page 13?) about the results from simulation studies from Carrico 2013. In her simulated studies, WQS performs well when the correlation between the mixture components and a continuous response, Y, is about 0.3 with environmentally relevant correlation among the mixture components (based on 11 chemicals). In comparison, her simulation results also indicated LAR LASSO tends to identify correct components but also incorrect components with high probability. We believe that further simulation studies are beyond the scope of this paper.

(M) - Because this work aims at considering PCBs without prior knowledge on their classification: (1) results for individual congener must be highlighted, and (2) other possible PCBs classifications (e.g. structure-based, mechanism-based other than Wolff 1997) must be considered in the Table 1 and discussed.

RESPONSE:
We appreciate the thought of including an additional classification scheme. We added a structure based grouping on the number of chlorines as described by Kezios et al 2012. Combining the empirical approach with these groupings is informative in generating hypotheses about TTP. For example from the new Table 2, it is evident that the highly chlorinated PCBs are most associated with longer TTP.

(M) - There is only little discussion.

RESPONSE:
We have greatly enhanced the discussion section of the paper as evidenced by the addition of more than 30 references.
Abstract:

The authors need to provide additional elements to help in interpreting the % provided in the Results section. The bootstrap analyses for the weight estimates has to be introduced in the Methods section. In the results, the comparisons of % are not useful, providing 60% and 23% is sufficient. The average weight of 0.47 specific to the PCB56 estimated in association with longer TTP should be mentioned in the abstract.

RESPONSE:
We have edited the abstract making the points as suggested.

(m) - page 4 “There is mounting of evidence... fetal growth”: This sentence and the 2 related references are not well appropriate for the present study since they are very recent and they question about a possible role of the current PCBs exposure levels on reproductive health. The question of the present study is how the PCBs levels found in humans in 1960s (when PCBs was still used) may impact on the reproductive health of the next generation. The authors should adapt their introduction in this way.

RESPONSE:
The references in this sentence refer to in utero exposure and we have clarified this in the manuscript.

(m) - page 4 “Increased TTP is most likely multifactorial... pubertal periods”: the authors should add references to support this sentence. The authors have to argue why they do not mention the adulthood exposure such as tobacco for instance.

RESPONSE:
We have deleted the sentence.

(m) -page 5: in order to compare with the novel method, the authors must provide what is the statistical model used by Cohn et al.

RESPONSE:
A description of the approach by Cohn et al is provided in the introduction with a statement of the results. Table 1 has been edited to include the results from Cohn et al. for easier comparison.

(m) - page 6: replace Table 1 by Table S1. What does “complex” correlation structure mean? The author should provide additional elements clarifying the reasons or the conditions that render standard regression strategies problematic in case of correlated factors. In fact, there is always correlation, and so this sentence deserves to be more precise.

RESPONSE:
We appreciate the reviewer noting the typo referring to Table S1; the correction has been made. We have provided a reference pointing to the problems inherent in standard regression with multicolinearity (from Myers 1990, Chapter 8). We have added a sentence about the weighted index stabilizing the ill conditioning with reference to Carrico 2013.

(M)- page 6: after “Instead, we use a non linear....”. Instead of summarizing the method that is explained in details in the Methods section, we would expect to read the rationale of the construction of the novel method: why is it nonlinear? Why the use of quartiles, what are their advantages in case of correlated exposures? What does the use of weight bring in the method?...
RESPONSE:
We have added such a discussion/description as suggested on pages 5-6.

(m) -page 6: Referring only to the National Collaborative Perinatal Project for concluding on the possible role of organochlorine exposure on TTP is too limited. The authors should consider other studies: Axmon 2004, 2006, Buck Louis 2008, Yang 2008, Law 2005, and more recently Chevrier 2013 and Buck Louis 2013
RESPONSE:
We have removed the sentence in the introduction and added a full discussion paragraph (pages 21-22) reviewing the suggested papers.

(d)- Methods: additional details are required in this manuscript concerning the selection of these 289 daughters of women having participated in the CHDS. For instance, is it a random sample? Discuss quickly about possible selection bias.
RESPONSE:
The methods section has been improved by adding a more complete description of the study as described by Cohn et al (2011).

(m)- Methods: when studying TTP, sensibility analyses are required such as, at least, doing same analyses on primipara women only.
RESPONSE:
We note that in our original publication on this sample, Table 3 gives results stratified by daughter’s gravidity. We do not repeat that information in the current manuscript, but we now refer to the effect of gravidity on findings in the discussion where we have added a paragraph to the discussion on page 18.

(d)-page 7: we do not need to know the reference Carrico et al. if it is under review. However it is useful to understand in what way this strategy extends the previous works proposed by Gennings and Christensen. The authors should add these elements.
RESPONSE:
We have changed Carrico et al to Carrico (2013), a dissertation, and included the point that Carrico’s extension was to include a bootstrap analysis of the weights on page 10.

(d)-page 7: “The assumption of a Weibull... function”. What did the author do if the assumption was not verified?
RESPONSE:
The assumption was verified through diagnostic plots (page 9).

(m)-page 8: correct the formula and clarify what is x, lambda and gamma. Where is delta “indicator of pregnancy or censoring” cited in the text?
RESPONSE:
We appreciate the reviewer noting the typo regarding x; the equation has been corrected. We have also defined the Weibull parameters below equation (1). We have removed the reference to delta since it is not a parameter in the hazard function.

(m)- page 8: Add reference justifying that the Truth region method is more stable for moderate
sample size.

**RESPONSE:**
A reference was added to the SAS documentation of the procedure (page 10).

(m)-page 9: Did the author use the NLP procedure for checking the estimation consistency or for choosing the best optimization techniques for the non linear model? The author should clarify that, and/or they should add results concerning these algorithms.

**RESPONSE:**
For space considerations we have elected to state that analyses conducted with other algorithms including the conjugate gradient method, the Newton-Raphson, the Nelder-Mead Simplex, and the Newton-Raphson with ridging, were consistent to what is presented here (page 10).

(m)- page 9: the number of samples and analyses is not clear. Right now, we understand that for one sample there is one analysis and this is not coherent with the “100 bootstrap samples” and the third bullet.

**RESPONSE:**
We have edited the bullets to make the approach more clear.

(M)-page 9: due to the huge number of samples and analyses, the authors should mention and justify the threshold used for statistical significance, used in the construction of the histograms and the computation of the average indices.

**RESPONSE:**
A sentence was added on page 11 to indicate we only average the weights that have evidence of a significant slope as weights estimated in samples with the slope near zero are not interpretable.

(d)-page 12: The 2 sentences starting with “On average...” must be inversed in order to keep the same order as in the Figure 2.

**RESPONSE:**
The sentences were reversed as suggested.

(d)- page 12: What does the Figure 3 and its related text bring to this work? Please clarify more.

**RESPONSE:**
We have added more text about the figure. In short, the figure demonstrates the rough comparison between the groups defined by the median cuts. For example, as stated, the combination with the shortest TTP (high concentrations of PCB+ and low concentrations of PCB-) has roughly a 10-20% increase in the probability of pregnancy compared to the combination with the longest TTP (low concentrations of PCB+ and high concentrations of PCB-). In addition, the plot indicates the cumulative probability of pregnancy increases most rapidly roughly within the first three months.

(d)-page 13: replace “the other group” by “unclassified”

**RESPONSE:**
The suggested edit was made.

(m)-page 13: the second paragraph is the same as the one in the introduction. Please choose its right place.
RESPONSE:
We thank the reviewer for the note. We have left the paragraph in the discussion section.

(m)- Summary: there is contradiction in the Summary section: the authors firstly summarize their results using the Wolff classification, and then they state that this method “can generate new hypotheses” without these a prioris. The authors should instead highlight their results on individual congeners that might not correspond to Wolff classification.

RESPONSE:
We appreciate the referee’s comment. Our point is that the empirical weights are not constrained by hypothesized categories; however, combining the weights with these categories may generate hypotheses. For example, the mono-, di-, and tri-ortho groups considered by Kezios were not important in predicting longer TTP – instead, the higher chlorinated group dominated the association.

(m)- Did the current PCBs levels in these daughters be assessed? If yes, the authors should take in consideration these current exposures. If not, this should be discussed, as well as the potential role of other risk factors that are missing in the study (tobacco consumption of the mother and the daughter, age…).

RESPONSE:
Current levels of PCBs were not assessed, but maternal and daughter exposures to tobacco (including passive smoking), alcohol and caffeine, daughter’s health history, including sexually transmitted diseases, race, education, income (mother and daughter), history of oral contraceptive use and other contraceptive use, marital status, age at menarche, abortion history (daughter) were assessed. We selected to use the same covariates as in Cohn et al (2011): race (African American vs all other) and whether the daughter was breastfed (yes or no). The final model was further adjusted for maternal variables: age, body mass index, and maternal lipids (triglycerides, cholesterol) as described on page 9.
REVIEWER #2
This is a very interesting paper and one that begins to address critical data gaps regarding POPs and human fecundity. Moreover, the authors are commended for developing novel approaches for considering mixtures of PCB congeners to more closely analyze exposure consistent with the manner in which couples are exposed.

My overall comments are specific and are offered with the goal of helping to interpret the findings in the context of evolving findings focusing on chemicals and human fecundity as measured by TTP.

1. It would be helpful to the reader if a succinct description of the methodology used in the original study were provided, rather than to make the reader pull this paper to help interpret the current work (many readers won’t bother). This description should make it clear if only one pregnancy per daughter was analyzed (and if so, the need to make it clear that this pregnancy is assumed to be representative of all pregnancies per woman), and the distribution of gravidity and parity for participating daughters. This is important as retrospective TTP studies such as this one and prospective TTP studies alike report shorter TTPs for parous than nulliparous women. [I am not advocating the inclusion of parity conditional on gravidity; rather, I am trying to suggest some discussion relative to interpretation.]

RESPONSE:
---- As requested, the methodology used in the original study has now been added to the methods section.

---- We note that in our original publication on this sample, Table 3 gives results stratified by daughter’s gravidity. We do not repeat that information in the current manuscript, but we now refer to the effect of gravidity on findings in the discussion where we have added the following sentences to the discussion:

We were unable to stratify analysis by gravidity, as sample sizes would have been too small for analysis by the methods we used here. However, we previously compared results for daughters who were nulligravidas (N=54), primigravidas (N=70) and multigravidas (N=159) in our prior publication that is the basis for this comparative methods study. (Table 3 in {Cohn, 2011 #7662}) In that comparison we showed, as expected, that TTPs were shorter for parous women, but that overall, PCB associations observed were consistent for all three gravidity groups.

2. It would be helpful to pull in some of the evolving literature that suggests female fecundity may be remodeled or reprogramed during pregnancy (or earlier), given that daughters’ exposures seem to be in utero exposures.

RESPONSE:
We have added an extensive paragraph on recent literature on the developing reproductive system may be influenced by maternal exposures on page 16.

3. Bloom and colleagues have reported that PCBs decline during pregnancy, particularly in relation to women not achieving pregnancy during 12 months of trying. As such, the postpartum concentrations may be an underestimate of in utero exposure. This may be worth noting.
RESPONSE:

There are several differences between the design of the Bloom et al. study and our own that make it difficult to predict how the use of postpartum samples influenced results. Below we note our thinking on this. We have now added discussion of these issues to the manuscript as well (see Summary below).

1) Our postpartum samples were drawn within 1-3 days of delivery, compared to 6 weeks postnatally for Bloom et al. Based on mean daily rate of change estimates presented in Bloom et al., (Table 4), these three days past delivery in our study would likely result in very small differences between in utero PCB levels and measured PCB levels in postpartum samples collected so soon after delivery (e.g. mean change of 0.008 ng/g serum per day for total PCBs, and near zero daily changes for PCB 118 and PCB153).

We note that the inter-individual difference in maternal PCB levels greatly exceeds the variability due to sample timing (e.g. for PCB 118, the median in our study was 0.44 ng/g serum with a range of 0.14- 2.46 ng/g serum).

2) We decided to use the early postpartum samples for our studies of TTP:
   a. First because we wished to conserve timed samples available in the first, second and third trimester for studies where exposure does depend strongly on timing of the sampling (for example thyroid hormone, steroid hormones); and
   b. Second because our primary objective was to rank CHDS daughters according to their mothers’ organochlorine levels during daughter’s gestation. Prior work by Longnecker and colleagues {Longnecker, 1999 #3629} had suggested that maternal postpartum samples (even 35-48 days post-delivery as in the Collaborative Perinatal Population) would accurately rank in utero exposure because inter-individual variability greatly exceeded intra-individual variability over time.

3) Regarding differences that Bloom and colleagues report for infertile women versus women who conceived: these findings are not relevant to our study since all CHDS mothers in this study of daughter’s TTP had by definition, conceived.

In summary, as we now note in the manuscript: It would have been ideal to measure maternal PCBs at multiple points in gestation in relation to daughter’s TTP 30 years later. However, it was our objective to conserve valuable maternal prenatal serum samples for other research questions that require precisely timed sampling, such as steroid hormones and thyroid hormones. We thus chose to use maternal postpartum samples as a proxy for daughter’s in utero exposure to PCBs. While it is possible that some PCB congeners are more subject to change during pregnancy, {Bloom, 2007 #8203} the postpartum blood samples in our study were drawn within 1-3 days of delivery, minimizing this source of misclassification. Based on rates of daily change in PCB levels reported by Bloom and colleagues, {Bloom, 2007 #8203} variability due to changes in PCB levels over three days after delivery would be small and considerably less than the substantial inter-individual variability of maternal PCB levels we observed in this sample. {Cohn, 2011 #7662} For example, based on mean daily rate of change estimates presented in Bloom et al., (Table 4), these three days past delivery in our study would likely
result in no or in extremely small differences between *in utero* PCB levels and measured PCB levels in postpartum samples collected so soon after delivery. Consistent with this argument, Longnecker et al. suggested that using postpartum maternal samples does not introduce serious misclassification in ranking daughters according to their *in utero* PCB exposures, since inter-individual variation in PCB levels greatly exceeds intra-individual variation due to timing of sampling. {Longnecker, 1999 #3629}

4. TTP is very challenging to measure irrespective of prospective or retrospective measurement, with the latter approach utilized in this work. A better description of how TTP was measured is needed, as it seems the authors lumped recalled TTP (actually trying) with at risk time but not necessarily trying. This latter approach is the “current duration” approach, but it does have methodologic implications and limitations as the authors of this work note. (See Keiding and Slama’s various papers). It would be helpful to have these issues addressed along with the rather poor validity for retrospective TTP when using the gold standard of prospective measurement. This work suggests poor validity for long-term recall (Cooney et al.), but good for short-term <2 months recall. (Zeilhaus et al.) It would be helpful if the authors address how digit preference, a common occurrence with retrospective recall, might impact the findings (Weinberg et al., Joffe et al.).

**RESPONSE:**
We have added such a paragraph to the discussion section as suggested on page 19.

5. While I understand the choice of PCB groupings utilized, it should be noted that this classification isn’t without error as the biologic activity of many PCB congeners is unknown. Only a subset of congeners is even considered, possibly a reflection of what was known about congeners at the time of classification. It would be helpful for the authors to comment on their rationale for using this classification rather than a completely untargeted approach or even other mode of classification. [There have been two recent supplements devoted to the modeling of biomarkers including chemical exposures – Epidemiology 2012 and Statistics and Medicine 2013.] I am not suggesting it isn’t appropriate to use; rather, I think it helps to delineate the assumptions with any choice.

**RESPONSE:**
We agree that functional/structural classification of PCBs has not only been controversial, but also sometimes limiting in the literature. The Wolff lab measured our PCBs and the congeners we report were state of the art at that time and for her lab. Measurement, conceptualization of PCB classification, and approaches to mixtures are rapidly evolving in this field.

We have added the following to the Discussion Section: When we published our first paper on in maternal PCBs and daughter’s TTP {Cohn, 2011 #7662}, there was active work to determine whether the classification proposed by Wolf and colleagues {Wolff, 1997 #2647} demonstrated predictive validity for a number of outcomes. By now it appears not to hold up well as we noted in our 2011 TTP paper {Cohn, 2011 #7662} and in a recent paper on maternal breast cancer {Cohn, 2012 #8165} and in other papers on infant health outcomes in the CHDS. {Kezios, 2012 #8032} It is for this very reason that we attempted an alternative approach that did not require an a priori classification. In reporting our results, we did consider the Wolff et al. classification, in order to provide information in this paper that could be compared to other papers based on this
widely used classification. We do not mean to suggest that this classification is better than any other, but we are suggesting that an empirical approach, independent of a priori classification, can be very useful in light of a gap in understanding the biological effects of simultaneous exposure to multiple congeners in different proportions.

6. There is other research than the CPP work focusing on PCB congeners and TPP, but this work is not cited with regard to the congeners deemed relevant here. The CPP work is limited to women achieving a birth; hence, not directly comparable to the work here that focuses on pregnancy and not just live birth (at least as I understand the work).

RESPONSE:
We have a multiple paragraphs to the discussion section addressing this (pages 20-22).

7. It might be helpful to interpret the findings in relation to PCB congeners and prospectively observed TTP among couples trying to get pregnant (Buck Louis et al. 2013). This paper likely overlapped with submission and publication. Still, this work identified reductions in TTP ranging from 18-21% for congeners #118, 167 and 209. More PCB congeners in men were associated with a reduction in TTP by 17-29% (#118, 138, 156, 157, 167, 170, 172). It might be helpful to interpret the current findings for the current generation of couples trying with the findings in offspring, at least for the congeners measured in both studies (#118, 1138, 56, 167, 170). Unlike the present work, none of the FORs were significantly above 1 for females and only 1 in males.

RESPONSE:
We have added the following paragraph to the discussion section.

In our study we found increased TTP with PCBs thought to be antiestrogenic (PCBs 66, 74, 105, 118, 156 and 167) as well as those not classified in any of the Wolf et al (1997) groupings (PCBs 56, 146). In contrast, shorter TTP was associated with PCBs thought to be anti-estrogenic (PCBs 66,74,105, 118, 156 and 167) and those thought to be PB-inducers (PCBs 99, 153, 180, 203 and 183). Buck Louis et al (2013) in the LIFE study also found reduced fecundability with maternal PCB congeners 118 and 167. They also found reduced fecundability with paternal PCB congeners 138, 167 and 170; for the latter two we found shorter TTP. Because the exact mechanisms by which these associations are likely, it is hard to discern why the results differed for paternal exposure patterns. One possible explanation is that we considered empirically derived mixtures of congeners while Buck Louis evaluated congeners individually.