Reviewer's report

Title: Application of a Weighted Quantile Sum approach to identifying Bad Actors in Maternal PCB exposure related to time to pregnancy in daughters

Version: 1 Date: 9 April 2013

Reviewer: cécile Chevrier

Reviewer's report:

25 comments:
(M): 5 major comments
(m): 14 minor comments
(d): 6 discretionary comments

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. I have several concerns for this paper.

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.

(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.

(M)- There is only little discussion.

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.

“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.

“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.

In order to compare with the novel method, the authors must provide what is the statistical model used by Cohn et al.

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 precised.

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?...

Refering 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

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.

When studying TTP, sensibility analyses are required such as, at least, doing same analyses on primipara women only.

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.

“The assumption of a Weibull… function”. What did the author do if
the assumption was not verified?

(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?

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

(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.

(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.

(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.

(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.

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

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

(m)-page 13: the second paragraph is the same as the one in the introduction. Please choose its right place.

(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.

(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…).

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

no competing interest to declare