PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort profile: Chiba study of Mother and Children's Health (C-MACH): cohort study with omics analyses
AUTHORS	Sakurai, Kenichi; Miyaso, Hidenobu; Eguchi, Akifumi; Matsuno, Yoshiharu; Yamamoto, Midori; Todaka, Emiko; Fukuoka, Hideoki; Hata, Akira; Mori, Chisato

VERSION 1 - REVIEW

REVIEWER	William Toscano, Professor of Molecular Toxicology University of Minnesota, School of Public Health, Division of Environmental Health Sciences, MMC 807, 420 Delaware St. SE Minneapolis, Minnesota 55455, United States of America
REVIEW RETURNED	23-Nov-2015

GENERAL COMMENTS	This is a well-written manuscript describing the cohort birth study based in Chiba Japan. The strength of the study is that exposure from chemicals, especially persistent organic pollutants (POPs) will be examined, using well accepted surrogates (PCBs) in the exposure assessment. The authors will examine both outdoor and indoor air sampling. In addition to the examination of chemical exposures, they investigate social determinants as well. They will combine their exposure studies with epigenetic studies, and health outcomes. This is a very modern approach to examining whether verse outcomes of physiology are related to exposures. This work will make a great contribution to the understanding and testing of the DoHaD hypothesis that suggests in utero and early childhood exposures are causes of later adult health or disease. Some minor suggestions include: page 2, line 17, suggesting should read supporting page 3, line 31 wider, should read wide page 3, line 36, by that of their mothers should read by that of the mother page 3, line 49, these recently developed methods, needs an explanation. There is no mention of the new methods to which the authors are referring Page 6, line 46, "delete, "Follow up after the age of 5 years will be considered later". This sentence dosen't mean anything. The authors could say Page 10, line 51, "...resulting in 408 women.", should read "...resulting in a final cohort of 408 women."

REVIEWER	Leda Chatzi University of Crete, Greece
REVIEW RETURNED	08-Dec-2015

GENERAL COMMENTS	In this manuscript authors present a new pregnancy cohort study, the Chiba study in Japan. Although potentially very valuable data

	sets will be generated with this project, there are some major concerns regarding the methods used, which need to be addressed. 1. Sample size: The main limitation in epidemiological studies including omics analyses is the small sample size. The sample size description in methods (page 5) is too general and does not address the specific hypotheses of this project as it is only based on outcome prevalence rates. The authors should carefully re-write this section and provide sample size estimations for the specific aims (1-5) as describe in page 4. For example, is the sample size of n=500 mother-child pairs enough to identify biomarkers of epigenetic regulation to predict children's health problem such as obesity, asthma etc (aim 1)? Similarly, is the sample size of n=100 pregnant women enough to investigate whether the gut microbiota of mothers affects the health of their children (aim 4)? Most probably the study is underpowered to detect these associations. 2. Study participants: Although the authors state that there were approximately 4000 potential candidates, consent to participate was obtained only by 433 women (~10% participation rate). This low participation rate is important as it may introduce selection bias. The authors do not provide any information on this issue. I strongly recommend to present comparisons between participants and non-participants in terms of sociodemographic characteristics, and discuss the main reasons of not participation in the study. This should be described in the results section, added as a new Table, and discussed in the limitations. 3. Introduction: The authors do not describe previous analyses from birth cohort research on this topic (epigenetics, micriobiota, metabolomics), while they inaccurately state that there are no previous studies on this topic (metabolomics). 4. Exposure assessment: While throughout the manuscript the authors state that they will investigate the effects of exposure to environmental chemicals, the only chemicals that will be measured are PCBs. Why did they choose the specific group of chemicals? Is exposure to these specific chemicals highly prevalent in this population? Do they have information on other environmental pollutants such as non persistent chemicals (ie BPA, phthalates), heavy metals, and air pollution levels? 5. Metabolomic analysis: It is not clear if the authors refer to targeted or untargeted metabolomic analysis. The authors just describe that "the composition of metabolites, such as amino acids, lipid metabolites, and vitamines in blood and urine and their associations with outcomes will be analyzed". More information should be provided on the methods for metabolomic analyses and their association with outcomes. 6. Methods, Covariates: The authors describe that they will perform biochemical tests and hormone-level measurements. Please provide more information on this: Which exactly biochemical and hormone measurements will be performed? (please indicate type of biological data, time of collection, method) 7. Statistical analysis: The description of the statistical analysis is not adequate for the proposed study. The authors will generate a big dataset with multiple exposures and outcomes, and they cannot answer these complex research questions by just performing multiple logistic regression analysis.

VERSION 1 – AUTHOR RESPONSE

Response to the reviewer 1
We are grateful to you for the critical comments and useful suggestions that have helped us to improve our paper. As indicated in the responses below, we have addressed all these comments and suggestions and revised the manuscript accordingly.

1. We have changed “suggesting” to “supporting” (page 4, line 8).
2. We have changed “wider” to “wide” (page 4, line 17).
3. We have changed “that of their mothers” to “that of the mother” (page 4, line 20).
4. We have mentioned some text describing our analytical methods to explain “these recently developed methods” (page 4, line 31).
5. We have deleted the sentence "Follow up after the age of 5 years will be considered later" (page 7, line 31).
6. We have changed “resulting in 408 women” to “resulting in a final cohort of 408 women” (page 12, line 9).

Response to the reviewer 2
We are grateful to you for the critical comments and useful suggestions that have helped us to improve our paper. As indicated in the responses that follow, we have addressed all these comments and suggestions and revised the manuscript accordingly.

1. It is difficult to accurately estimate the required sample size for metabolomics analysis and other omics studies using orthogonal projections to latent structures analysis (OPLS), l1/l2 regularized regression analysis, and gradient boosting. Therefore we referred to a previous study that used metabolomics (page 6, line 29). And we agree your suggestion that some analyses are underpowered to conclude the study, so we changed a word “Ascertain” to “Explore” in the INTRODUCTION section (page 5, lines 24 and 26).
2. We added the following sentences to the FINDINGS TO DATE section (page 12, line 12): “These participants were almost 10 % of the potential candidates. We randomly contacted approximately one-fifth of all of the subjects who came to the hospitals and clinics. Around half of them consented to the study.” We do not have the records of non-participants, so we cannot make a new table as you suggested. Although we contacted the candidates randomly, there may be potential selection bias. We have mentioned that possibility in the STRENGTHS AND LIMITATIONS section (page 14, line 23). We added the following sentences to the STRENGTHS AND LIMITATIONS section “We could not contact all potential candidates because of the limited number of staff. We could contact one-fifth of potential candidates although the ratio of contacted subjects to all potential candidates was slightly different for each medical institution.”
3. We have mentioned the previous studies using omics analyses and added some references to the INTRODUCTION section (page 4, line 25).
4. We focused on the POPs in this study, so we evaluate PCB concentrations. As mentioned on page 8, line 21, we have demonstrated that PCB concentrations in blood and umbilical cord samples are correlated with the concentrations of various POPs. Therefore, we will use PCB concentration as a surrogate marker in this study
5. We added more detail about the methods of the metabolomics to the METHODS section (page 10, line 26).
6. We added the type of biological data measured using biospecimens in Methods, Covariates section (page 11, line 7).
7. We added additional information about statistical methods to the METHODS section. The detailed procedures will also be described in other articles in the future. (Page 11, line 15)