

Dear John

We are very pleased that you have given us the chance to revise our paper, Manuscript ID BMJ.2018.043743 entitled "Prenatal biochemical screening and long-term risk of maternal cardiovascular disease"

We respond to every point below, and have used track changes in the manuscript as well. If we have not addressed things as you would like, please let me know, and we will work with you to get this paper fit for publication in the BMJ.

Thank you, John

Joel

### Detailed comments from the meeting:

1. First, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below.

→ Done

3. Our statistician made the following observations:

a. It is not clear why a CVD composite (which covers a number of different events) was chosen as the primary outcome. Is it not more appropriate to select just one event?

→ Prior works suggests that maternal placental syndromes (e.g., preeclampsia) are associated with a higher future risk of coronary artery disease, cerebrovascular disease, heart failure and dysrhythmias, and peripheral arterial disease. Accordingly, we chose a composite of these outcomes, together. Doing so would maximize the number of outcome events, all of which are premature in nature.

→ Certainly, we agree with the statistician, and now also present: i) coronary artery disease alone; and ii) cerebrovascular disease alone. These are the top-2 most common CVD outcomes.

→ We have added these two outcomes to Table 2 and within the Methods and Results.

b. One pregnancy per woman was chosen at random (to simplify the statistical analysis). Analyses are said to have been adjusted for gravidity, but would a previous history of abnormal prenatal screening results (for those with gravidity greater than zero) be a possible confounder? For those women with greater than one pregnancy during the study period, what proportion had more than one abnormal screening result?

→ We followed the reviewer and present those proportions in Supplementary file 9 ("Description of the number of women who had at least 1 pregnancy with a measurement of a given serum analyte (column a), the percentage of those women who had 2 or more pregnancies in which that given analyte was measured (column b), and the percentage of those women with 2 or more abnormal results for that given analyte (column c)."

→ Next, in Supplementary file 10 (Additional analysis 5), we re-ran the Main model for CVD, restricted to women with 2 or more pregnancies with prenatal biochemical screening, and further adjusted for the number of prior pregnancies with a given abnormal serum analyte. The findings did not change from the Main model in Table 2.

c. Table 4. The analyses here relating to the number of abnormal analytes show results for 855,536 women. How is this possible when PAPP-A was carried out on only 499,790 women and DIA on only 98,160 women (Table 1)? What assumptions have been made here?

→ AFP, hCG and uE3 serum analytes were available in all years; PAPP-A starting in 1999; and DIA starting in 2003. As such, the maximum number of abnormal analytes that a woman could have in a given pregnancy varied depending on the year of collection.

→ To deal with this distinction, we revised Table 4 (Additional analysis 2) to show results separately for specimens in the era of 1993 to 2002, and those in the era of 2003 to 2011 -- that is, before and after all 5 tests were available. In lieu of this change, and a splitting of the number of outcomes into two era, we therefore present the results for 0, 1 and  $\geq 2$  abnormal analytes.

→ These new results are also updated in the manuscript and in the Abstract.

d. Table 2. It is not clear how the smaller incidence of the CVD composite outcome for those with abnormal readings on hCG compared to those with normal readings (5.8 per 10,000 vs 6.0 per 10,000) leads to a unadjusted hazard ratio of 1.3.

→ We spent a fair amount of time ensuring that this was not an error or spurious result.

→ We also consulted two epi colleagues: The incidence rate ratio (IRR) is only similar to the hazard ratio under specific conditions: The conditions are proportional hazards (which we confirmed were not violated for any of the analytes) and exponentially distributed hazards.

→ The noted discrepancy between the IRR and the HR for hCG suggests that we may have non-exponentially distributed hazards. We did not pursue this examination, given that the Cox proportional hazards model relaxes the assumption of the exponentially distributed hazards, unlike the Poisson model, and the HRs are more likely to validly express the risk (i.e., hazard) between the low vs. not-low hCG exposed pregnancies.

4. The editors all thought that your observations are at the beginning of a story and do not have any direct clinical implications at the moment. We do not often publish research that has no direct clinical application but have made an exception this time because we think some of our readers will enjoy thinking about the information available from screening tests in a new way. Your discussion is suitably reserved about practical implications.

→ Thank you.

5. Some editors were unsure why you had excluded births with congenital or chromosomal anomalies from the analysis. Please can you add a short explanation as to why this is a good idea?

→ Births with anomalies, as well as pregnancies with MoM outliers beyond 0.2<sup>nd</sup> or 99.8<sup>th</sup> percentiles, were removed only in the first step in our statistical analysis (i.e., the fractional polynomial regression that was run to identify the best fitting plots and selection of abnormal cut-points for each serum analyte)

→ These records with anomalies were subsequently included in all other analyses.

→ The first step was undertaken so as to remove the influence of extreme outliers on the selection of an meaningful abnormal cut-point.

→ We revised the description of this process in the Statistical analysis section as: “After removing livebirths or stillbirths with a congenital or chromosomal anomaly diagnosis on the maternal or newborn hospital record, as well as outliers of MoM beyond the 0.2<sup>nd</sup> or 99.8<sup>th</sup> percentiles<sup>17</sup> from the fractional polynomial models (but not from the study cohort), the best-fit plots were more interpretable (Supplementary file 3).”

6. Please can you present a table of the frequency of the various cardiovascular outcomes in this study? As these are women of childbearing age the pattern of cardiac outcomes may be different from the older men that are so frequently the participants in research with cardiovascular outcomes.

**→ Supplementary file 7 shows the sub-type of cardiovascular disease among the 6209 women who experienced a cardiovascular disease composite outcome event. This has been added to the Results section as well.**

7. If the numbers are large enough to make sense, please can you present the main analysis using the most common single cardiac conditions as outcomes as well as the composite?

**→Table 2 is now revised to also include i) coronary artery disease alone, and ii) cerebrovascular disease alone, as these are the two most common CVD sub-types in our study. See also “Detailed comments from the meeting” 3a.**

**Reviewer: 1**

Specific questions:

1- On Page 5 line 47: "Of all remaining deliveries, we randomly selected one pregnancy per woman as the index pregnancy in order to simplify the data analyses." Why not select the first pregnancy as the index pregnancy?

**→ In previous work on pregnancy and long-term outcomes, reviewers have suggested that the random selection of a pregnancy avoids limiting the eligible cohort to just women who are primigravid. In doing so, selection bias is reduced, and the findings are also more generalizable to a "typical" pregnancy.**

2- On Page 7 line 42: "HRs were adjusted for maternal age (continuous), gravidity (1, 2+, missing), neighbourhood income quintile (1, 2, 3, 4, 5, missing), rural residence (rural, urban, missing), ethnicity (Asian, "Oriental", Black, Caucasian, Hispanic, other, missing) and gestational age (continuous) – each at the time of prenatal biochemical screening – as well as maternal diabetes mellitus, chronic hypertension, renal disease, tobacco/drug use and dyslipidemia within 365 days before or after the start of the index pregnancy (i.e. within 2 years preceding time zero)." What % had preeclampsia or eclampsia? Did they have higher CVD events on followup?

**→ The prevalence of preeclampsia/eclampsia among all livebirth or stillbirth pregnancies in our cohort was 1.9% (see Table 1).**

**→ Around 2.1% of women with a preeclampsia/eclampsia diagnosis went on to have a cardiac outcome, while 0.6% of women without a history of preeclampsia/eclampsia had a cardiac outcome. This is certainly in keeping with prior studies on this topic.**

**→ Certainly, preeclampsia/eclampsia could be a mediator between abnormal prenatal biochemical screening and CVD. It is for this reason, among others, that we did not adjust for preeclampsia/eclampsia.**

**→ Rather, in Additional analysis 4, each biochemical analyte was evaluated in the absence or co-presence of a maternal placental syndrome (preeclampsia, gestational hypertension, or placental abruption or infarction), which is shown in Figure 1b, lower, and in the related text.**

## Reviewer: 2

1. The uptake of the prenatal screening has big variation geographically, thus the cohort went for the screening could be biased. The models were adjusted for rural residence, but it would be more useful if the cluster of regions could be taken into account.

→ Previous studies in Ontario (e.g. Reference 11 [Hayeems RZ, Campitelli M, Ma X, Huang T, Walker M, Guttman A. Rates of prenatal screening across health care regions in Ontario, Canada: a retrospective cohort study. *CMAJ Open* 2015;3:E236-4]) have shown that lower uptake of prenatal screening is associated with rural residence, as well as lower income quintile, low maternal age and multiparity. We suspect that much of the regional variation in Ontario would be explained by differences in rurality and socioeconomic factors, which we accounted for. This is especially so within a universal healthcare system and standardized prenatal biochemical screening across the province of Ontario.

→ Moreover, hospitalization for CVD are likely to occur, regardless of regional profiles, given the seriousness of the outcome. Hence, we probably have sufficiently adjusted for the potential influence of rural/urban and SES income status as a potential confounder within our models.

2. The study randomly selected one pregnancy per woman as the index pregnancy, and analysed all the data by adjusting for gravidity. Although the authors have acknowledged it as limitation, it could be investigated as a sensitivity analysis. A stratified analysis by gravidity for the selected cohort could clarify whether there is any effect of abnormal screening results from repeat pregnancies.

→ Upon adjusting for number of prior abnormal screening results we found no change in the estimates for any of analytes or outcomes. Please see our full response above to #3b in the 'Detailed comments from the meeting'.

### Reviewer: 3

1. What was the basis for the definition of composite CVD outcome which was used as the primary outcome? Peripheral arterial disease and heart failure are normally outcomes that are assessed separately.

→ We agree. It is for this reason that we added to Table 2 the individual outcomes of coronary artery disease, as well as cerebrovascular disease.

→ PAD and HF are conditions that are highly connected to CVD, both in terms of risk factors (HTN, diabetes, obesity, lipids, etc), and mechanisms (arterial occlusion, vascular remodeling, endothelial dysfunction). So, having them within the CVD composite is still consistent with what we know about how, and in whom, these conditions arise prematurely.

2. Why were the major adverse cardiovascular events (MACE) not rather considered as the primary outcomes? Why was all-cause mortality included in the definition for MACE?

→ See #1 above

→ There is a debate about what exactly to include in “MACE” outcome: Many include cardiovascular mortality, others include sudden cardiac death, and others include all-cause mortality. Another example (<http://bmjopen.bmj.com/content/6/1/e009598>) used yet another broader definition (“ACS, a resuscitation setting with ventricular tachycardia or fibrillation as first rhythm or a (sudden) death presumed to be of cardiac origin according to the medical file, stroke (ischaemic and haemorrhagic), TIA and acute peripheral ischaemia of the legs”). In our study, establishing cardiovascular death would be difficult, while an out-of-hospital cardiac arrest (sudden cardiac death) could not be adjudicated. Thus, we included all-cause mortality in MACE, and, for this reason as well, made MACE a secondary outcome.

3. The authors will need to review their outcomes again as not appropriate. In their conclusions, the authors recommend replication of these findings in other studies. Can this easily be done given the nature of the primary outcome specified?

→ Certainly, our study can be replicated by others, either using the ICD codes that we have used, or, a modification of those outcomes to be more or less broad.

### Specific comments

What is already known on this subject

1. This section needs to be revised as it does not reflect the title. The authors could provide some background evidence that led to their hypothesis

→ The reviewer is correct. We clarified that section, and added: “Abnormal prenatal biochemical screening is related to a higher risk of preeclampsia, and preeclampsia is linked to premature cardiovascular disease (CVD) in women.”

2. “Abnormal prenatal biochemical screening for trisomies and birth defects has been completed among millions of women” This statement is very vague and need to be completed or revised.

→ That was a typo. We now state: “Prenatal biochemical screening for trisomies and birth defects has been completed among millions of women.”

## Abstract

1. Remove “also” from the first sentence.

→ Done. We also removed “*completed among millions of women*” from the first sentence.

2. “It is unknown if abnormal prenatal biochemical screening is associated with premature CVD after pregnancy, the goal of the current study.” Please revise this and be specific about the aim of the study. Do the same in background.

→ We revised the Abstract to: “**Objective: Abnormal prenatal biochemical screening for trisomies and birth defects is related to a higher risk of preeclampsia. Preeclampsia is linked to premature cardiovascular disease (CVD). The current study examined whether abnormal prenatal biochemical screening is associated with a higher risk of premature CVD after pregnancy.**”

→ In the main text Background section, we think that the current wording is clear and sets the objective properly.

3. Please provide the mean/median follow-up time in years in place of/in addition to person-years

→ The first sentence of the Results now states: “**Results: Among 1,209,690 pregnancies, and after a median of 11.6 (IQR 7.3-17.0) years of follow-up, each of the five prenatal biochemical screening analytes was associated with a higher risk of CVD, especially DIA.**”

4. Please specify that the estimates were multivariate-adjusted

→ Done

## Background

Third paragraph lines 29-39. Most of this does not belong here.

→ We did not entirely agree. First, the Background section is already short. Second, at some point the reader needs to be introduced to prenatal biochemical screening, why it is done, and what its constituents are. We feel that it needs to be here, so that a non-obstetrics/non-genetics reader can follow the concept development.

## Methods

1. The authors indicated in the abstract that the goal of the study was to evaluate if abnormal prenatal biochemical screening is associated with PREMATURE CVD after pregnancy. The term PREMATURE was however not defined/discussed in the methods section.

→ In the Background section, we added: “**Several risk factors for placental vascular disease, especially for preeclampsia<sup>3,4</sup>, are shared with those for cardiovascular disease (CVD) (see Supplementary file 1). A maternal placental syndrome in pregnancy appears to forecast a woman’s cardiovascular health in the years that follow, including premature onset before age 65 years of coronary artery disease<sup>5,6</sup>, heart failure and dysrhythmias<sup>7</sup>, and death after coronary revascularization<sup>8</sup>. Different guidelines for the prevention of CVD recommend screening for CVD risk factors in women with a prior maternal placental syndrome<sup>9</sup>.**”

→ The in the Discussion section, we changed to: “**Women with abnormal prenatal biochemical screening were found to be at modestly higher risk of a broad premature CVD composite outcome, largely arising before age 50 years, especially those with an elevated serum DIA.**”



→ To be clear, otherwise, it was not a requisite that the premature CVD outcome had to arise before a specific age cut-off. Rather, it is apparent from our study design, and duration of follow-up, that all (or nearly all CVD events), would arise in women in their 50's or earlier.

2. Did the authors adjust for BMI at prenatal screening?

→ We lacked BMI data but we did have maternal weight at the time of prenatal screening for 90% of pregnancies.

→ Additional analysis 1, Supplementary file 5, did adjust for weight, and the findings did not change.

→ We also state in the Discussion section "Still, about 10% of pregnancies lacked information on maternal weight, while height and menopausal status were entirely unknown."

## Results

1. How robust are these results? Did the authors consider undiagnosed CVD as partly explaining their findings? The authors should conduct a sensitivity analyses which excludes the first two years of follow-up.

→ We have performed several additional analyses, suggesting that the findings are quite robust. However, we see no reason to also evaluate the CVD outcome starting at 2 years after the index pregnancy. Already starting at 1 year enables any immediate or latent pregnancy factors to "wash out".

2. Consider doing a sensitivity analysis on exclusion of pre-existing renal disease, chronic hypertension, or dyslipidaemia, which are all risk factors for CVD. Exclusion is a much appropriate way of demonstrating robustness rather than adjustment.

→ We did this, as suggested. It is presented in the Results, and in Supplementary file 6. Doing so did not meaningfully change in the estimates.

3. The authors have a large and rich data source and should be able to conduct subgroup analyses by relevant characteristics such as age (average vs advanced maternal age), ethnicity (compare major ethnicities), type of pregnancy etc

→ We added a new subgroup analysis for the cardiac composite outcome by the co-presence of advanced maternal age  $\geq 35$  years (Additional analysis 4, Figure 1c).

→ We previously reported the results for the CVD composite outcome according to livebirth vs non-livebirth pregnancy outcomes (Additional analysis 4, Figure 1b, upper). However, there were too few events for us to report results for stillbirths, miscarriages and induced abortions separately.

→ For ethnicity, we always adjusted for this variable in all models. Yet, we did not believe that the addition of 60 more lines to Figure 1 (i.e., 6 ethnic group x 5 analytes x normal/abnormal state for each analyte) would enhance the paper, especially given low statistical power issues that might arise from some of the non-Caucasian group.

#### Reviewer: 4

1. However, some doubts should be addressed. My main criticism to the study is regarding to how the traditional risk factors were analyzed. Specific definitions of covariates (diabetes mellitus, chronic hypertension, dyslipidemia) should be showed in the main text. Also, the way used to analyze these variables in the regression model should be clarified (enter vs stepwise, enter and excluded p values, collinearity). Furthermore, showing more data of the regression model could be useful to the reader in order to understand the results. No data were provided regarding the treatment of traditional risk factors. In the same way, how women with previous cardiac, cerebrovascular or peripheral disease were identified should be specified.

→ **First, we show the definitions of every variable in a transparent manner, in Supplementary file 2. To add these to the text would require a lot of unnecessary space, and perhaps, distract the reader.**

→ **In the Statistical analysis section, we now clarify this point: “HRs were adjusted for variables chosen *a priori*, based on the existing literature<sup>3-7</sup>, including ...”**

→ **Certainly, we show the prevalence of the conventional CVD risk factors within Table 1, and also in Supplementary file 11. However, it is unconventional, and perhaps, distracting, to present all beta coefficients for all covariates.**

→ **In Supplementary file 2, we show all Exclusion criteria within  $\leq 5$  years before the maternal serum screening date in the index pregnancy, including previous cardiac, cerebrovascular or peripheral arterial disease.**

2. Regarding the clinical relevance of the findings, I disagree with the authors. In order to evaluate the utility of adding prenatal screening to the traditional risk evaluation, an analysis showing risk reclassification becomes necessary. In other words, how many women would change their risk level if the values of prenatal screening were added to conventional risk estimators?

→ **We completely agree with the reviewer. In the “Clinical policy and policy relevance section” we added: “Regardless, we require better data about whether prenatal biochemical screening offers additive information over that provided by conventional CVD risk factors<sup>4</sup> and adverse events in pregnancy<sup>5-8</sup>, such as the maternal placental syndromes or preterm delivery. One approach might be an analysis of risk reclassification, assessing the proportion of women who risk level changes if the values of prenatal screening are added to a list of conventional risk factors.”**

3. Finally, I would suggest showing the results using Kaplan-Meier curves as well (for example: no, 1, 2, 3... abnormal prenatal biochemical screening).

→ **We have numerous figures and supplementary files. This would mean creating multiple additional figures, and we are not sure they will add much to that shown in the existing Tables and Figure 1.**