Author’s response to reviews

Title: Maternal caffeine intake during pregnancy is associated with birth weight but not with gestational length: results from a large prospective observational cohort study

Authors:

Verena Sengpiel (verena.sengpiel@obgyn.gu.se)
Elisabeth Elind (elielind@online.no)
Jonas Bacelis (jonas.bacelis@gmail.com)
Staffan Nilsson (staffan@chalmers.se)
Jakob Grove (grove@hum-gen.au.dk)
Ronny Myhre (ronny.myhre@fhi.no)
Margaretha Haugen (margaretha.haugen@fhi.no)
Helle M Meltzer (helle.margrete.meltzer@fhi.no)
Jan Alexander (jan.alexander@fhi.no)
Bo Jacobsson (bo.jacobsson@obgyn.gu.se)
Anne-Lise Brantsæter (anne.lise.brantsaeter@fhi.no)

Version: 2 Date: 21 October 2012

Author’s response to reviews: see over
To The BioMed Central Editorial Team
Dr Lin Lee
BMC Medicine
BioMed Central
236 Gray's Inn Road
London WC1X 8HB
United Kingdom

Gothenburg, October 21st, 2012

Dear Dr Lin Lee,

Thank you for considering our paper “Maternal caffeine intake during pregnancy is associated with birth weight but not with gestational length: results from a large prospective observational cohort study” (MS: 4584083108059329) for publication in BMC Medicine. We want to thank Ina Santos and Alastair Hay for their valuable comments on our manuscript. All changes have been highlighted in the revised paper and a detailed response to all comments follows below.

Yours sincerely,

Verena Sengpiel, on behalf of all authors

Verena Sengpiel, MD, PhD
Dept of Obstetrics and Gynaecology
Sahlgrenska University Hospital/Östra
SE-416 85 Gothenburg
Sweden
Reviewer 1, Ina Santos:

Minor Essential Revisions:

1) Introduction, page 6, line 19 and 20: it seems to have an inconsistency between what the authors say on page 19 ("preterm delivery and to be small for gestational age at birth are “common” conditions) and on page 20 “low prevalences such as these... "). Common conditions should be frequent.

The passage has been changed, page 6 line 20: Despite these prevalences, the complexity makes it difficult to measure the effect of a single environmental factor, except in large studies.

Page 23 line 1: Due to the large study sample, there were 1,411 cases defined as spontaneous PTD and 852 (Marsal), 4,503 (Skjaerven) or 4,733 (Gardosi) cases of SGA in the study population.

2) Methods, page 11, lines 14-16: this sentence is not clear.

The paragraph has been changed, page 12 line 4: In Q1 and Q3, women reported their coffee, tea and caffeinated soft drink consumption in cups or glasses/d. These data allowed following a participant’s caffeine consumption from the three main caffeine sources from the time before pregnancy until gestational week 30.

3) Results, page 14, line 6: limits of the quartiles categories are confusing. Does the value 8.4 belong to the first or to the second quartile? Idem for value 40.7 in the second and third quartiles.

Now, the more exact instead of the rounded data are presented, page 14 line 9: no intake and for the remaining subjects quartiles 0-8.38, 8.39-40.71, 40.72-110.52, >110.52 mg/d.
4) References, page 39, reference #62: authors’ names are lacking.


5) Table 2: the value 24.9 is present in two categories of the BMI variable.

The borders are defined as <18.5, 18.5-24.9, 25-29.9 and ≥30 kg/m² and have been corrected in the table.

6) Idem for values 7.9, 9.4, and 11.1 of the variable “Quartiles of energy intake in MJ” (Table 2).

The exact borders are <7.90, 7.90-9.35, 9.36-11.14 and >11.14 MJ and have been corrected in the table.

7) Idem for caffeine intake groups on Table 7.

The exact borders are 0-50, 51-200, 201-300 and >300 mg/d and have been corrected in the table.

8) Because the infant birth weight outcome was calculated as the percentage of the expected birth weight (3,600 g), it is not clear whether the results shown on Table 5 are controlled for gestational age at birth or whether only term infants were included in that analyses.

The analysis was performed for all 59,123 participants in the study including both spontaneous and iatrogenic as well as preterm and term deliveries. Analyses were not adjusted
for gestational age because gestational age was part of the outcome variable due to its
definition as percentage of the expected birthweight (BW). The expected BW was defined
according to Marsal, Skjaerven and Gardosi, and all models give the expected BW for a
certain gestational age.

We tried to clarify the procedure in the Methods section page 10 line 13: Difference between
BW and expected BW was calculated as percentage of expected BW. This implicates that
gestational length is taken into account by the definition of our outcome variable and analysis
were not adjusted for gestational age.

9) Table 7 should present p-value for the entire variable (“total caffeine intake
groups”) in each column, instead of the p-values for each category of the
variable.

The idea with Table 7 is to give some risk estimate for women following the different
recommendations. It presents the results of a Cox regression comparing the effect of these
different caffeine intake categories to the lowest caffeine intake category of 0-50 mg/d of total
caffeine. As such, the Cox regression only provides p-values comparing the different
categories to the reference category.

When adding the categorical total caffeine intake variable in a second step after adjusting for
all other covariates before, the model improves with a p= 0.0008 for SGA (Marsal), p <10^{-11}
SGA (Skjaerven) and p <10^{-7} SGA (Gardosi). However, presenting that there is an association
of overall caffeine intake and birthweight was not the intention with this table but table 5.
Reviewer 2, Alastair Hay:

Before dealing with Alastair Hay’s valuable remarks in detail, we want to address his general remarks in pointing out that the only association with increased PTD risk was found for black tea caffeine in the subgroup of early spontaneous PTD. With a p-value of 0.01- much less extreme than for the association with SGA - confounding cannot be ruled out as an explanation in a big study like this.

Major compulsory revision

1) The above findings are important and confirm those of a study published in 2008 (the CARE study, ref 62 in this manuscript). Like the CARE study the authors of this manuscript validated their tool for assessing caffeine intake. However, unlike the CARE study group which had published detailed comparisons of caffeine intake from different sources, including use of biomarkers to confirm consumption, the paper the current authors refer to as validation for caffeine intake merely refers to whether intake was correctly classified. It is thus not possible to make a detailed assessment of their measure of caffeine intake and this is something which needs addressing, perhaps by inclusion of an additional table or figure showing how well the questionnaire (used in this study) performed in relation to weighed food intakes. In fairness the authors do claim that there is a good agreement between their food frequency questionnaire and coffee intake. This being the case they should show the data as the measure of intake is crucial as referred to earlier.

We agree that careful measurement and validation of caffeine intake is important. The MoBa FFQ included detailed questions about different types of coffee and distinguished between cola-drinks, energy drinks and other soft drinks. The validation study did not include evaluation of agreement for estimated caffeine intake (unlike the CARE study), but focused on food groups including the main caffeine sources. On page 22 in the discussion we did report several correlation coefficients reflecting the agreement between our questionnaire and
a weighed food diary and biological markers. The references cited under Methods and Discussion also included correlation coefficients, and not only whether intake was correctly classified.

In order to address the suggestion by Alastair Hay to report how well the questionnaire performed in relation to weighed food intakes, we have now calculated caffeine intake by the FFQ and the food diary for the 119 women in the validation study. High agreement was observed for total caffeine \((r=0.70, \text{CI}: 0.59-0.78)\) as well as for caffeine from coffee \((r=0.80, \text{CI}: 0.72-0.86)\) and tea \((r=0.47, \text{CI}: 0.32-0.60)\). The paragraphs about validation have been revised in the Methods and Discussion section.

As additional file we offer a Bland-Altman plot showing the differences in caffeine intake between the FFQ and food diary measurements (bias) against the mean caffeine intake by the two methods showing that the mean difference was small and not biased towards any of the methods (Additional file 1).

Page 11, line 16: The FFQ has been extensively validated in a MoBa sub-population \((n=119)\) using a four-day weighed food diary and biological markers in blood and urine as reference measures [42, 43]. The validation study showed that the MoBa FFQ is a valid instrument for assessing habitual diet during the first four to five months of pregnancy. The agreement between the FFQ and the food diary was particularly high for coffee \((r=0.80, \text{CI}: 0.72-0.86)\), and was high for tea \((r=0.53, \text{CI}: 0.39-0.65)\) and soft drinks \((r=0.48 \text{ CI}: 0.33-0.61)\). Estimated caffeine intake was not evaluated at the time, but when caffeine concentrations (Table 1) were combined with consumption data for women in the validation study, high agreement was observed between the FFQ and the food dairy for total caffeine \((r=0.70 \text{ CI}: 0.59-0.78)\). The median (IQR) caffeine intake in the validation study sample was 40 mg/d (18-88 mg/d) by the FFQ and 38 mg/d (10-99 mg/d) by the four-day weighed food diary. Caffeine from coffee and tea showed similar high agreement as for the beverages, while poorer agreement was seen for caffeine from chocolate \((r=0.20, \text{CI}: 0.02-0.36)\). No participants in the validation study had intake of caffeine from soft drinks. Food items like soft drinks, chocolate and sweets are more likely to be misreported than most other food items.
Page 23, line 15: Relying exclusively on self-reported data, without a biological marker to confirm the accuracy of estimated caffeine exposure is a weakness. For the present study we evaluated the agreement between caffeine intake estimated by the FFQ and a food diary. The estimated caffeine intake did not differ between the methods, and a high correlation was observed (r=0.70 CI: 0.59-0.78). Furthermore, the MoBa FFQ has been extensively validated in a MoBa sub-population using the four-day weighed food diary and several biomarkers as reference measures [42, 43]. The agreement between the FFQ and the food diary was particularly high for coffee and tea, which are the main sources of caffeine in this study.

Coffee intake according to both the FFQ and the food diary correlated with serum beta-carotene (0.31 and 0.36, respectively), which can be explained by interplay between antioxidants in coffee with beta-carotene, as also reported by Svilaas et al. [66]. Likewise, tea intake according to both the FFQ and the food diary correlated with kaempferol, a flavonoid found in tea (r=0.41 and 0.50 for the FFQ and food diary, respectively [42]). Similar, but slightly weaker correlations were observed for estimated caffeine contributed by coffee and tea. A Bland-Altman plot for the differences in caffeine intake between the FFQ and the food diary is available as Additional file 1.
Minor essential revision

As I interpret the data the authors found no threshold of effect for total caffeine intake for their SGA group (small for date) with assessments using all three variations on defining SGA. Their data on birth weight is less easy to interpret but it would be helpful if the authors were to say whether they observed any threshold effect for this outcome.

We tested for a threshold effect by applying the same regression model for sixtiles of total caffeine intake finding increased odds ratios even for the lower sixtiles, results are presented in Figure 4.

Page 16, line 16: To test if there was a threshold effect, we performed the same logistic regression with sextiles of total caffeine intake (0-14.645, 14.646-32.093, 32.094-57.265, 57.266-96.029, 9603-163.806, >163.806 mg/d). In all three models the caffeine intake categories were associated with increasing odds ratios for SGA as compared to the lowest intake group, Figure 4.

Page 22 line 12: We could not find a threshold for the association of caffeine consumption and SGA risk. Until there is clarity if there is a causal association between caffeine intake and increased risk for SGA, women might be advised to reduce their caffeine consumption as much as possible during pregnancy.
As the findings of this study and the earlier CARE study are similar it has implications for regulators. Advice to mothers to be in the UK changed following publication of the CARE study with pregnant women advised to limit caffeine intake and to stay below 200mg per day. The CARE study did not identify a threshold for reduction in fetal growth restriction either. Taken together with this study even more cautious advice for pregnant women may be necessary.

We do agree, that the results from this comprehensive study, especially if taken together with the findings from the CARE Study ask for further investigation of this association. Until there is clarity if there is a causal association between caffeine intake and increased risk for SGA, women might be advised to reduce their caffeine consumption as much as possible during pregnancy. We aimed to stress this in several places throughout the discussion:

Page 21 line 24: Especially the data from some of the largest observational studies published so far, are consistent with our findings, moreover with comparable effect size of a decrease in BW by 60-70 g for >200 mg/d [45] or 28 g for 100 mg/d caffeine consumption [50].

Page 22, line 9 However, as our results confirm earlier findings [45] that the increase in SGA risk can already be found in women following the current recommendations by Norwegian Authorities, further studies are needed to establish the impact of caffeine on neonatal morbidity and mortality. Until there is clarity if there is a causal association between caffeine intake and increased risk for SGA, women might be advised to reduce their caffeine consumption as much as possible during pregnancy.

Page 25, line 11: As the risk for SGA increases even if pregnant women follow official recommendations in Norway of a maximum caffeine intake of 200 mg/d, this association should be further investigated and recommendations might have to be re-evaluated.
Discretionary revisions

This is a valuable addition to the literature on the effect of caffeine in pregnancy. The major findings in this manuscript of an effect of caffeine on birthweight were identified in the earlier CARE study which used a well validated measure of caffeine intake. It would help to stress this in the paper if only to indicate to others contemplating similar investigations that intake must be measured carefully.

We totally agree that a careful measurement of total caffeine intake from different sources is a must when studying the effect of caffeine intake on pregnancy outcomes and tried to emphasize it even more throughout the discussion:

Page 18, line 9: Thus, findings attributed to increasing total caffeine intake might be due to a changing distribution of caffeine sources. These results emphasize what Peck et al. and the CARE Study group pointed out: if the aim of an epidemiologic study is to assess the effect of caffeine, it is not correct to study only coffee caffeine [20, 45].

Page 20, line 1: Caffeine from several sources was assessed, rather than caffeine intake from a single source, usually coffee, as in many studies [30, 53-55]. For some populations using only caffeine from coffee would implicate that a major – if not the major – part of total caffeine intake was not considered at all, e.g. in a UK-study black tea contributed with 62% to daily caffeine intake while coffee and cola drinks accounted for about 12-14% each [45].