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

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Pregnancy associated cancers and birth outcomes in children: a Danish and Swedish population-based register study</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Momen, Natalie; Arendt, Linn; Ernst, Andreas; Olsen, Jørn; Li, Jiong; Gissler, Mika; Ramlau-Hansen, CH</td>
</tr>
</tbody>
</table>

VERSION 1 – REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Marina Vafeiadi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>03-Apr-2018</td>
</tr>
</tbody>
</table>

GENERAL COMMENTS

Manuscript ID bmjopen-2018-022946

**Title:** "Pregnancy associated cancers and birth outcomes in children: a population based register study"

This manuscript aims to explore the association between pregnancy-associated cancers and birth outcomes. The study is well-designed and the results are of public health relevance. Below are some minor comments and/or suggestions.

1. A “Methods” and “Background” section is missing from the abstract.

2. Abstract, lines 27-29 should be deleted, the outcomes are already mentioned above

3. Abstract, the first line in the results section should be re-written and “exposed” should be explained.

4. The introduction section is very short, authors should expand this part describing any previous work that has
been done on this field and possible underlying mechanisms

5. Methods, page 7, lines 12-19, CS should be included in the outcomes list and relevant categories should be explained.

6. Methods, page 7, lines 19-29, there is no need to repeat all outcomes, just provide the source of the data and the dates.

7. More details on the procedure used to identify confounders should be given in the statistical analysis section.

8. Results, page 10, lines 15-25 and Table 1, were any of these differences statistically significant?

9. Results, page 15 and Table 2, the number of subjects in each specific maternal cancer should be given in order to better understand the importance of these findings.

10. Discussion, page 20, lines 5-6, this sentence should be moved to the limitations section.

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Frédéric Amant</th>
</tr>
</thead>
<tbody>
<tr>
<td>KU Leuven, Belgium</td>
<td></td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>12-Jun-2018</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GENERAL COMMENTS</th>
<th>Pregnancy associated cancers and birth outcomes in children: a population based register study.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The authors aimed to estimate the association between pregnancy-associated maternal cancers (diagnosed both prenatally and postnatally) and birth outcomes. Therefore, they used existing data from different national registers of Denmark and Sweden. An important strength of the study is the large population used. However, I have some comments/questions:</td>
</tr>
</tbody>
</table>
Abstract (p2-3)
- P2, line 10: “Population-based register study” instead of “Population based register study” (change this everywhere).
- P2, line 27-29: When describing the primary and secondary outcomes the article repeats the outcomes with a new sentence at the end: “Gestational age, birth weight...”. This should be left out. (repetition)
- P2, line 40: For postnatal diagnoses: ”some (associations) remained statistically significant”. It is unclear which ones are significant and it is interesting to have this information in your abstract. I propose that you clearly write down which associations are significant.
- P2, line 43: “results suggest an association between..” instead of “results suggest an association exists between...”.
- P2, line 50-52: I suggest to leave out “(as well as stress, medical intervention or treatment)”. This does not add that much information relevant to the abstract. This article also does not give that much information about the effects of stress, medical interventions on fetal development and these effects are still investigated to a wide extent.
- P3, line 3-5: It is unclear what you mean with “to include some postnatal years”, something is wrong with this sentence. Do you mean “by including some postnatal years”?

Article summary (p3)
- P3, line 14: “makes use of Danish and Swedish...” instead of “makes use of the Danish and Swedish...”
- P3, line 26: “exposure window and investigate” instead of “exposure window further and investigate”.
- P3, line 31-33: “needs to be” instead of “need to be”, and “their impact” instead of “their impacts”.

Introduction (p4-5)
- P4, line 39: “population-based register study” instead of “register-based study” (be consistent).
- P4, line 55- P5, line 3: In the hypothesis you write about “the effects of both the cancer itself and its treatment” and “maternal stress or lifestyle changes following diagnosis influence birth outcomes”. However you do not support this information with references. I propose you add this to your introduction. This also for information about CS.
- P5, line 5: “observed risk should be limited to children of ...” instead of “observed risk should be limited to be limited to children of ...” (repetition)
Material and methods (p6-9)

Ascertainment of maternal cancer diagnoses (p6-7)

- P6, line 40-54: You include maternal cancers diagnosed “up to 15 years postpartum”, but you do not elaborate why you choose this specific timeframe. Only at the end of the paper you give a brief explanation (P21). It is best to already explain it here and why specifically up to 15 years after pregnancy. This is unclear. It is also not stated why you choose the frames <6 years postpartum and 6-15 years postpartum. Why after 6 years and not for example 5 years?
- Was there any data in the registers about cancer treatment during pregnancy? This could also be an interesting factor to look at.
- Are there any exclusion criteria beside the time of diagnosis? They are not well-defined. The women of whom a follow-up is missing until 15 years after delivery are for example a group a women that can be excluded from the data.

Covariates (p7-8)

- P7, line 53: You have 3 groups for maternal age, why these specific ones? And why the oldest group already starts at the age of 31?
- P7, line 56: Why is the year of birth of the child included? And why is it included as a continuous variable and not for example divided in groups? This variable is only mentioned here. In all the other sections of the article this covariant is not mentioned anymore and it is unclear what the associations are.
- P7, line 56: The article describes 3 possibilities for highest education at time of childbirth “<14 years, >15 years, missing”. However, in Table 1 (p11) you describe 4 groups: Low, <9 years; Middle, 10-14 years; High, >15 years; and Missing. Change your groups in the covariate section, so this is consistent.
- P8, line 5: Multiplicity is an interesting covariant, but there are also no results reported about the associations with this covariant.
- There are other interesting covariates that may possibly influence the results (e.g. other maternal illnesses during pregnancy, lifestyle factors such as smoking/ drugs/ medication during pregnancy, cancer treatment, parity etc.). The register may not provide information about all these factors, but did you consider taking them into account? They can have an important impact on the results.

Because so few possible confounders are taken into account in this study, it is hard to make hard conclusions with this study design. A case-control study comparing the birth outcomes of women with or without a cancer diagnosis, by matching for possible confounders would be interesting.
- In your table with descriptive characteristics (Table 1, P11) you mention parity as characteristic, but you never mention this variable in your Method section.

Results (p10-18)
Characteristics of the study population (p10)
- I think it can be interesting to give information about the numbers of people categorized by specific maternal cancer type (e.g. 1000 women with a breast cancer diagnosis etc.). Especially since you differentiate these types of cancer in Table 2 (P16-17). This enables the reader to know how many woman are in each group and helps the reader for example to interpret the results carefully if it is a very uncommon cancer diagnosis in the used population.

Associations between any type of maternal cancer and adverse birth outcomes
- Figure 1, P27: There is no reference to the unexposed group being the reference category for these 3 groups. Can you add this in the title or as a remark?
- P13, line 28: The article indicates that the increase in risk is “borderline, statistically significant”, but it is unclear when a RRR is considered as significant or not in this study (nor here, nor in the method section). What value did you take as being statistically significant (and for Borderline)? It is best to mention this somewhere.
- P13, line 51: We observed that point estimates…” instead of “We observed point estimates went towards..”
- P14, line 3-5: The construction of this sentence seems to be incorrect. I think you should leave out “and those whose mothers were diagnosed 6-15 years postpartum” at the end of the sentence. I think it is a repetition.
- P14, line 11: “was an increased risk of …” instead of “were increased risks of…”

Associations between types of cancer and adverse birth outcomes
- P15: Do I understand it correctly that the exposed group is all children of women with a cancer diagnosis 2 years before pregnancy up to 15 years after pregnancy? Maybe this can be added to the text? Is it not interesting to do analyses for the different periods of diagnoses too? Definitely given that the latency periods to clinical diagnoses for cancers can be different related to type of cancer? You can add these results to your Table 2.
- P15: Why no 95% CI’s here? (be consistent)
- P15, line 11: “statistical significance”: when statistically significant?
- P17: The upper part of Table 2 is missing (above ‘Endocrine’).
- P18: Can this paragraph be moved to P15?
Discussion (p19-23)
Interpretation (p19-20)
- P43-49: You talk about the relation between early CS, preterm birth, and birth weight. Maybe include references to support this? The relationship between preterm birth and low birth weight is quite straightforward. Could this also be controlled for in analyses?
- P20, line 5: The article talks about “residual confounding”. This is good because many possible influencing factors are not taken into account. Is it possible to give an example for the readers? It is also advisable to include this in the ‘Strengths and limitations’-section of this paper.
- P20, line 37-41: Some or your findings confirm the results of previous research, some results are different. It not clear when your findings confirm or contradict previous findings. Can you state this more clearly and if the results are different, can you try to suggest a reason? This should be elaborated.
- P20, line 52-57: The sentence “However, this is not proof…” is quite suggestive. We do not know. No research findings about this.

Strengths and limitations (p21-22)
- What about potential biases? Generalizability?
- I miss more recommendations for future research!

**VERSION 1 – AUTHOR RESPONSE**

Reviewer 1

1. A “Methods” and “Background” section is missing from the abstract. We have followed the BMJ Open guideline for abstracts, which does not include these sections.

2. Abstract, lines 27-29 should be deleted, the outcomes are already mentioned above Amended as suggested.

3. Abstract, the first line in the results section should be re-written and “exposed” should be explained. Amended. This now reads: “In this study 2% of children were born to mothers with a diagnosis of cancer’exposed’.”

4. The introduction section is very short, authors should expand this part describing any previous work that has been done on this field and possible underlying mechanisms. We have lengthened the introduction with information on previous work and potential mechanisms, as suggested.

5. Methods, page 7, lines 12-19, CS should be included in the outcomes list and relevant categories should be explained. Amended as suggested.

6. Methods, page 7, lines 19-29, there is no need to repeat all outcomes, just provide the source of the data and the dates Amended as suggested.
7. More details on the procedure used to identify confounders should be given in the statistical analysis section. This has been amended to mention the use of directed acyclic graphs to identify confounders.

8. Results, page 10, lines 15-25 and Table 1, were any of these differences statistically significant? We have added a sentence on page 10 about statistical significance in the description of characteristics.

9. Results, page 15 and Table 2, the number of subjects in each specific maternal cancer should be given in order to better understand the importance of these findings. To avoid making the table difficult to read, we have added number of children exposed to specific cancer groups in the text (page 10, paragraph 2).

10. Discussion, page 20, lines 5-6, this sentence should be moved to the limitations section. Amended as suggested.

Reviewer 2

Abstract (p2-3)
- P2, line 10: “Population-based register study” instead of “Population based register study” (change this everywhere). Amended as suggested.

- P2, line 27-29: When describing the primary and secondary outcomes the article repeats the outcomes with a new sentence at the end: “Gestational age, birth weight…”. This should be left out. (repetition) Amended as suggested.

- P2, line 40: For postnatal diagnoses: “some (associations) remained statistically significant”. It is unclear which ones are significant and it is interesting to have this information in your abstract. I propose that you clearly write down which associations are significant. Amended to include the significant associations as suggested.

- P2, line 43: “results suggest an association between..” instead of “results suggest an association exists between…”. Amended as suggested.

- P2, line 50-52: I suggest to leave out “(as well as stress, medical intervention or treatment)”. This does not add that much information relevant to the abstract. This article also does not give that much information about the effects of stress, medical interventions on fetal development and these effects are still investigated to a wide extent. Amended as suggested.

- P3, line 3-5: It is unclear what you mean with “to include some postnatal years”, something is wrong with this sentence. Do you mean “by including some postnatal years”? Amended. The sentence now reads:

Future studies on maternal cancer during pregnancy should consider including some postnatal years in their exposure window.
Article summary (p3)
- P3, line 14: "makes use of Danish and Swedish..." instead of "makes use of the Danish and Swedish...". Amended as suggested.

- P3, line 26: "exposure window and investigate" instead of "exposure window further and investigate". Amended as suggested.

- P3, line 31-33: "needs to be" instead of "need to be", and "their impact" instead of "their impacts". Amended as suggested.

Introduction (p4-5)
- P4, line 39: "population-based register study" instead of "register-based study" (be consistent). Amended as suggested.

- P4, line 55- P5, line 3: In the hypothesis you write about "the effects of both the cancer itself and its treatment" and "maternal stress or lifestyle changes following diagnosis influence birth outcomes". However you do not support this information with references. I propose you add this to your introduction. This also for information about CS. An additional paragraph containing information about potential mechanisms and references has been added (page 4, paragraph 3).

- P5, line 5: "observed risk should be limited to children of …" instead of "observed risk should be limited to be limited to children of …" (repetition) Amended as suggested.

Material and methods (p6-9)
Ascertainment of maternal cancer diagnoses (p6-7)
- P6, line 40-54: You include maternal cancers diagnosed “up to 15 years postpartum”, but you do not elaborate why you choose this specific timeframe. Only at the end of the paper you give a brief explanation (P21). It is best to already explain it here and why specifically up to 15 years after pregnancy. This is unclear. It is also not stated why you choose the frames <6 years postpartum and 6-15 years postpartum. Why after 6 years and not for example 5 years? The decision to follow up 15 years postpartum, with further analysis looking at frames <6 years postpartum and 6-15 years postpartum, was arbitrary. As latency times vary and we cannot know when a cancer was initiated, we chose follow-up times we thought to be reasonable. Division of the exposure window was used to see if any observed increase in risk would return to normal when it became less likely that the cancer had been present during pregnancy/advanced enough during pregnancy to affect the fetus. We have added to the text to state as such.

- Was there any data in the registers about cancer treatment during pregnancy? This could also be an interesting factor to look at. We do not have access to this data, however it is available. The heterogeneity of treatment within different types of cancers and across different types of cancers would mean numbers were very small and would these analyses underpowered. Further, type of treatment for different cancers have changed dramatically over time. However it would be an interesting factor to look at and could be considered in the future.

- Are there any exclusion criteria beside the time of diagnosis? They are not well-defined. The
women of whom a follow-up is missing until 15 years after delivery are for example a group a
women that can be excluded from the data. There were no exclusion criteria – all children who were born in Denmark and linked to their mother
were included. Exclusion of children who did not have a full 15 years of follow-up before the end of
the data was briefly considered, but it would exclude a large proportion of the children (all those born
from 1991 in Sweden and 1993 in Denmark), so it was decided against. We recognize that this may
lead to misclassification (some children who were exposed would be considered unexposed) and
have included this as a limitation in the discussion (page 23, paragraph 3).

Covariates (p7-8)
- P7, line 53: You have 3 groups for maternal age, why these specific ones? And why the oldest
group already starts at the age of 31?
These were chosen a priori. They are groupings often used in reproductive epidemiology in the
Danish register data and provide fairly evenly sized groups.

- P7, line 56: Why is the year of birth of the child included? And why is it included as a
continuous variable and not for example divided in groups? This variable is only mentioned
here. In all the other sections of the article this covariant is not mentioned anymore and it is
unclear what the associations are.
Birth year was included as it is expected to be associated with some of the birth outcomes.
Additionally, maternal cancer diagnoses may be associated with birth year; in particular, children born
later in the study period, who cannot be followed-up for a full fifteen years, would be less likely to be
exposed. However, this has now been changed to a categorical variable and included in table 1, to
show its relation with exposure. Results have been updated.

- P7, line 56: The article describes 3 possibilities for highest education at time of childbirth
"<14 years, >15 years, missing". However, in Table 1 (p11) you describe 4 groups: Low, <9
years; Middle, 10-14 years; High, >15 years; and Missing. Change your groups in the
covariate section, so this is consistent.
This has been amended.

- P8, line 5: Multiplicity is an interesting covariant, but there are also no results reported about
the associations with this covariant.
It is displayed in table 1, however it has been amended from “singleton” so that the language used is
consistent

- There are other interesting covariates that may possibly influence the results (e.g. other
maternal illnesses during pregnancy, lifestyle factors such as smoking/ drugs/ medication
during pregnancy, cancer treatment, parity etc.). The register may not provide information
about all these factors, but did you consider taking them into account? They can have an
important impact on the results.
Because so few possible confounders are taken into account in this study, it is hard to make
hard conclusions with this study design. A case-control study comparing the birth outcomes of
women with or without a cancer diagnosis, by matching for possible confounders would be
interesting.
One of the limitations of register based research is that relevant confounders are not always available.
We have expanded the discussion section to include this as a limitation (page 23, paragraph 5).

- In your table with descriptive characteristics (Table 1, P11) you mention parity as
characteristic, but you never mention this variable in your Method section.
Parity has now been added to the regression models.

Results (p10-18)
Characteristics of the study population (p10)
- I think it can be interesting to give information about the numbers of people categorized by
specific maternal cancer type (e.g. 1000 women with a breast cancer diagnosis etc.).
Especially since you differentiate these types of cancer in Table 2 (P16-17). This enables the
reader to know how many woman are in each group and helps the reader for example to
interpret the results carefully if it is a very uncommon cancer diagnosis in the used population.
We have added number of children exposed to specific maternal cancer groups in the text (page 10, paragraph 2).

Associations between any type of maternal cancer and adverse birth outcomes
- Figure 1, P27: There is no reference to the unexposed group being the reference category for these 3 groups. Can you add this in the title or as a remark?
This has been added to the text, as suggested.

- P13, line 28: The article indicates that the increase in risk is “borderline, statistically significant”, but it is unclear when a RRR is considered as significant or not in this study (nor here, nor in the method section). What value did you take as being statistically significant (and for Borderline)? It is best to mention this somewhere.
A 95% confidence interval which did not include the value of 1 (the null hypothesis value) was taken as statistically significant. This is a standard way to consider statistical significance, but an addition was made to the text (page 8, paragraph 2) to clarify this. References to “borderline” significance have been removed as it is less objective.

- P13, line 51: We observed that point estimates…” instead of “We observed point estimates went towards…”
Amended as suggested.

- P14, line 3-5: The construction of this sentence seems to be incorrect. I think you should leave out “and those whose mothers were diagnosed 6-15 years postpartum” at the end of the sentence. I think it is a repetition.
This has been amended.

- P14, line 5: The supplementary Figure 1 is missing in the article, I can only find the supplementary figure 2.
Supplementary figure 1 has been uploaded.

- P14, line 11: “was an increased risk of…” instead of “were increased risks of…”
Amended as suggested.

Associations between types of cancer and adverse birth outcomes
- P15: Do I understand it correctly that the exposed group is all children of women with a cancer diagnosis 2 years before pregnancy up to 15 years after pregnancy? Maybe this can be added to the text?
Amended as suggested.

Is it not interesting to do analyses for the different periods of diagnoses too? Definitely given that the latency periods to clinical diagnoses for cancers can be different related to type of cancer? You can add these results to your Table 2.
This has been included, but in a supplementary table. Numbers of exposed cases are quite small, especially for prenatal diagnoses, where there are sometimes no exposed cases or numbers are not high enough to report (as Statistics Denmark’s reporting regulations were not met).

- P15: Why no 95% CI’s here? (be consistent)
95% CIs have been added, as suggested

- P15, line 11: “statistical significance”: when statistically significant?
As above, a 95% confidence interval which did not include the value of 1 (the null hypothesis value) was taken as statistically significant. An addition was made to the text (page 8, paragraph 2) to clarify this.

- P17: The upper part of Table 2 is missing (above ‘Endocrine’).
This is a large table which spans two pages and seems to be submitted correctly. The top line on page 2 of the table is the upper confidence interval for CNS cancers.

- P18: Can this paragraph be moved to P15?
The text has been edited to allow this.

Discussion (p19-23)
Interpretation (p19-20)
- P43-49: You talk about the relation between early CS, preterm birth, and birth weight. Maybe include references to support this? The relationship between preterm birth and low birth weight is quite straightforward. Could this also be controlled for in analyses? CS, preterm birth and birth weight have not been included as covariates in analyses where others have been the outcome, because they are likely to be mediators on the causal path from the exposure to the outcomes of interest and should therefore not be adjusted for.[1,2]

- P20, line 5: The article talks about “residual confounding”. This is good because many possible influencing factors are not taken into account. Is it possible to given an example for the readers? It is also advisable to include this in the ‘Strengths and limitations’-section of this paper. We have expanded on our discussion of residual confounding around the examples of maternal smoking during pregnancy and maternal nutrition (page 23, paragraph 5).

- P20, line 37-41: Some of your findings confirm the results of previous research, some results are different. It not clear when your findings confirm or contradict previous findings. Can you state this more clearly and if the results are different, can you try to suggest a reason? This should be elaborated. We have made some additions to this paragraph to make it clearer whether findings confirm or contradict previous findings, and to discuss these further.

- P20, line 52-57: The sentence “However, this is not proof…” is quite suggestive. We do not know. No research findings about this. This has been amended make it less suggestive.

Strengths and limitations (p21-22)
- What about potential biases? Generalizability?
   We have discussed some potential sources of bias e.g. misclassification, live birth bias. The study includes all children born in Denmark and Sweden, so findings may be generalizable to other countries with similar population/healthcare setting etc to those of Denmark and Sweden, however findings may be different in studies of this association in different populations.

- I miss more recommendations for future research!
   Some more recommendations for future research have been added throughout the discussion. Specifically, we suggest future research could look at treatment during pregnancy, and consider the postpartum period further and give more precise consideration to latency periods.
