31/3/2018

Response to Editors and Reviewers: - Manuscript ID BMJ.2018.043336 entitled *"Risks of ovarian, breast and corpus uteri cancer in women treated with assisted reproductive technology; 2.2 million person years of observation in Great Britain"*

Dear Dr Weber and colleagues,

Thank you very much for considering our above paper for publication in the BMJ.

In the table below we have attempted to address the reviewers concerns and where appropriate made alterations to our manuscript, also detailed in the table below. We also attach our revised manuscript.

Once again thank you for your time in considering our submission,

Yours,

Professor Alastair Sutcliffe, Dr Carrie Williams and colleagues

Comment		Comment	Author Response
From	**No.**		
Manuscript committee meeting	1	We thought your study addresses an interesting and potentially important research question	Thank you
	2	There's a recent and very large systematic review you might want to cite (ref included)	Thank you for highlighting this. This reference has now been included. The introduction has now been re-worked as suggested by reviewer 4.
	3	You use "old fashioned" statistical methods comparing observed rates with national rates to compute standardised incidence ratios. One of the reviewers thinks this is inappropriate and Cox models should be used. We don't think that it would make much of a difference to the findings – maybe change the magnitude of the relationships a little, but not much.	If we had access to an appropriate comparison group from the general population, with individual-level information on potential founding factors, we agree that a Cox regression or survival analysis (as suggested by Referee 4), [or matching to comparable controls for a nested case-control approach (as suggested by Referee 2)] may be appropriate. However, a suitable external comparison group with measurements on potential confounding factors was not available, nor was it within the scope of this study to assemble one. Therefore we used a standard method to compare observed rates with national population rates to compute standardised incidence ratios (SIRs). An advantage of this method is that we were able to present 'exact' 95% confidence intervals for SIRs by reference to the Poisson distribution.
	4	You make no adjustments for confounding – so difficult to draw causality conclusions, but you are careful in not making them, just describing the observed risk relationships.	As above, given that there was no available unexposed cohort with individual level confounders, we were unable to make specific adjustments for such confounding factors. However, we are able to stratify our analysis by confounding factors, providing more information about specific risks. As mentioned, we are careful not to over-interpret results.
	5	The absolute risk is not explained well. Errors in the confidence interval calculation and not reported widely.	Further clarification about absolute risk calculation has been added (Statistical analysis paragraph). Further absolute risk have been added and all related CI's are now correct.

	6	Tables could be better organised.	We have reviewed the tables and cannot find a suitable way of re-organizing these without losing information. However, we would be happy to consider any suggestion of how this might be done.
Reviewer 1	7	For the analyses on invasive and borderline ovarian tumors, the authors should also report results for sensitivity analyses restricting to the cohort to male factor as reason for infertility.	As suggested we have added results for male factor infertility to the text (pages 11 & 12). The results were originally only presented in table 4.
	8	Others confounders such as socio-economic status, co-morbidities or co-medication would have been of interest to control for confounding. This might be preferably analyzed using a conventional cohort/nested case-control design where women who underwent assisted reproductive technology are e.g. age- or frequency-matched to comparable controls. Corresponding multivariable regression analyses could be used to control for further confounders here. However, I assume that such potential data was not available for analysis?	Unfortunately as the reviewer suggests, appropriate data were not available for the further analysis suggested. This limitation is mentioned in the discussion (Strengths and Limitations, paragraph 1).
	9	The authors should explain a bit more in detail what they mean with „Sensitivity analyses excluded the first 12 months of follow-up“. Does that mean that they restricted the cohort of women who underwent assisted reproduction to the ones who were cancer free at least for the first 12 months after the first cycle?	Yes, the reviewer is correct in interpreting this phrase. This type of sensitivity analysis is used in number of similar studies as it reduces the risk of surveillance bias. We have added a clarification as a footnote to tables S3, S4, S5.
	10	Page 6: „Overall 7% of women undergoing assisted reproduction, 1991-2010, were not included in this study, representing a loss of less than 1% of person-years follow-up“. Why were they not included? Due to missing values for important variables or because they refrained to give their consent?	They were not included due to lack of consent, as indicted in the previous lines. This has been clarified further (STUDY POPULATION, paragraph 2).

	11	Statistical analysis, first sentence, page 7: „Person-years at risk were calculated from date of first treatment…". Maybe change „Person-years….." to „Follow-up was calculated from…" to make it more consistent throughout the manuscript.	Thank you. Changes have been made (Statistical analysis- first sentence).
Reviewer 2	12	The subgroup analyses depend on the accuracy of clinical diagnoses of causes of infertility. Approaches to diagnosis have relaxed over the duration of the study as the inexorable advance of IVF as a solution to all causes of infertility has reduced the clinical need for investigative rigour. For example, laparoscopy is performed far less frequently than in the early part of the century so diagnosis of endometriosis (an important factor linked to diagnosis of ovarian cancer in this study) is made less often and on the basis of soft clinical signs than previously. This point is not adequately addressed in the discussion.	Thank you. Additions have been made to the discussion (Strengths and limitations paragraph). We have stated that we have no data on how diagnoses were made but have not gone any further with this statement as we have no specific evidence about the rigour with which diagnoses were made over the study period. However we agree with your statement.
	13	Ref 3 describes an effect of oral clomifene only – not injectable FSH as is used in >90% of IVF cycles. This should be mentioned.	Changes have been made to reflect this (Introduction, paragraph 2).
	14	Comparison data are only available in 5 year age bands. Patient data are presented per year. Could the authors be more clear about how they corrected for this potential mismatch?	The comparison data (rates from the general population) were available in 5-year age bands. Patient information was available in finer detail. As is standard for this type of analysis we grouped patient age into corresponding five-year age groups matching the same age grouping of the comparison data. This is described in the Statistical Analyses.
	15	Are data available for pregnancy as well as livebirth. Was the protective effect of livebirth also observed for other pregnancies.	Self-reported data about pregnancy were also available. The protective effect of live birth on all ovarian cancer and invasive ovarian cancer was also seen for pregnancy, but to a lesser degree.

<table>
<tr><td></td><td></td><td></td><td>
<table>
<tr><th>Number of Pregnancies</th><th>Ovary</th><th>Invasive ovary</th></tr>
<tr><td>0</td><td>1.67
(1.42, 1.96)</td><td>1.76
(1.45, 2.14)</td></tr>
<tr><td>1</td><td>1.32
(1.10, 1.57)</td><td>1.46
(1.18, 1.80)</td></tr>
<tr><td>2+</td><td>1.13
(0.93, 1.37)</td><td>1.10
(0.86, 1.40)</td></tr>
</table>

This is likely to be because some of the pregnancies included in this data may be of short duration.

This data has not been included as we feel the data included about live births is clearer and is the standardly reported outcome reported for ART.</td></tr>
<tr><td></td><td>16</td><td>I would have appreciated information to describe the magnitude of the effect of risk of ovarian cancer. This paper will raise anxiety amongst patients in the “at risk” categories, and amongst clinicians. We are only provided with statements of relative, not absolute risk. The discussion should try to contextualise the positive findings. The current version is similar to papers linking oral contraceptives to risk of venous thrombosis or the WHI study publications, which over empathised the risk of harm by presenting only relative risks.</td><td>Thank you. As mentioned in responses to point 5 above, more absolute risks have been included in the paper.</td></tr>
<tr><td></td><td>17</td><td>Why might a twin pregnancy be more protective against uterine cancer?</td><td>This result is difficult to interpret, as it is one of very few such estimates to have been made. The small number of such findings to date are inconsistent and ours might be a chance finding.</td></tr>
<tr><td></td><td>18</td><td>The authors should update ref 24 – viz ICMART report from 2016</td><td>Done- thank you.</td></tr>
</table>

Reviewer 3	19	This report is an authoritative analysis of the cancer risk in women following exposure to the medications and procedures undertaken using assisting reproduction. It is well written, clear in its aims and acknowledges its weaknesses. The linked access to mandatory notification of the procedures and the cancer registry has provided a unique opportunity for assessment of the potential risks. The authors consist of authorities in the relevant disciplines, notably epidemiology and the experience of employing similar research discipline of epidemiology for cancer screening and risk, now applied to fertility interventions and possible risk of cancer. The reassuring observations reported here will be welcomed by clinicians and patients alike.	Thank you
	20	It would be useful to know if the authors plan to analyse the risk of malignancy in women who act as oocyte? Egg "donors" who have been recruited due to their normal reproductive health as there is a increasing demand for oocytes from fertile young women.	Unfortunately, under current HFEA legislation access to data required for such analysis is not available.
Reviewer 4	21	My main concern about this study is the analytic approach used which, to my mind, is not fully optimal. The analytic approach used is a so-called 'SIR'-analysis, where the observed number of cancer cases in the exposed population is compared to the expected number of cancer cases in the general population. Such a comparison takes age and calendar time into account but does not has the ability to consider other important confounders (i.e. parity, causes of infertility, endometriosis, oral contraceptives, etc.)	As mentioned in reply to point 3 above, if we had access to an appropriate comparison group from the general population, with additional individual level data on potential founding factors, we agree that a Cox regression or survival analysis may be appropriate. However, a suitable external comparison group was not available. Therefore we used a standard method to compare observed rates with national population rates to compute standardised incidence ratios (SIRs). The size of this study has enabled us to instead stratify by

		Therefore, the calculated relative risks (SIRs) will most likely be confounded (most likely resulting in an overestimation of the risk estimates) and therefore difficult to interpret. On page 14, lines 38-45, the authors states that their study has '…sufficient size to stratify by potential confounding factors and thereby investigate characteristics of associations.'. This is correct, but it does not eliminate the problems with missing proper adjustment for confounding variables.	potential confounders. We argue that stratification, whilst not completely comprehensive in adjusting for potential confounders, does enable us to infer quite a lot about associations and even about potential causation. For example, it is reassuring that most cancer risks are not different to population based risks in the cohort of women who have ART due to male factor only infertility, suggesting ART itself is less likely to be mediating detected increased risks. Similarly, lack of dose response (except in in-situ breast cancer) and lack of increased cancer risk in women without existing known risk factors are reassuring. Given the strengths of this study and the caution used in interpreting results, we argue that it represents a valuable source of information to women who have and are considering undergoing ART, and their clinicians.
	22	Another problem with the analytic approach is that the risk estimates might be slightly underestimated, as the background comparison population also includes the exposed women.	We agree that background comparison population also includes exposed women. We made this point in the text (page 16) and point out that in our situation the bias would be minimal: "*Whilst comparator rates do include cohort participants, <5% of the population of reproductive-age women underwent assisted reproduction and our SIRs were generally <2.0; therefore resulting bias will have been minimal*[33]". This is supported by an appropriate reference that allows the size of bias to be calculated, and within this study the bias would be minimal.
	23	Lastly, the authors have made no attempt to censor for different kinds of gynecological surgery (in the exposed women) and have therefore overestimated the number of person-years at risk in the analysis.	Data for such censoring is not available, and population rates also do not correct for gynecological surgery. Whilst rates of gynecological surgery may differ between these cohorts, it is unlikely that the person-years would be altered enough to change observed results.

Reviewer 4 (continued)	24	I suggest that the number of (exposed) women could be mentioned instead of the number of PY, as I think the former is more informative.	We consider both to be equally informative and include the number of women clearly in our abstract. We are not inclined to change the title, but would consider doing so if editors felt this added clarity, for example '**Risks of ovarian, breast and corpus uteri cancer in 250,000 women treated with assisted reproductive technology; long-term follow-up in Great Britain'**
	26	Page 5, lines 16-32: I am not convinced that the references mentioned are sufficient. E.g., only one study is mentioned when referring to earlier studies on fertility drugs and ovarian cancer risk. Please make a more thorough description of the earlier publications in this research area.	Thank you. We have made appropriate changes to the introduction.
	27	Page 7, lines 19: Have you also excluded non-melanoma skin cancer before follow-up? Please specify. I do not think that it necessary to exclude women with this cancer type.	Women with non-melanoma skin cancer (ICD9 173/ ICD 10 C44) before treatment were excluded (n=222). We have amended the text to clarify this (OUTCOME DATA, methods section).
	28	The risk of breast cancer is split up in in-situ and invasive breast cancers and correspondingly, ovarian tumors are split up in invasive and borderline tumors. I suggest that the authors also look at the association with the different histological subtypes of respectively breast- and ovarian cancer as it is well-known that the aetiology of the different histotypes may differ, and it could therefore be interesting to see if that is also the case for the association with ART/infertility. Also, if the information is available, uterine cancers could also be split up in type 1 and type 2 endometrial cancers.	Unfortunately, the reference comparison rates for the general population required for these analyses are currently unavailable.
	29	IVF vs. ICSI: The exposure variable seems to be ART	This information could potentially be available (with further

		only. As ICSI is a more invasive method than IVF, and it can be hypothesised that it may affect cancer risk to a higher degree than IVF. If the authors have differentiated information on IVF and ICSI, I strongly suggest that these should be analysed as separate exposures. If this information is not available, it could be mentioned as a limitation	approvals and data extraction). However, the main difference between IVF and ICSI practically is the method of fertilization, which happens *in-vitro* in both procedures. Given that the two accepted potential mechanisms by which ART may increase cancer risk in women are 1) exposure to exogenous hormones and 2) exposure to multiple ovarian punctures, it is hard to hypothesize that IVF and ICSI would result in different cancer risks in women (this is obviously different when considering outcomes in children). In the UK the same stimulation protocols are used for both IVF and ICSI. Differences in success rates between the 2 are accounted for as we stratify by number of cycles. Differences in the background fertility of patients undergoing IVF and ICSI are accounted for as cause of fertility is stratified for. Therefore we suggest that further analysis of sub-groups without clear underlying hypotheses is not justified.