Manuscript ID BMJ.2016.037083 : "Antidepressants during pregnancy and offspring autism: population-based cohort study"

RESPONSE TO EDITORIAL COMMITTEE AND REVIEWERS

Dear Dr Weber,

We are very grateful for the time and effort that your reviewers and the editorial committee spent considering our paper. The comments were very useful and helped us improve the quality and readability of our paper. We were pleased that the reviewers and editorial committee felt the paper addressed an important issue and merits further consideration. We have tried to fully address all of the issues raised, and are very pleased to submit a revised manuscript. We hope this will enable you to reach a decision regarding publication. Please do not hesitate to contact me if you require any further information.

Yours sincerely,

Dheeraj Rai, corresponding author.

****Response to report from the BMJ's manuscript committee meeting****

BMJ Committee comment: *We wondered if the severity of depression varies between the groups and can be properly accounted for, despite propensity analysis (there is no adjustment for severity). Some covariates vary considerably (i.e. smoking) and may not reflect true rates of that risk factor.*

Authors: We agree, and continue to acknowledge that severity of depression may differ between women prescribed/taking antidepressants during pregnancy as opposed to those who do not. We also agree that despite being a potentially important confounder, accounting for it adequately is a challenge in register based studies such as this, and others published previously. To address this to the best of our ability, we carried out the propensity score matched analysis in addition to conventional regression analyses. To construct our propensity scores, we included several characteristics detailed in the methods supplement, which could be proxies of severity of depression: for example, these included the type of care (such as inpatient or outpatient) and number of contacts, number of diagnoses and polypharmacy all of which may be indicators of severity. We included a larger range of clinical and socioeconomic/demographic characteristics with the aim of achieving a balancing the groups based on their propensity scores. We stated this limitation originally in our submission (page 9): "The absence of detailed measures of depression severity during pregnancy was a limitation and so we used propensity scores to match and therefore balance exposure groups using a wide range of relevant characteristics". We have now described further the propensity score methods in the paper and in our response below, and also provided information on balance of covariates following propensity score matching.

BMJ Committee comment: *The focus on the sibling analysis drops off the headlines as the result is negative - it shouldn't disappear.*

Authors: We wrote in the abstract, and in the main paper that the direction of the results of the sibling analysis was consistent with the other analyses, although the confidence intervals were wide. To bring more prominence to these results- we have now amended the abstract slightly and also added an additional line in the discussion (Page 10) alongside the discussion of the possibility of genetic confounding: "Although the results of our sibling-control analyses were consistent with the other approaches we used, the numbers were low leading to an imprecise result. Larger sibling comparisons in the future may elucidate the role of genetic confounding in this relationship."

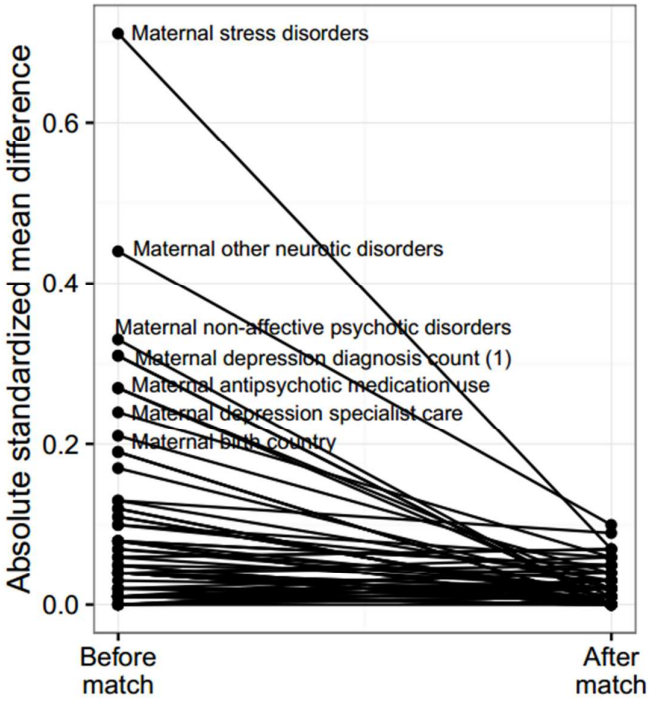
BMJ Committee comment: *Is the negative control really a negative control? We would have thought that using a group comprising mothers taking some other medication (e.g. hypnotics or treatments for irritable bowel disease) would be more appropriate.*

Authors: Using fathers is considered a strong negative control design in studies of in-utero exposures (for discussion see for example, Lewis S et al Approaches for strengthening causal inference regarding prenatal risk factors for childhood behavioural and psychiatric disorders. J Child Psychol and Psychiatry Volume 54, Issue 10, October 2013, 1095–1108). Parents often share a wide range of characteristics such as unmeasured family environment that could be potential confounding factors. If an association of a similar magnitude was also observed between fathers' antidepressant use during pregnancy and autism, it would indicate that the mechanism is unlikely to involve in-utero pathways and point towards unmeasured confounding. As the committee suggests, alternative medication exposures in mothers could also potentially act as negative controls but the results would be difficult to interpret if such medications were also implicated in autism aetiology. For example,

hypnotics are commonly prescribed in women with psychiatric conditions, and IBS and other GI problems have been linked with autism. We therefore think the paternal use of antidepressants makes for a stronger negative control.

BMJ Committee comment: *We can't see how the size of the cohort used after matching in the propensity score analysis, nor can we see how well the matching was; there is no table of matched characteristics with standardized differences.*

Authors: We had provided the n's in the methods supplement although we should also have provided matching diagnostics. The average standardized absolute mean difference (ASAM) of our PS matching was estimated at 0.032, suggesting a good match. We have re-estimated the boosted CART propensity score models as before in order to provide information on matched characteristics for each individual variable. Because boosting is a stochastic process, we needed to repeat the matching and n's of the matched samples and effect estimates varied slightly in our re-estimation. We have clearly marked the changes and updated the manuscript, although these changes do not affect the results in a major way, or our interpretation. We also took the opportunity to improve the covariate balance with additional refinement of the matching algorithm to include exact matching on anxiety disorders and obsessive-compulsive disorder (in addition to the previous match specifications). The ASAM calculated on re-estimation decreased to 0.022. In lieu of a balance table of the 66 covariates and covariate categories, we provide a graphical assessment of balance before and after matching below. These details have been included in the methods supplement.



BMJ Committee comment: *There was some missing data (e.g., BMI), how was this handled in any of the analyses, more importantly what else was missing and how much?*

Authors: We have added the following information to the Methods Supplement. In the study sample N = 254,610, there were missing data for the following characteristics: maternal age: N=5 (0.002%); paternal age: N=2284 (0.9%); family income quintile: N=117 (0.05%);

maternal education: N=1395 (0.55%). These data were not imputed due to the small number missing. In Analysis 1, Model 3, missingness on these characteristics resulted in a sample loss of 1.5%. Missingness on these characteristics did not influence Analysis 2 because the propensity score estimation technique was able to incorporate the missingness in the prediction. The sibling models in Analysis 3 did not adjust for these characteristics. In Analysis 4, missingness on these characteristics resulted in a sample loss of 1.0%. Therefore, missingness on these characteristics was not deemed significant.

Missingness was also observed for maternal smoking: N=34,249 (13.4%) and maternal BMI: N=52,893 (20.8%). These variables were not used in our main analyses because we have previously found no evidence of an association of maternal smoking (Lee BK et al. J Autism Dev Disord. 2012 Sep;42(9):2000-5) and BMI (Gardner RM et al. Int J Epidemiol. 2015 Jun;44(3):870-83) with autism. We did include these covariates in the propensity score model as they could potentially suggest important differences between exposure groups. The boosted CART model as a non-parametric technique is able to incorporate missing values in prediction and as a result no sample loss due to missingness on smoking/BMI occurred in the propensity score analyses. Furthermore, since previously identified predictors of smoking/BMI missingness (e.g., birth year) were also incorporated in the propensity score model; and because of the null associations of those characteristics with ASD, the missingness of these two characteristics is unlikely to meaningfully influence estimates.

BMJ Committee comment: *You might want to cite this recent systematic review: Reprod Toxicol. 2016 Dec;66:31-43. doi: 10.1016/j.reprotox.2016.09.013. Epub 2016 Sep 22. Prenatal selective serotonin reuptake inhibitor use and the risk of autism spectrum disorder in children: A systematic review and meta-analysis.*
https://www.ncbi.nlm.nih.gov/pubmed/27667009

Authors: We have now cited this systematic review

****Response to Reviewers****

Reviewer: 1

Reviewer comment: I would suggest a brief discussion about results emerging from this study and those reported in the review by Gentile, 2015 (Prenatal antidepressant exposure and the risk of Autism Spectrum Disorders in children. Are we looking at the falls of Gods?" Journal of Affective Disorders. 182:132-7

Authors: We have now cited this useful review

Reviewer: 2

Reviewer comment: *1. Intro - clarify what is meant by 'updated total population cohort'*

Authors: We have deleted the word ‘updated’ from the introduction. As also requested by reviewer 3, we make it clearer that the cohort has been updated since our previous study and that this manuscript contains a range of further analyses.

Reviewer comment: *a. Patient involvement - delete sentence starting with 'We will disseminate...'*

Authors: We were specifically asked by the editorial staff to include this paragraph so we have included this information (now further clarified after guidance from the editorial assistant).

Reviewer comment:*b. Medication use during pregnancy - last line on pg4 and first on pg5 is not clear.*

Authors: We agree the original sentence was difficult to follow. We have deleted the latter half and hope it reads OK now: "We considered exposure to antidepressants if there was a record of these in either the MBR or the PDR."

Reviewer comment: *c. Depression and other psychiatric disorders in parents - specify which 'other diagnoses of psychiatric disorders' were included, and during which time periods these diagnoses were considered, e.g., during pregnancy? anytime before delivery of study child? anytime including after delivery of study child? Only diagnoses made prior to the birth of the study child should be included.*

Authors: To clarify, we have now amended this sentence to: "We used ICD-9 and 10 diagnoses recorded in the National Patient Register, which covers inpatient (with complete national coverage since 1973), and outpatient specialist care (since 2001), to ascertain depression and other maternal and paternal diagnoses of other psychiatric indications for antidepressants (anxiety disorder, bipolar disorder, non-affective psychoses, obsessive-compulsive disorder, other stress-related and neurotic disorders) disorders any time before the birth of the child as described elsewhere."

Reviewer comment: *d. Analysis 1 - how was the comparison group defined? Both psychiatric disorder but no antidepressant use during pregnancy? or psychiatric disorder anytime but no antidepressant use during pregnancy? What proportion of mothers who used antidepressants had a psychiatric diagnosis?*

Authors: The comparison group was defined as having a maternal diagnosis of a major psychiatric indication for antidepressant use any time before the birth of the child, and was not exposed to antidepressants during pregnancy. Maternal indications for antidepressant use were: anxiety disorders (F40-41), bipolar disorder (F30-31), depression/mood disorder (F32-39), non-affective psychoses (F20-29), obsessive compulsive disorders (F42), stress related disorders (F43), and other neurotic disorders (F44-48). We have clarified this in the text and in Figure 1. In total, 1,866 of the 3,342 (55.8%) antidepressant-exposed children had mothers with any of these indications diagnosed prior to birth.

Reviewer comment: *e. Analysis 2 - it is not clear which children were included in this analysis. The numbers in the supplement are very low (204 ASD cases, 68 exposed) compared to the total shown at the bottom of Figure 1 (489 ASD cases, 136 exposed).*

Authors: The propensity score matching process matches similar exposed with unexposed persons and removes the dissimilar from analysis. We have now added the n's in Table 2 with further explanation in the methods supplement. Figure 1 shows the derivation of the analytic sample for the main analysis (analysis 1). We have now clarified this in the Figure 1 text.

Reviewer comment:*3. Results - last sentence in this section is incomplete*

Authors: Sorry this sentence was a remnant from an earlier draft where we had discussed the result of supplementary table S1 before deciding that it fitted better in the second paragraph of the results section. This incomplete sentence has now been deleted.

Reviewer comment: *4. Discussion a. Second sentence typo "particularly autism with intellectual disability". Should read "particularly autism without intellectual disability"*

Authors: Thanks, this has been corrected.

Reviewer comment: *b. pg 10 - lines 48-49 - this sentence doesn't follow the rest of the paragraph, which seems to be saying that confounding is possible but not that likely, at least not that likely to completely account for the results.*

Authors: We have moved this sentence further up in the paragraph and hope it now reads logically.

Reviewer comment:*c. very last sentence of manuscript is not clear. Important to investigate possible biologic mechanisms underlying the observed association with prenatal antidepressant exposure? alternative autism prevention strategies - alternative to what?*

Authors: We have deleted the word 'alternative' and think it conveys the meaning of this sentence more accurately.

Reviewer comment: TAble 1 - how are women categorized into 'Maternal depression dx before birth by dx type' categories? Are they mutually exclusive?

Authors: By including these variables, we were interested in capturing whether the women had any of the diagnoses for which antidepressants are prescribed. The categories were not mutually exclusive.

Reviewer comment: Table 3 - add N exposed cases

Authors: We have now added extra columns with the numbers.

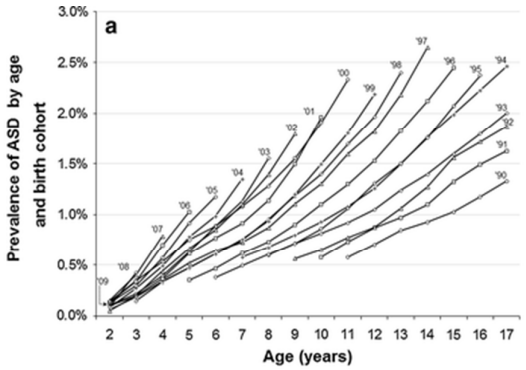
Reviewer: 3

Reviewer comment: *It would be helpful to make an explicit statement of study hypothesis at the end of the introduction, perhaps following p 3, line 55.*

Authors: We have now added the following sentence: "We hypothesised that if the association between antidepressant use during pregnancy and autism were likely to be causal, the results would be consistent across a range of analytic methods with different strengths and limitations and underlying assumptions."

Reviewer comment: In describing the methods, was there any trend in ASD prevalence over the study period? Exclusion of children under age 4 is reasonable, but by birth cohort did the prevalence change over time, and did the pattern of age at first diagnosis change over time (this reviewer would hypothesis that median age occurred earlier in 2011 than in 2001).

Authors: We have previously described the changing trends in ASD diagnoses in the population of Stockholm County making up our cohort (Idring S, et al. PLoS One 2012;7(7); Idring S, et al. J Autism Dev Disord 2015;45(6):1766-73). A figure from this previous work below highlights that at any particular age, newer birth cohorts have a higher prevalence of autism than older ones broadly consistent with the reviewer's hypothesis that the median age at diagnosis is decreasing in newer cohorts.



The question is whether these temporal changes are differential in relation to the exposure to antidepressant use (which have also increased over time in this cohort). If these changes represent a non-differential trend, our results would be biased towards the null. Below, we have investigated whether our results are influenced by mothers taking antidepressant in recent cohorts and these mothers have children diagnosed with ASD earlier. We ruled out this possibility by repeating the analysis 1 for autism without intellectual disability stratified by 3 year birth cohorts (shown below): we found similar estimates with overlapping confidence intervals across these time periods (the point estimates in the youngest cohort were lower than others with wide confidence intervals, probably indicating smaller numbers and a more limited follow up time for an autism diagnosis).

1996-1998: 2.14 (1.05,4.37)

1999-2001: 2.42 (1.46, 4.04)

2002-2004: 1.61 (1.02, 2.54)

2005-2007: 1.14 (0.68, 1.91)

Reviewer comment: *Unless required by the journal, the paragraph on patient involvement, p 4, could be removed, or summarized in a sentence attached to the end of the previous paragraph.*

Authors: The Journal editorial staff suggested to us that we add this paragraph to our manuscript when we initially submitted the paper.

Reviewer comment: *Ascertainment of ASD and ID are potential issues for this study. It appears (p 5, lines 18-36) that all data were obtained from ICD and DSM codes in the several registers used as input data sources. The text notes that validation studies were done (lines 32-36), but can the authors verify that a case diagnosed in 2004-6 has similar clinical features to those diagnosed in 2010-11?*

Authors: The diagnostic criteria for autism have not changed during the period of the study (DSM-V was introduced in 2013 after the end of follow up) so we have no reason to believe that the clinical features of those diagnoses in 2004-06 would be different from those diagnosed in 2010-11. However, we do not have symptom level data on all the individuals diagnosed in our cohort.

Reviewer comment: *In describing diagnosis of depression and other psychiatric disorders, these were apparently solely from ICD codes. Did the authors also supplement these with DSM diagnostic codes? Is it possible some cases not publicly funded were missed since they are not in the SAPCR?*

Authors: The two registers that we extracted the psychiatric diagnoses in parents (the National patient register and the Stockholm Adult Psychiatric Care Register) both use ICD-10 codes. We did use DSM-IV codes to supplement the autism diagnoses in children when using the Stockholm Child and Adolescent Mental Health Register where the DSM system is in use.

The vast majority of the population in Stockholm County use publicly financed care. Whilst it is possible that we missed some cases of psychiatric disorders diagnosed and treated privately, we do not think the non-ascertainment would substantially bias the results.

Reviewer comment: *In the description of the analysis, p 6, was age at diagnosis controlled in the analysis?*

Authors: We did not control for age at diagnosis. We do not think age at diagnosis of autism can be strictly considered a confounder since it occurs after the exposure rather than preceding it.

Reviewer comment: *The results from analysis 4 (p 7) are for only two years of birth cohort, and these children were only ages 4-5 at the conclusion of the study. Thus, these results should be interpreted with extreme care. The paper could be more straightforward if this analysis were removed.*

Authors: Our aim in this study was to use a variety of methods which can strengthen causal inference. We think that analysis 4, although based on a smaller sample, provides interesting additional information complementing our other analytic strategies. We agree these results should be interpreted with care but we hope that the reviewer will agree that we base no conclusions on this one analysis, but rather provide an overall picture of results triangulating across a range of different analyses.

Reviewer comment: *Based on results in first para of results on p 8, the prevalence of ASD was 2.1% among nonexposed pregnancies. This is considerably higher than what is reported in many other prevalence studies. It deserves some comment, as if the prevalence overall is higher the study findings may be less generalizable to other settings.*

Authors: The oldest 'children' with an ASD diagnosis in our cohort were 27 years of age and the above figure represents cumulative prevalence until this age. Whilst the prevalence appears higher than some (but not all – for example, see Kim et al Am J Psychiatry. 2011 Sep;168(9):904-12 which reported a prevalence of 2.6% in a South Korean total population

cohort) other prevalence studies, it should be noted those often restrict follow up time to a younger age (eg. The latest US CDC estimates of 1.46% are reported in children up to 8 years old: MMWR Surveill Summ. 2016 Apr 1;65(3):1-23). We think this difference in age when prevalence was estimated explains this difference.

Reviewer comment: *The last few sentences in the results need some work (p 9, lines 11-15). What is the specific point of focusing on differences in point estimates with overlapping CIs? Was a statistical test done to determine whether the estimates differ significantly? And the last sentence ends 'including . . .', seems something was left off.*

Authors: We wrote: "The point estimates for the risk of autism in users of low/moderate SERT affinity appeared to be greater than that of users of high SERT affinity antidepressants although the confidence intervals overlapped." Although the confidence intervals overlapped for these analyses (suggesting p values for any statistical test would be large), we thought it would be useful to present these analyses in order to provide a balanced view to our paper since these results, if replicated in larger cohorts, may actually suggest confounding by severity of depression since low SERT affinity antidepressants are more often prescribed for treatment resistant depression. We would prefer to keep this within our draft if the reviewer and editor do not have any objection. As for the last incomplete sentence in the results, as mentioned above, it crept in the submitted draft in error and we have now deleted it.

Reviewer comment: *The previous study used some of this study for a case-control design and was also published in BMJ (p 10, line 8). The discussion should more clearly outline how the present study provides stronger evidence, and whether its findings provide clarity to the prior results.*

Authors: We have amended the sentence to: "This study builds on our previous case-control study, now enhanced with a larger sample, a more stringent comparison group of mothers with psychiatric disorders, and a range of causal inference methods including propensity score matching, a sibling comparison, and a negative control design to strengthen confidence in the results."

Reviewer comment: *The second full para on p 10 begins by mentioning that this study has strengths and limitations, however, no limitations are explicitly addressed. There most definitely are limitations and these should be fully discussed here.*

Authors: We had discussed several limitations in the preceding two paragraphs above the one the reviewer has noted: including the possibility of exposure misclassification, inability to accurately measure severity of depression, small numbers, inability to carry out trimester or dose specific comparisons, imprecision in some estimates, inability to study other negative controls. In the paragraph we have discussed the possibility of genetic confounding and confounding by severity of depression.

Reviewer comment: *The discussion concludes by suggesting the attributable fraction for antidepressants is small, likely 2% at most. This might be interpreted in the context of other studies seeking to determine the relative contributions of identified risk factors, for example Schieve LA et al, Ann Epidemiol 2014;24:260-266.*

Authors: We have considered this suggestion but think weighing the relative importance of different risk factors that have been investigated for autism has the potential to be misleading because i) there are few such factors where evidence for causality is strong and ii) PAFS are sensitive to exposure and outcome prevalence and so may vary across studies carried out in different populations. We hope the reviewer will agree that making such comparisons may therefore be misinterpreted.

Reviewer comment: *p 3 line 32 - consider 'autism-like' rather than 'autism like'*

Authors: Done

Reviewer comment: p 5 line 47 - this might be an American v British English difference, but in the US we would use 'publicly' rather than 'publically'

Authors: Thanks- this was a typo and British spelling is also ‘publicly’ – corrected.

Reviewer comment: p 5 the paragraph headed 'Ascertainment of autism' should to titled 'Ascertainment of autism and intellectual disability

Authors: Done

Reviewer comment: p 8 line 29 consider space between % and CI

Authors: Done

Thank you again to the reviewers and editorial committee for all the important comments above which have been very helpful in improving our manuscript significantly.