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

TITLE (PROVISIONAL)	Neonatal and Childhood Neurodevelopmental, Health and Educational Outcomes of Children Exposed to Antidepressants and Maternal Depression During Pregnancy: Protocol for a Retrospective Population-Based Cohort Study Using Linked Administrative Data
AUTHORS	Singal, Deepa; Brownell, Marni; Chateau, Dan; Ruth, Chelsea; Katz, Laurence

VERSION 1 - REVIEW

REVIEWER	Salvatore Gentile, MD, Ph.D Psychiatrist, Neurologist ASL Salerno - Department of Mental Health - Piazza Galdi 84013 - Cava de' Tirreni (Salerno) - Italy - Adjunct Professor University of Naples - Italy - Medical School "Federico II" - Department of Neurosciences - Perinatal Psychiatry - Via S. Pansini, 5 80131 Naples
REVIEW RETURNED	13-Jul-2016

GENERAL COMMENTS	This is a very interesting study. However, the Authors will lump together data from the unexposed population suffering from different psychiatric disorders. Antenatal untreated anxiety disorders and depressive disorders have been associated with different, specific impacts on pregnancy and neonatal outcomes. For this reasons, data on women with anxiety disorders and depression should be analysed and presented disagreggated. I'm also unable to understand the difference between depressive disorders and mood disorders. Some references should be replaced with more recent articles.

REVIEWER	Bengt Källén Tornblad Institute, Lund University, Sweden
REVIEW RETURNED	15-Jul-2016

GENERAL COMMENTS	This manuscript describes plans for a study which will make use of a large set of different registers in Manitoba, Canada, 1996-2014. The purpose is to evaluate the effects of maternal use of antidepressants (notably SSRI and SNRI) during pregnancy and during specific parts of the pregnancy and the outcome in the offspring, from neonatal findings to long-term follow-up including childhood educational problems, neurodevelopmental disorders and mental disorders. A control group of women will be identified with a diagnosis of mood or

anxiety disorder during pregnancy but without use of antidepressants.

Major comments

1. The manuscript is unusual because it presents no findings, only plans for making a study. It is more like a grant application than a scientific article. It is of course an editorial decision if such a manuscript should be published or not.

2. The manuscript begins with an over-view of the large literature on antidepressant use during pregnancy. This is not a very useful summary and some critical aspects are not covered. To take an example: in a series of Danish studies which found an increased risk for congenital malformations after all use of antidepressants, data on malformations were obtained solely from discharge diagnoses. As infants born of mothers who have used such drugs are more often than other infants transferred to neonatal wards for other reasons than malformations, a biased reporting will be obtained. Page 9, line 48 and following would be a suitable place for this comment. On the other hand, in refs. 20 and 22, consideration was taken to pregnancy duration and maternal BMI so line 30 on p. 10 is not correct. Among the number of complicating factors which the authors mention (p. 8, line 32) no mention is made on the possible effects of the excess use of various other drug categories, not only drugs with psychiatric indications. On p. 8, line 13 the authors point out that no RCTs have been done – I think it should be clearly said that such studies are not ethical possible.

3. The manuscript presents the numerous registers one intends to use. There is, however, no information on the number of births which will be studied and the expected number of events. The latter figures would make it possible to evaluate the quality of the information. We know, for instance, relatively well the true rate of major congenital malformations, of diabetes type 1 and of epilepsy during pregnancy. I think such an initial tabulation should give valuable information on the quality and could also give a basis for power calculations.

4. Registers: One likely confounder is socio-economic level, notably for the long-.term follow-up studies. On p. 17 line 10 one talks about area level SES. This will probably be a crude estimate compared with individual data which may be available in the census information. On p. 23, line 41 it is indicated that individual data will be used. On this page the use of alcohol and smoking is also mentioned but it is not clear from where this information is obtained. An important confounder is maternal BMI but I cannot find any mentioning of it. - Another register is mentioned, Employment and Income assistance, but how this information is going to be used is not clear. A register of Education data seems to be used for evaluation of the children but not to give the education of the parents. Or? Medical diagnoses will be obtained from Hospital discharge abstracts – see comment 2 above – and from Medical/physician reimbursement claims. The latter source is restricted to one diagnosis on the level of three ICD digits which seems a rather crude and uncertain source of information. Registers 8-11 are described but no information is given on what information in them will be used or how it will be used.

5. A critical point in the study is the use, as a control to the antidepressant-exposed infants, of a group of women with a history

	of physician visits with a diagnosis for depressive disorder, affective psychosis or adjustment reaction (p. 20, line 6) or mood/anxiety disorder (p. 21, line 39). If we consider two important outcomes, ADHD and ASD, there is a rather strong heritability and such adult patients often get treatment with antidepressant. Maternal ADHD or ASD will be a strong confounder and I don't see how this will be treated. Paternal ADHD or ASD will be no confounder because it can hardly affect maternal antidepressant use (or can it?). If this cannot be disentangled, the study will anyway be inconclusive on these important points. 6. Timing of antidepressant use during pregnancy will be uncertain. One will demand two prescriptions during pregnancy to count the infant as exposed. For the issue of congenital malformations, one is quite enough (if she actually took the drug) – and this will mean that women who were judged not to have taken antidepressant may actually have done it. There are also other sources for obtaining drugs, via internet or relatives and friends, which also will make the definition of the unexposed group uncertain. It has been shown that prescription data may identify only about half of the actual exposures (based on interview) during the first trimester (Kallen et al., Eur J Clin Pharm 2011¸ 67: 839-845). To this can be added the uncertainty about when the woman (if at all) took the prescribed drug. It is a wide-spread knowledge that women should avoid drug use during the first trimester. This issue is not as easy as the authors apparently think and will undermine their belief that they are constructing a water-tight study. – A detail: in the list of SSRI, sertraline is not mentioned but I think it is used in Canada. 7. The authors will exclude women who used antipsychotic medication, benzodiazepines or opioids. In this way they will abstain from possibilities to study possible interactions between antidepressants and these drug groups. Notably the sedative/hypnotics group is relatively large. I think it would be better to keep them and make analyses on subgroups (if numbers will be enough). 8. Like in all studies using propensity scores, the effectiveness will depend on how adequate the measures of severity are. One component here is the number of visits to a physician or psychiatrist, I suppose with a diagnosis of depression/anxiety. Given that only one diagnosis is available at physician visits (and 3-figure ICD level) this variable seems very uncertain. If a woman has a pneumonia and depression it is likely that she will be registered as pneumonia! 9. The diagnosis of ADHD is a rather uncertain one. Have the authors considered the use of prescriptions for specific drugs like methylphenidate as a criterion instead? It would probably identify severe but relatively certain cases as these drugs have little use for other reasons. 10. I notice in Appendix 1 that VSD and ASD are apparently not included in cardiovascular defects in spite of the fact that the observations made on a possible increased risk of malformations after antidepressants to a large extent refer to those conditions. Why are they not included? Is it because the severity of the conditions cannot be evaluated from ICD codes? This is true also for hypospadias, for example, which is included. 11. The study plan indicates that a large number of comparisons will

	be, the exact number is not stated. Have you considered the multiple testing problem, and how do you intend to handle it?

REVIEWER	Simone Vigod Women's College Hospital, Toronto, Ontario
REVIEW RETURNED	30-Jul-2016

GENERAL COMMENTS	This is a population-based cohort study in Manitoba, Canada, studying maternal and child outcomes associated with antidepressant use in pregnancy among women with prenatally diagnosed mood and anxiety disorders. The strengths of the study lie in the comprehensive linked datasets with long-term child outcomes that have been rarely measured on a population-level in this field. The inverse probability of treatment-weights are a good approach to address confounding. However, one of the biggest problems in research in this field in general is whether the exposed and unexposed groups overlap enough on confounder distribution for these statistical techniques to balance them adequately. Hence, there is always potential for residual confounding. In the current protocol, it is discussed as a strength that the comparison group also has depression/anxiety, but I think that there are still limitations (unavoidable) to this approach. Most women will only take medication in pregnancy if they really feel they have no other choice. Almost by definition, they have a more severe form of illness, and statistical adjustment or propensity weighting or matching may not eliminate the confounding by indication situation. First, if exposed women have a more severe form of illness, they may also have a greater biological risk for mental disorders, and thus their children may have a greater biological risk for neurodevelopmental disorders and psychiatric disorders - so no amount of “adjustment” will be able to isolate the effect of antidepressants from the genetic risk - this is why some researchers have been adding discordant sibling-matched analyses to their protocols in this area, in efforts to better account for genetic and environmental risks. Second, those with more severe illness may also be at greater risk for postpartum mental health symptoms, with the negative implications that has for child development - and there does not appear to be a measure of postnatal maternal health longitudinally even though some of the outcomes are fairly long-term. Using health administrative data to identify women with mood and anxiety disorder diagnoses is unfortunately only a crude measure of illness. There has not been a validation of these codes, and it is hard to measure severity except with proxies like prior hospitalizations, or the use of antidepressant medication itself. The authors should make sure to acknowledge these issues both in the protocol and in the interpretation of any results. Particularly where the effect sizes are very small, we need to be very careful to acknowledge the possibility of a type I error, as women and providers need to understand these subtleties when making decisions about treatment. There are a few inconsistencies in the protocol, and also some minor points for consideration: 1. Antidepressant exposure: - Two antidepressant prescriptions during pregnancy makes sense to define exposure; but in the first section, it also says that it could

	be one prescription overlapping with gestation. This might be problematic given that women may have stopped taking their medications when pregnant even if they filled a prescription. I would be more apt to exclude such women from the analysis if that is their only exposure. Of note, when the exposed groups are later described, the exposure group definitions appear to have changed, so perhaps this is just a clarification point. - It is important not to over-play the accuracy of the exposure data as we really don't know how many women fill prescriptions and never take them. It is fairly standard in this work, not only to require 2 prescriptions during pregnancy, but to require them to be consecutive. 2. Exclusion criteria: These are reasonable but there are many exposures that could impact pregnancy outcomes, and I would rather see a measured approach to each outcome, with exclusions perhaps based on the logical mechanism underlying each of the outcomes. If you exclude all of these women at the outset, then you lose the opportunity to generalize the study to more complex women, and we need answers about them too. 3. The definition of the unexposed group needs some clarification. Some places it states that the look-back for mood and anxiety disorders is 3 months prior to pregnancy, other places 6 months, and in one place 12 months. In the section on "unexposed group", it starts to discuss women with new diagnoses in pregnancy. 4. As an extension of the point above about the severity of maternal mental illness, the look-back periods for indicators of severity also need to be defined. 5. Can the authors explain Gamma sensitivity analysis in the protocol?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1		
This is a very interesting study. However, the Authors will lump together data from the unexposed population suffering from different psychiatric disorders. Antenatal untreated anxiety disorders and depressive disorders have been associated with different, specific impacts on pregnancy and neonatal outcomes. For this reasons, data on women with anxiety disorders and depression should be analyzed and presented disaggregated. I'm also unable to understand the difference between depressive disorders and mood disorders.	Thank you for your comments Dr. Gentile, we appreciate your feedback. For the purposes of this research project we have decided to use the aggregated definition of mood and anxiety disorders. This is based on the limitations of the administrative data and the diagnostic capabilities of the providers diagnosing the majority of patients on antidepressants. There are high rates of comorbidities between mood and anxiety disorders and given that only one diagnosis is entered in the administrative database those identified as having one disorder or another would be frequently incorrect as patients often have both. Also, given that most patients are seen in primary care where the accuracy of diagnosis in sorting those out in the brief	

	time those providers have is questionable. Therefore, a panel of experts in Manitoba have decided that it is more accurate to use this aggregate definition in our work. A multitude of reports and papers have been published using this aggregate definition, including the following: • Kingston D, Heaman M, Brownell M, Ekuma O (2015) Predictors of Childhood Anxiety: A Population-Based Cohort Study. PLoS ONE 10(7): e0129339. doi:10.1371/journal.pone.0129339 • *Bolton JM, Walld R, Chateau D, Finlayson G, Sareen. J. Risk of suicide and suicide attempts associated with physical disorders: A population-based, balancing score-matched analysis. Psychological Medicine 2015;45:495-504.* • *Bolton JM, Au W, Chateau D, Walld R, Leslie WD, Enns J, Martens PJ, Katz LY, Logsetty S, Sareen J. Bereavement after sibling death: A population-based longitudinal case-control study. World Psychiatry 2016;15(1):59-66.* • *Enns J, Gawaziuk JP, Khan S, Chateau D, Bolton JM, Sareen J, Stone J, Doupe M, Logsetty S. Mental and physical health outcomes in parents of children with burn injuries as compared with matched controls. J Burn Care Res 2016;37(1):e18-26.* • *Leong C, Enns MW, Sareen J, Alessi-Severini S, Bolton J, Prior HJ, Chateau D. New antidepressant use in older adults: A Canadian population-based study (1997-2013). Aging & Mental Health 2016;Epub ahead of print.* • *Fransoo R, Martens P, The Need to Know Team, Prior H, Burchill C, Koseva I, Bailly A, Allegro E. The 2013 RHA Indicators Atlas. Winnipeg, MB: Manitoba Centre for Health Policy, 2013.* • *Brownell M, Chartier M, Santos R, Ekuma O, Au W, Sarkar J, MacWilliam L, Burland E, Koseva I, Guenette W. How are Manitoba's Children Doing? Winnipeg, MB: Manitoba Centre for Health Policy, 2012* • *Brownell M, Chartier M, Au W, MacWilliam L, Schultz J, Guenette*	

	W, Valdivia J. The Educational Outcomes of Children in Care in Manitoba. Winnipeg, MB: Manitoba • *Martens PJ, Fransoo R, McKeen N, The Need to Know Team, Burland E, Jebamani L, Burchill C, De Coster C, Ekuma O, Prior H, Chateau D, Robinson R, Metge C. Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study. Winnipeg, MB: Manitoba Centre for Health Policy, 2004.* We do however, appreciate this nuance and will conduct a sensitivity analysis using a disaggregated definition for key outcomes to see if there is a difference in outcomes between women with just one or the other diagnosis. We have also included this in our limitations section. Mood disorders and depressive disorders are the same thing – we have clarified this in the paper.	Page 23
Some references should be replaced with more recent articles.	In response to reviewer 2's concerns about the literature review we have edited it considerably to make it more concise, and in response to your comment about updating the references, the following references from 2016 and 2015 have been added. Our new background section now reflects the most current literature in this topic. 1. Vigod SN, Wilson CA, Howard LM. Depression in pregnancy. *BMJ (Clinical research ed.).* 2016;352:i1547. 2. Gentile S. Untreated depression during pregnancy: Short- and long-term effects in offspring. A	Page 6-10

systematic review. *Neuroscience.* 2015.

3. Gentile S, Fusco ML. Placental and fetal effects of antenatal exposure to antidepressants or untreated maternal depression. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet.* 2016:1-11.

4. Furu K, Kieler H, Haglund B, et al. Selective serotonin reuptake inhibitors and venlafaxine in early pregnancy and risk of birth defects: population based cohort study and sibling design. *BMJ (Clinical research ed.).* 2015;350:h1798.

5. Malm H, Brown AS, Gissler M, et al. Gestational Exposure to Selective Serotonin Reuptake Inhibitors and Offspring Psychiatric Disorders: A National Register-Based Study. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2016;55:359-66.

6. Brandlistuen RE, Ystrom E, Eberhard-Gran M, Nulman I, Koren G, Nordeng H. Behavioural effects of fetal antidepressant exposure in a Norwegian cohort of discordant siblings. *International journal of epidemiology.* 2015;44:1397-407.

7. Handal M, Skurtveit S, Furu K, et al. Motor development in children prenatally exposed to selective serotonin reuptake inhibitors: a large population-based pregnancy cohort study. *BJOG : an international journal of obstetrics and gynaecology.* 2015.

8. Harrington RA, Lee LC, Crum RM, Zimmerman AW, Hertz-Picciotto I. Prenatal SSRI use and offspring with autism spectrum disorder or developmental delay. *Pediatrics.*

2014;133:e1241-8.

9. Huybrechts KF, Bateman BT, Palmsten K, et al. Antidepressant use late in pregnancy and risk of persistent pulmonary hypertension of the newborn. *Jama.* 2015;313:2142-51.

10. Huybrechts KF, Hernandez-Diaz S, Avorn J. Antidepressant use in pregnancy and the risk of cardiac defects. *The New England journal of medicine.* 2014;371:1168-9.

11. Huybrechts KF, Sanghani RS, Avorn J, Urato AC. Preterm birth and antidepressant medication use during pregnancy: a systematic review and meta-analysis. *PloS one.* 2014;9:e92778.

12. Skurtveit S, Selmer R, Roth C, Hernandez-Diaz S, Handal M. Prenatal exposure to antidepressants and language competence at age three: results from a large population-based pregnancy cohort in Norway. *BJOG : an international journal of obstetrics and gynaecology.* 2014;121:1621-31.

13. Sorensen MJ, Gronborg TK, Christensen J, et al. Antidepressant exposure in pregnancy and risk of autism spectrum disorders. *Clinical epidemiology.* 2013;5:449-59.

14. Hanley GE, Mintzes B. Patterns of psychotropic medicine use in pregnancy in the United States from 2006 to 2011 among women with private insurance. *BMC pregnancy and childbirth.* 2014;14:242.

15. Nulman I, Koren G, Rovet J, Barrera M, Streiner DL, Feldman BM. Neurodevelopment of children prenatally exposed to selective reuptake inhibitor antidepressants: Toronto sibling study. *The Journal of clinical psychiatry.* 2015;76:e842-

	7. 16. Castro VM, Kong SW, Clements CC, et al. Absence of evidence for increase in risk for autism or attention-deficit hyperactivity disorder following antidepressant exposure during pregnancy: a replication study. *Translational psychiatry.* 2016;6:e708. 17. Hviid A, Melbye M, Pasternak B. Use of selective serotonin reuptake inhibitors during pregnancy and risk of autism. *The New England journal of medicine.* 2013;369:2406-15. 18. Kaplan YC, Keskin-Arslan E, Acar S, Sozmen K. Prenatal Selective Serotonin Reuptake Inhibitor Use and the Risk of Autism Spectrum Disorder in the Children: A Systematic Review and Meta-Analysis. *Reproductive toxicology (Elmsford, N.Y.).* 2016.	
Reviewer 2		
The manuscript is unusual because it presents no findings, only plans for making a study. It is more like a grant application than a scientific article. It is of course an editorial decision if such a manuscript should be published or not.	Thank you for your comments Dr. Kallen, we appreciate your feedback. BMJ Open published study protocols which are manuscripts that outline the methodology of proposed studies in order to ensure that work is not duplicated and that methodology is disseminated to the larger research community. To make our purpose clearer, we have also stated the purpose of this paper on page 15: **"The purpose of this protocol paper is to: (1) provide dissemination of our research activity to prevent duplication of work and encourage collaboration; (2) provide a citation for a detailed study methodology that can be referenced for future papers produced from this**	N/A Page 10

	research to enhance transparency."	
The manuscript begins with an over-view of the large literature on antidepressant use during pregnancy. This is not a very useful summary and some critical aspects are not covered. To take an example: in a series of Danish studies which found an increased risk for congenital malformations after all use of antidepressants, data on malformations were obtained solely from discharge diagnoses. As infants born of mothers who have used such drugs are more often than other infants transferred to neonatal wards for other reasons than malformations, a biased reporting will be obtained. Page 9, line 48 and following would be a suitable place for this comment. On the other hand, in refs. 20 and 22, consideration was taken to pregnancy duration and maternal BMI so line 30 on p. 10 is not correct.	Thank you for pointing out the weakness in this literature review. We have re-structured it, made it more concise, shortened it and included more recent studies, listed below. These references reflect the most current literature on this topic. 1. Vigod SN, Wilson CA, Howard LM. Depression in pregnancy. *BMJ (Clinical research ed.).* 2016;352:i1547. 2. Gentile S. Untreated depression during pregnancy: Short- and long-term effects in offspring. A systematic review. *Neuroscience.* 2015. 3. Gentile S, Fusco ML. Placental and fetal effects of antenatal exposure to antidepressants or untreated maternal depression. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet.* 2016:1-11. 4. Furu K, Kieler H, Haglund B, et al. Selective serotonin reuptake inhibitors and venlafaxine in early pregnancy and risk of birth defects: population based cohort study and sibling design. *BMJ (Clinical research ed.).* 2015;350:h1798. 5. Malm H, Brown AS, Gissler M, et al. Gestational Exposure to Selective Serotonin Reuptake Inhibitors and Offspring Psychiatric Disorders: A National Register-Based Study. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2016;55:359-66.	Page 6-10

6. Brandlistuen RE, Ystrom E, Eberhard-Gran M, Nulman I, Koren G, Nordeng H. Behavioural effects of fetal antidepressant exposure in a Norwegian cohort of discordant siblings. *International journal of epidemiology.* 2015;44:1397-407.

7. Handal M, Skurtveit S, Furu K, et al. Motor development in children prenatally exposed to selective serotonin reuptake inhibitors: a large population-based pregnancy cohort study. *BJOG : an international journal of obstetrics and gynaecology.* 2015.

8. Harrington RA, Lee LC, Crum RM, Zimmerman AW, Hertz-Picciotto I. Prenatal SSRI use and offspring with autism spectrum disorder or developmental delay. *Pediatrics.* 2014;133:e1241-8.

9. Huybrechts KF, Bateman BT, Palmsten K, et al. Antidepressant use late in pregnancy and risk of persistent pulmonary hypertension of the newborn. *Jama.* 2015;313:2142-51.

10. Huybrechts KF, Hernandez-Diaz S, Avorn J. Antidepressant use in pregnancy and the risk of cardiac defects. *The New England journal of medicine.* 2014;371:1168-9.

11. Huybrechts KF, Sanghani RS, Avorn J, Urato AC. Preterm birth and antidepressant medication use during pregnancy: a systematic review and meta-analysis. *PloS one.* 2014;9:e92778.

12. Skurtveit S, Selmer R, Roth C, Hernandez-Diaz S, Handal M. Prenatal exposure to antidepressants and language competence at age three: results from a large population-based pregnancy cohort in Norway. *BJOG : an international journal of obstetrics and gynaecology.*

2014;121:1621-31.

13. Sorensen MJ, Gronborg TK, Christensen J, et al. Antidepressant exposure in pregnancy and risk of autism spectrum disorders. *Clinical epidemiology.* 2013;5:449-59.

14. Hanley GE, Mintzes B. Patterns of psychotropic medicine use in pregnancy in the United States from 2006 to 2011 among women with private insurance. *BMC pregnancy and childbirth.* 2014;14:242.

15. Nulman I, Koren G, Rovet J, Barrera M, Streiner DL, Feldman BM. Neurodevelopment of children prenatally exposed to selective reuptake inhibitor antidepressants: Toronto sibling study. *The Journal of clinical psychiatry.* 2015;76:e842-7.

16. Castro VM, Kong SW, Clements CC, et al. Absence of evidence for increase in risk for autism or attention-deficit hyperactivity disorder following antidepressant exposure during pregnancy: a replication study. *Translational psychiatry.* 2016;6:e708.

17. Hviid A, Melbye M, Pasternak B. Use of selective serotonin reuptake inhibitors during pregnancy and risk of autism. *The New England journal of medicine.* 2013;369:2406-15.

18. Kaplan YC, Keskin-Arslan E, Acar S, Sozmen K. Prenatal Selective Serotonin Reuptake Inhibitor Use and the Risk of Autism Spectrum Disorder in the Children: A Systematic Review and Meta-Analysis. *Reproductive toxicology (Elmsford, N.Y.).* 2016.

Among the number of complicating factors which the authors mention (p. 8, line 32) no mention is made on the possible effects of the excess use of various other drug categories, not only drugs with psychiatric indications.	Thank you for pointing out this oversight, we concur that this is a very important confounding factor and should be mentioned, therefore, we have added the following sentence: "In observational studies, it is also often difficult to control for confounding factors that may alter the relationship between the development of adverse neonatal and childhood outcomes, such as socioeconomic status (SES), maternal nutrition, co-morbid mental and physical illness, prenatal substance use, **and the use of other medications including those with and without psychiatric indications**".	Page 7
On p. 8, line 13 the authors point out that no RCTs have been done – I think it should be clearly said that such studies are not ethical possible.	We have included the statement "**Due to ethical considerations, there are no randomized controlled trials that…."**	Page 6
This manuscript presents the numbers of registers one intends to use. There is, however, no information on the number of births which will be studied and the expected number of events. The latter figures would make it possible to evaluate the quality of the information. We know, for instance, relatively well the true rate of major congenital malformations, of diabetes type 1 and of epilepsy during pregnancy. I think such an initial tabulation should give valuable information on the quality and could also give a basis for power calculations.	Thank you for bringing to our attention that the statement we have included on page 25 that describes our exposed number of births that will be studied was not effectively placed. To highlight that the number of births that will be studied, we have moved this statement to page 12 under "Study Design and Population" where it will be clearer to the reader and a more appropriate fit for this information. **"Based on a study on the perinatal health of women in Manitoba by MCHP, 7.5% of women were diagnosed with prenatal psychological distress (including depression) out of 15,000 births in 2008/2009[91]. This will give us approximately 21,375 women diagnosed with pre-natal psychological distress during our study period. Based on**	Page 12

	previous studies, the majority of these mothers (>90%) will be linkable to their children, obtaining the largest Canadian sample to date of children exposed to prenatal antidepressants and/or prenatal mood/anxiety disorders."	
Registers: One likely confounder is socio-economic level, notably for the long-.term follow-up studies. On p. 17 line 10 one talks about area level SES. This will probably be a crude estimate compared with individual data which may be available in the census information. On p. 23, line 41 it is indicated that individual data will be used. On this page the use of alcohol and smoking is also mentioned but it is not clear from where this information is obtained. Another register is mentioned, Employment and Income assistance, but how this information is going to be used is not clear. A register of Education data seems to be used for evaluation of the children but not to give the education of the parents. Or? Registers 8-11 are described but no information is given on what information in them will be used or	Thank-you for bringing up this area of clarification. Area level data will come from the census information, as stated in our definition on page 14. Individual data will come from income assistance data, which is a good indicator for poverty. A report by MCHP reported that approximately 15% of pregnant women were on income assistance, making this indicator a quality indicator to reflect the social complexities of our study cohort. This has been clarified in our manuscript. Thank you for bringing to our attention that we have not clearly linked the explanation of our registers to the information that we will be obtaining from them. To provide a greater understanding for our data to our readers and greater transparency in our methods we have included a new table (TABLE 1) entitled: "**Description of data sets used for analysis and types of information retrieved.**" This table includes the name of the dataset, the description of the dataset, years of data used, and information retrieved. See our new Table 1. Alcohol and smoking information will be available through the Families First/Babies First Data We have access to education data for both	Page 33 Page 33 Page 33

how it will be used.	the mothers and the children. Income assistance will provide individual level SES data. See our new Table 1. See our new Table 1.	Page 33
Medical diagnoses will be obtained from Hospital discharge abstracts – see comment 2 above – and from Medical/physician reimbursement claims. The latter source is restricted to one diagnosis on the level of three ICD digits which seems a rather crude and uncertain source of information.	The majority of the definitions that we use to obtain medical diagnoses include not only information from hospital discharge abstracts and from physician reimbursement claims, but also the use of prescription drug data (please see Appendix 1). This strengthens the accuracy of the diagnosis. We do agree that for certain diagnoses, coding to the 3rd digit ICD code is not specific enough, however, for the purposes of mood and anxiety disorders, all diagnosis that would be coded using the fourth or fifth digits are included, thus all sub-classifications are included in three digit coding. These definitions have been utilized in various studies and government reports, and moreover, have been validated to be used in health services research (see references cited below): 1. Coleman N, Halas G, Peeler W, Casaclang N, Williamson T, Katz A. From patient care to research: a validation study examining the factors contributing to data quality in a primary care electronic medical record database. BMC family practice 2015;**16**:11 doi: 10.1186/s12875-015-0223-z[published Online First: Epub Date]|. 2. Katz A, Soodeen RA, Bogdanovic B, De Coster C, Chateau D. Can the quality of care in family practice be measured using administrative data? Health services research 2006;**41**(6):2238-54 doi: 10.1111/j.1475-6773.2006.00589.x[published	

Online First: Epub Date]|.

3. Marrie RA, Fisk JD, Stadnyk KJ, et al. Performance of administrative case definitions for comorbidity in multiple sclerosis in Manitoba and Nova Scotia. Chronic diseases and injuries in Canada 2014;**34**(2-3):145-53

4. Marrie RA, Yu BN, Leung S, et al. Prevalence and incidence of ischemic heart disease in multiple sclerosis: A population-based validation study. Multiple sclerosis and related disorders 2013;**2**(4):355-61 doi: 10.1016/j.msard.2013.03.001[published Online First: Epub Date]|.

5. Marrie RA, Yu BN, Leung S, et al. The incidence and prevalence of fibromyalgia are higher in multiple sclerosis than the general population: A population-based study. Multiple sclerosis and related disorders 2012;**1**(4):162-7 doi: 10.1016/j.msard.2012.06.001[published Online First: Epub Date]|.

6. Marrie RA, Yu BN, Leung S, et al. The incidence and prevalence of thyroid disease do not differ in the multiple sclerosis and general populations: a validation study using administrative data. Neuroepidemiology 2012;**39**(2):135-42 doi: 10.1159/000339757[published Online First: Epub Date]|.

7. Metcalfe A, Lix LM, Johnson JA, et al. Validation of an obstetric comorbidity index in an external population. BJOG : an international journal of obstetrics and gynaecology 2015;**122**(13):1748-55 doi: 10.1111/1471-0528.13254[published Online First: Epub Date]|.

8. Nuttall M, van der Meulen J, Emberton M. Charlson scores based on ICD-10 administrative data were valid in assessing comorbidity in patients

	undergoing urological cancer surgery. Journal of clinical epidemiology 2006;**59**(3):265-73 doi: 10.1016/j.jclinepi.2005.07.015[published Online First: Epub Date]|. 9. Wittie M, Nelson LM, Usher S, Ward K, Benatar M. Utility of capture-recapture methodology to assess completeness of amyotrophic lateral sclerosis case ascertainment. Neuroepidemiology 2013;**40**(2):133-41 doi: 10.1159/000342156[published Online First: Epub Date]|. 10. Lix L, Yogendran M, Burchill C, Metge C, McKeen N, Moore D, Bond R. Defining and Validating Chronic Diseases: An Administrative Data Approach. Winnipeg, Manitoba Centre for Health Policy, July 2006.	
An important confounder is maternal BMI but I cannot find any mentioning of it. –	While maternal BMI is an important confounder, we cannot obtain it from our data, therefore, we have included this as a limitation in our limitation section.	Page 22
A critical point in the study is the use, as a control to the antidepressant-exposed infants, of a group of women with a history of physician visits with a diagnosis for depressive disorder, affective psychosis or adjustment reaction (p. 20, line 6) or mood/anxiety disorder (p. 21, line 39). If we consider two important outcomes, ADHD and ASD, there is a rather strong heritability and such adult patients often get treatment with antidepressant. Maternal ADHD or ASD will be a strong confounder and I don't see how this will be treated. Paternal ADHD or ASD will be no confounder because it can hardly affect maternal	Thank you for highlighting this important oversight on our part. We are indeed going to control for maternal ADHD as well as maternal ASD - we had incorrectly assumed that we had explained on page 24 that "**additional covariates would be taken into consideration for known or suspected risk factors for certain outcomes….."** Because of the importance of potential for confounding by maternal ADHD or ASD we have now included this in this statement on page 24: **"*Additional covariates that will be taken into consideration include known or suspected risk factors for certain***	Page 17

antidepressant use (or can it?). If this cannot be disentangled, the study will anyway be inconclusive on these important points.	***outcomes; for example, known risk factors for congenital cardiac malformation include: multiple gestation, maternal hypertension, diabetes, renal disease, use of other psychotropic medications. Moreover, specifically in the investigation of ADHD or ASD in children, we will control for maternal ADHD or ASD, as there is a possibility of genetic heritability for these disorders. A unique model will be built for each outcome that takes into account important risk factors and covariates*.”** While paternal ADHD and ASD may effect maternal antidepressant use, we feel that this is outside the scope of this research. Future research in this area should be done as paternal influences are also important in fetal development.	
Timing of antidepressant use during pregnancy will be uncertain. One will demand two prescriptions during pregnancy to count the infant as exposed. For the issue of congenital malformations, one is quite enough (if she actually took the drug) – and this will mean that women who were judged not to have taken antidepressant may actually have done it. There are also other sources for obtaining drugs, via internet or relatives and friends, which also will make the definition of the unexposed group uncertain. It has been shown that prescription data may identify only about half of the actual exposures (based on interview) during the first trimester (Kallen et al., Eur J Clin Pharm 2011¸ 67: 839-845). To this can be added the uncertainty about when the woman (if at all) took the prescribed drug. It is a wide-spread knowledge that women should avoid drug use during the first trimester. This issue is not as easy as the authors	We respectfully acknowledge that studying this important research quandary is a very challenging undertaking and there will be limitations in our work that are unavoidable given the nature of administrative data claims. We are working within the constraints of this data and are attempting to develop methods that will give us the most rigorous study methodology in our capacity, but acknowledge there are important limitations beyond our control. We have therefore, edited our limitations section to include several of the limitations raised by the reviewers. Please see our new expanded limitations sections: **“Limitations: We acknowledge that conducting research in this area is extremely complex, and there are potential limitations that will warrant our careful consideration, including: lack of data on non-pharmaceutical treatment of mood/anxiety disorders (e.g., cognitive-behaviour therapy given by psychologists); limited information on disease severity; lack of information on maternal BMI; and potential for**	Page 22

apparently think and will undermine their belief that they are constructing a water-tight study. – A detail: in the list of SSRI, sertraline is not mentioned but I think it is used in Canada.	**unmeasured confounding. Despite our efforts in controlling for confounding by indication, women who take medications during pregnancy may have a more severe form of illness, and no amount of statistical adjustment using information from administrative data claims can eliminate this type of confounding. Furthermore, some women may also have a greater biological risk for mental disorders, and thus their children may have an increased biological risk for adverse childhood outcomes, such as neurodevelopmental disorders and psychiatric disorders. Our study may not fully be able to isolate the effect of antidepressants from genetic risk, despite controlling for family history. Moreover, as with all studies utilizing administrative data claims, the diagnoses are reliant upon the accuracy of physician data. As well, we are utilizing an aggregate definition of mood and anxiety disorders due to limitations of the data and the diagnostic capabilities of the providers diagnosing the majority of patients taking antidepressants. There are high rates of comorbidities between mood and anxiety disorders and given that only one diagnosis is entered in the administrative database, women having one disorder or another would be frequently incorrect as patients often have both. Also, given that most patients are seen in primary care settings where the accuracy of diagnosis in sorting between these two disorders is questionable, a panel of experts in Manitoba has decided that it is more accurate to use this aggregate definition in our work. However, we acknowledge that antenatal untreated mood and anxiety disorders may be associated with different impacts on neonatal impacts and will conduct sensitivity analysis using a disaggregated definition for key outcomes to assess if there is a difference between women with one or the other diagnosis. Furthermore, as with all studies utilizing drug data, we do not know if women actually took the**	

	medications they were prescribed, or if they stopped taking their medications after they filled their prescription. To help account for this possibility we have only included women who have more than one prescription throughout their pregnancy.. We also cannot account for the confounding effect of illicit drugs or illegally obtained prescription drugs. Finally, due to the exploratory nature of this study, there are multiple comparisons being performed and we acknowledge the potential for an inflated type 1 error, which is a limitation of this work. Future research should be conducted that corroborates the results from this proposed study." We do agree that one prescription may be enough to cause congenital malformations and by stating women need two prescriptions to be counted as an exposed infant, we are being extra precautious and conservative. If women who did have one antidepressant prescription during pregnancy are counted incorrectly in our unexposed group, this will actually serve to weaken rather than strengthen any findings. This consideration also applies to women who obtained drugs via others sources such as the internet or relatives and friends. Moreover, the uncertainty about when and if the woman at all took the prescribed drug during pregnancy has been included as an important limitation of this work, and is a limitation in any study ever conducted in this area. We have added sertraline to the list of SSRIs on page 13.	Page 13
The authors will exclude women who used antipsychotic medication, benzodiazepines or opioids. In this way they will abstain from possibilities to study possible interactions between antidepressants and these drug groups. Notably the	We absolutely agree that the investigation of the use of antipsychotic medications, benzodiazepines and opioids is essential, as well as the examination of the possible interactions between antidepressants and these drug groups. We actually have funding and approval to conduct a future study that is investigating these exact	n/a

sedative/hypnotics group is relatively large. I think it would be better to keep them and make analyses on subgroups (if numbers will be enough).	questions. Because this investigation requires great detail, methodology, as well as additional funds, we felt it was beyond the scope of this current study and did not include it in this current study protocol. We will be conducting another series of studies that look at these very objectives and hope you will be able to review those as well.	
Like in all studies using propensity scores, the effectiveness will depend on how adequate the measures of severity are. One component here is the number of visits to a physician or psychiatrist, I suppose with a diagnosis of depression/anxiety. Given that only one diagnosis is available at physician visits (and 3-figure ICD level) this variable seems very uncertain. If a woman has a pneumonia and depression it is likely that she will be registered as pneumonia!	While we agree, if a woman presents with both pneumonia and depression she will likely be coded as pneumonia, we are making the assumption that this woman will likely also present for depression at one point and that she will be included in our work. This is an inherent limitation that is present with all work utilizing administrative data claims.	n/a
The diagnosis of ADHD is a rather uncertain one. Have the authors considered the use of prescriptions for specific drugs like methylphenidate as a criterion instead? It would probably identify severe but relatively certain cases as these drugs have little use for other reasons.	In appendix one we do state that our definition of ADHD includes prescription drugs. Methylphenidate is included in this definition of ADHD which has been used for multiple research studies by multiple health care professionals and psychiatrists in Manitoba.	n/a
I notice in Appendix 1 that VSD and ASD are apparently not included in cardiovascular defects in spite of the fact that the observations made on a possible increased risk of malformations after antidepressants to a large extent refer to those conditions. Why are they not included? Is it because the severity of the conditions cannot be evaluated from ICD codes? This is true also for hypospadias, for example, which is included.	Thank you for pointing this out. We are planning to look at both Ventricular Septal Defects (VSD) and Atrial Septal Defects, which are both acquired and congenital heart defects. We had made the assumption that including “congenital hear defects” these two disorders would be implied, but have no added them to table 1 for clarity.	Page 34

The study plan indicates that a large number of comparisons will be, the exact number is not stated. Have you considered the multiple testing problem, and how do you intend to handle it?	Thank you for bringing this to our attention. We have acknowledged that the multiple testing problem will be a limitation of our work and have included it in our limitation section: **"Finally, due to the exploratory nature of this study, there are multiple comparisons being performed and we acknowledge the potential for an inflated type 1 error, which is a limitation of this work. Future research should be conducted that corroborates the results from this proposed study."**	Pg 22
Reviewer 3: Dr. Simone Vigod		
The inverse probability of treatment-weights are a good approach to address confounding. However, one of the biggest problems in research in this field in general is whether the exposed and unexposed groups overlap enough on confounder distribution for these statistical techniques to balance them adequately. Hence, there is always potential for residual confounding.	Thank you for your comments Dr. Vigod, we appreciate your feedback. We strongly concur that there is the potential for residual, or unmeasured confounding when applying treatment weights. Although we cannot possibly control for all potential unmeasured confounding, we can conduct gamma sensitivity analysis that allows us to examine how strong an unmeasured confounder would have to be to nullify any statistically significant findings. Please see the addition to our sensitivity analysis section on page 18 for further detail. We have also included this in our limitations section.	Page 18
In the current protocol, it is discussed as a strength that the comparison group also has depression/anxiety, but I think that there are still limitations (unavoidable) to this approach. Most women will only take medication in pregnancy if they really feel they have no other choice. Almost by definition, they have a more severe form of illness, and statistical adjustment or propensity weighting or matching may not eliminate the confounding by indication situation. First, if exposed women have a more	Thank you for these important comments. Unfortunately in any study in this field there will be unavoidable limitations, as you have pointed out. We have expanded our limitations section to include some of these limitations that you have pointed out.	Page 22

severe form of illness, they may also have a greater biological risk for mental disorders, and thus their children may have a greater biological risk for neurodevelopmental disorders and psychiatric disorders - so no amount of "adjustment" will be able to isolate the effect of antidepressants from the genetic risk - this is why some researchers have been adding discordant sibling-matched analyses to their protocols in this area, in efforts to better account for genetic and environmental risks. Second, those with more severe illness may also be at greater risk for postpartum mental health symptoms, with the negative implications that has for child development - and there does not appear to be a measure of postnatal maternal health longitudinally even though some of the outcomes are fairly long-term. Using health administrative data to identify women with mood and anxiety disorder diagnoses is unfortunately only a crude measure of illness. There has not been a validation of these codes, and it is hard to measure severity except with proxies like prior hospitalizations, or the use of antidepressant medication itself. The authors should make sure to acknowledge these issues both in the protocol and in the interpretation of any results. Particularly where the effect sizes are very small, we need to be very careful to acknowledge the possibility of a type I error, as women and providers need to understand these subtleties when making decisions about treatment.	See our expanded limitation section below: **"Limitations: We acknowledge that conducting research in this area is extremely complex, and there are potential limitations that will warrant our careful consideration, including: lack of data on non-pharmaceutical treatment of mood/anxiety disorders (e.g., cognitive-behaviour therapy given by psychologists); limited information on disease severity; lack of information on maternal BMI; and potential for unmeasured confounding. Despite our efforts in controlling for confounding by indication, women who take medications during pregnancy may have a more severe form of illness, and no amount of statistical adjustment using information from administrative data claims can eliminate this type of confounding. Furthermore, some women may also have a greater biological risk for mental disorders, and thus their children may have an increased biological risk for adverse childhood outcomes, such as neurodevelopmental disorders and psychiatric disorders. Our study may not fully be able to isolate the effect of antidepressants from genetic risk, despite controlling for family history. Moreover, as with all studies utilizing administrative data claims, the diagnoses are reliant upon the accuracy of physician data. As well, we are utilizing an aggregate definition of mood and anxiety disorders due to limitations of the data and the diagnostic capabilities of the providers diagnosing the majority of patients taking antidepressants. There are high rates of comorbidities between mood and anxiety disorders and given that only one diagnosis is entered in the administrative database, women having one disorder or another would be frequently incorrect as patients often have both. Also, given that most patients are seen in primary care settings where**	

	the accuracy of diagnosis in sorting between these two disorders is questionable, a panel of experts in Manitoba has decided that it is more accurate to use this aggregate definition in our work. However, we acknowledge that antenatal untreated mood and anxiety disorders may be associated with different impacts on neonatal impacts and will conduct sensitivity analysis using a disaggregated definition for key outcomes to assess if there is a difference between women with one or the other diagnosis. Furthermore, as with all studies utilizing drug data, we do not know if women actually took the medications they were prescribed, or if they stopped taking their medications after they filled their prescription. To help account for this possibility we have only included women who have more than one prescription throughout their pregnancy.. We also cannot account for the confounding effect of illicit drugs or illegally obtained prescription drugs. Finally, due to the exploratory nature of this study, there are multiple comparisons being performed and we acknowledge the potential for an inflated type 1 error, which is a limitation of this work. Future research should be conducted that corroborates the results from this proposed study." To address your very good suggestion of adding discordant sibling-matched analyses – we have given a lot of thought to this and feel that this type of analysis presents its own challenges. First, obtaining a large sample of women who have had multiple pregnancies and have taken medications during one and not the other will be extremely difficult. Furthermore, we may not be correct in our assumption that the women did indeed take medications throughout one pregnancy and not the other since we cannot be sure that women take their medications during pregnancy at all. Furthermore, women may not have had diagnosed depression during one pregnancy and were diagnosed with	

	depression in a subsequent pregnancy (divorce, stressful life event etc.) and therefore, present difficulties accounting for untreated depression, which we feel is a priority. Once this initial analysis is completed, we may attempt to conduct future work that uses this type of analysis. Thank you for this suggestion.	
Antidepressant exposure: - Two antidepressant prescriptions during pregnancy makes sense to define exposure; but in the first section, it also says that it could be one prescription overlapping with gestation. This might be problematic given that women may have stopped taking their medications when pregnant even if they filled a prescription. I would be more apt to exclude such women from the analysis if that is their only exposure. Of note, when the exposed groups are later described, the exposure group definitions appear to have changed, so perhaps this is just a clarification point. - It is important not to over-play the accuracy of the exposure data as we really don't know how many women fill prescriptions and never take them. It is fairly standard in this work, not only to require 2 prescriptions during pregnancy, but to require them to be consecutive.	Thank-you for highlighting the need to clarify this important point. Two antidepressant prescriptions during pregnancy will be needed to define antidepressant exposure. These two prescriptions can include one prescription that overlaps with gestation. – we have edited the sentence to explicitly state this. We strongly agree with your point that women may have stopped taking their medications when pregnant even if they filled a prescription, therefore, we are only counting women as exposed if they did indeed fill two prescriptions. We are also reiterating the limitations of the exposure data in our newly expanded limitations section (see above response).	Page 22
Exclusion criteria: If you exclude all of these women at the outset, then you lose the opportunity to generalize the study to more complex women, and we need answers about them too.	We appreciate that women with complex mental health needs also need answers about the effects that medications may have on their children and agree that this is a very important research area. For the purposes of this study we have excluded women exposed to antipsychotic medications, and/or benzodiazepines, and/or opioids as these medications may	N/A

	affect child outcomes over antidepressants and are potential confounding variables, as are the disorders themselves. We feel that including these women as a subset analysis is outside the scope of this study and involves extra resources and data. Therefore, we are applying for additional funds to start a future study utilizing these very groups of women mentioned above. We hope you will have the opportunity to review or provide expertise on our upcoming studies as well.	
The definition of the unexposed group needs some clarification. Some places it states that the look-back for mood and anxiety disorders is 3 months prior to pregnancy, other places 6 months, and in one place 12 months. In the section on "unexposed group", it starts to discuss women with new diagnoses in pregnancy.	We apologize for our inconsistency with describing our look-back period for mood and anxiety. This should always be three months prior to pregnancy. We have gone through the paper and corrected this. We have also removed the statement that discusses women with new diagnosis versus previous diagnoses during pregnancy as this was confusing.	
As an extension of the point above about the severity of maternal mental illness, the look-back periods for indicators of severity also need to be defined.	The look back periods for the indicators of severity will be three year prior to the delivery date for the current pregnancy.	
Can the authors explain Gamma sensitivity analysis in the protocol?	We have included the following paragraph in our manuscript, however, we are unsure if this is to much detail: **Propensity score methods rest on the assumption that adjustments made in the models control for both measured and unmeasured confounding. Although this assumption cannot be directly tested, sensitivity to unmeasured confounding can be assed using gamma sensitivity analysis. This analysis allows us to examine how strong an unmeasured confounder would have to be to invalidate any statistically significant findings[94]. Examples of such confounders that may differ between our study groups and be associated with childhood outcomes**	Page 17

	could include maternal diet, exercise and genetic factors. Using established convention[94] we present an example of a dichotomous confounding variable. Equations 1 and 2 will be used to illustrate this sensitivity analysis. Equation 1 represents the estimated relationship between prenatal antidepressant use (PAU) and autism, and equation 2 represents the true relationship between PAU and autism, given the unmeasured confounder: β represents the true relationship between PAU and autism and γ represents the unaccounted for relationship between the confounder and Autism. Sensitivity analyses identify the minimum strength of γ that would result in a non-statistically significant β; i.e., the minimum γ that would produce a null-effect. These sensitivity analyses are outlined in detail, elsewhere[94]. $$Autism = \widehat{\beta} * PAU$$ **(equation 1)** $$Autism = \beta * PAU + \gamma * Confounder$$ **(equation 2)**	

VERSION 2 – REVIEW

REVIEWER	Salvatore Gentile ASL Salerno University of Naples Medical School "Federico II" Dept. Neurosciences Italy
REVIEW RETURNED	08-Oct-2016

GENERAL COMMENTS	I didn't find relevant changes

REVIEWER	Simone Vigod Women's College Hospital and University of Toronto, Canada.
REVIEW RETURNED	10-Oct-2016

GENERAL COMMENTS	Thank you for the opportunity to re-review this protocol. It is much more clear and consistent. A few minor points/comments for the authors to consider that might help in ensuring meaningful interpretation of the future results:

	1. There is still some lack of clarity about who will be in the study population. Looking at the codes, it appears that women with bipolar disorder will be included (i.e. outpatient ICD-9 296), so it would be more accurate to always use "mood and anxiety disorders" when describing the study population, and not to follow this with "e.g. depression/anxiety" since it is reasonable for the reader to understand that women with bipolar disorder will be included. 2. In light of the inclusion of women with bipolar disorder in the analyses, this does mean that the exclusion criteria will introduce some inconsistencies (i.e. antipsychotic exposures are excluded, but not mood stabilizers) and the authors may wish to think this through a little bit more. I read the response to my request to include (and then balance out) exposures to these other medications. If you exclude them, you cannot study them, and they do not represent an exhaustive list --- so as it stands, lithium exposures would be included, as well as valproate, for example, but not antipsychotics. I still think that it would be cleaner to consider including them and let the propensity scores balance them out, as that is part of the advantage of using propensity scores. But I respect the response. 3. In reviewing the list of antidepressants to be considered as exposures: a) There are some details that need to be clarified: i) there is an SSRI missing (I imagine that escitalopram is used in Manitoba) and an SNRI missing (I imagine that desvenlafaxine is also used); ii) generic names should be used for all drugs (e.g. fetzima and savella are brand names); iii) there are non-SRI antidepressants missing --- e.g. buproprion, mirtazapine, vortioxetine --- if these are not to be included in the exposure group, will users of these be excluded or included in the control group? b) It appears that the "other antidepressant category" are to be included as exposures. i) Are these antidepressants also theoretically related to some of the outcomes (e.g. would the serotonergic hypothesis of autism be expected to apply to an MAOI --- it technically impacts the serotonin system, but is a different breed of antidepressant)? and ii) Since TCAs may also be used in low doses for sleep, for example, will TCA users look so much like the controls that they will be overweighted in the IPTW analyses and skew the results (meaning --- there will not be very many of them but they will be strongly weighted)? and/or will full-dose TCA users and MAOI users be so different from controls due to their treatment-refractory illnesses that it will be difficult to find overlapping propensity scores? If the focus of this research is on SSRI/SNRI effects since that is where the vast majority of clinical decision-making needs to be made, then perhaps these antidepressant users should be excluded from the analyses. 4. A last point of clarification is the author statement that the propensity scores will address both measured and unmeasured confounding on page 17 (I note that later it is clarified that this is not necessarily the case). Randomization will, on average, result in both measured and unmeasured covariates being balanced between treatment groups. Conditioning on the propensity score will, on average, result in measured baseline covariates being balanced between treatment groups. However, conditioning on the propensity

	score will not necessarily balance unmeasured covariates (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3144483/). The proposed sensitivity analyses will help to estimate the potential effect of unmeasured confounding.

VERSION 2 – AUTHOR RESPONSE

Editor/Reviewer Comments	Authors Response	Reference Page Where the change appears in the revised manuscript
Reviewer 3: Dr. Simone Vigod		
1. There is still some lack of clarity about who will be in the study population. Looking at the codes, it appears that women with bipolar disorder will be included (i.e. outpatient ICD-9 296), so it would be more accurate to always use "mood and anxiety disorders" when describing the study population, and not to follow this with "e.g. depression/anxiety" since it is reasonable for the reader to understand that women with bipolar disorder will be included.	Thank you for your thoughtful comments Dr. Vigod, we appreciate the time you have taken to provide your feedback. I have gone through the manuscript and made sure that we always use “mood and anxiety disorders” and not to follow with the example of depression/anxiety.	Pg 13 – mood and anxiety disorder deleted
2. In light of the inclusion of women with bipolar disorder in the analyses, this does mean that the exclusion criteria will introduce some inconsistencies (i.e. antipsychotic exposures are excluded, but not mood stabilizers) and the authors may wish to think this through a little bit more. I read the response to my request to include (and then balance out) exposures to these other medications. If you exclude them, you cannot study them, and they do not represent an exhaustive list --- so as it stands, lithium exposures would be	We are currently in the process of conducting work that examines outcomes of children exposed to antipsychotics and benzodiazepines. We will be conducting a comparative analysis of this ongoing work with the results of this present study. Some of your concerns may be addressed through this future set of analysis. For the purposes of this protocol paper we will include your concern that the exclusion criteria may not be exhaustive in our limitations section.	Page 23

included, as well as valproate, for example, but not antipsychotics. I still think that it would be cleaner to consider including them and let the propensity scores balance them out, as that is part of the advantage of using propensity scores. But I respect the response.	Added on page 23: "Also, while we have excluded women who were prescribed antipsychotic medications, benzodiazepines, and antidepressants other than SSRIs and SNRIS, this exclusion criteria is not exhaustive. We cannot be sure that women are not taking other psychiatric drugs that may cause confounding.	
3. In reviewing the list of antidepressants to be considered as exposures: a) There are some details that need to be clarified: i) there is an SSRI missing (I imagine that escitalopram is used in Manitoba) and an SNRI missing (I imagine that desvenlafaxine is also used); ii) generic names should be used for all drugs (e.g. fetzima and savella are brand names); iii) there are non-SRI antidepressants missing --- e.g. buproprion, mirtazapine, vortioxetine --- if these are not to be included in the exposure group, will users of these be excluded or included in the control group? b) It appears that the "other antidepressant category" are to be included as exposures. i) Are these antidepressants also theoretically related to some of the outcomes (e.g. would the serotonergic hypothesis of autism be expected to apply to an MAOI --- it technically impacts the serotonin system, but is a different breed of antidepressant)? and ii) Since TCAs may also be used in low doses for sleep, for example, will TCA users look so much like the controls	I have added escitalopram and desvenlafaxine in this list and ensured there are only generic names. This list was meant to be examples and not at all an exhaustive list. When we conduct the proposed studies we will make the list that we use available to other researchers on request. Thank you for highlighting this important point. Upon consideration of your concerns and reexamination of our objectives, we agree that "other" antidepressant users should be excluded and that the focus of this research is SSRIs and SNRIs. We have made changes in the manuscript	Pg 2 – abstract changed Pg 9

that they will be overweighted in the IPTW analyses and skew the results (meaning --- there will not be very many of them but they will be strongly weighted)? and/or will full-dose TCA users and MAOI users be so different from controls due to their treatment-refractory illnesses that it will be difficult to find overlapping propensity scores? If the focus of this research is on SSRI/SNRI effects since that is where the vast majority of clinical decision-making needs to be made, then perhaps these antidepressant users should be excluded from the analyses.	that reflects this change, i.e. where we have said antidepressant use, we have changed it to SSRI and SNRI. We have changed our research objectives on page 10 to specific this, and our inclusion and exclusion criteria.	Pg 10 Pg 13 Pg 14
4. A last point of clarification is the author statement that the propensity scores will address both measured and unmeasured confounding on page 17 (I note that later it is clarified that this is not necessarily the case). Randomization will, on average, result in both measured and unmeasured covariates being balanced between treatment groups. Conditioning on the propensity score will, on average, result in measured baseline covariates being balanced between treatment groups. However, conditioning on the propensity score will not necessarily balance	Yes, we agree with you and have further clarified this in the paper by removing the statement that propensity scores make the assumption that they address both measured and unmeasured confounding and just directly explaining the sensitivity analysis.	Pg 17

unmeasured covariates (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3144483/). The proposed sensitivity analyses will help to estimate the potential effect of unmeasured confounding.		