Obstetrics & Gynecology

NOTICE: This document contains correspondence generated during peer review and subsequent revisions but before transmittal to production for composition and copyediting:

- Comments from the reviewers and editors (email to author requesting revisions)
- Response from the author (cover letter submitted with revised manuscript)*

The corresponding author has opted to make this information publicly available.

Personal or nonessential information may be redacted at the editor's discretion.

Questions about these materials may be directed to the *Obstetrics & Gynecology* editorial office: obgyn@greenjournal.org.

Date: 09/22/2022

To: "Ann M Bruno"

From: "The Green Journal" em@greenjournal.org

Subject: Your Submission ONG-22-1479

RE: Manuscript Number ONG-22-1479

Association between periviable delivery and new onset or exacerbation of existing mental health disorders

Dear Dr. Bruno:

Thank you for sending us your work for consideration for publication in Obstetrics & Gynecology. Your manuscript has been reviewed by the Editorial Board and by special expert referees. The Editors would like to invite you to submit a revised version for further consideration.

If you wish to revise your manuscript, please read the following comments submitted by the reviewers and Editors. Each point raised requires a response, by either revising your manuscript or making a clear argument as to why no revision is needed in the cover letter.

To facilitate our review, we prefer that the cover letter you submit with your revised manuscript include each reviewer and Editor comment below, followed by your response. That is, a point-by-point response is required to each of the EDITOR COMMENTS (if applicable), REVIEWER COMMENTS, and STATISTICAL EDITOR COMMENTS (if applicable) below.
The revised manuscript should indicate the position of all changes made. Please use the "track changes" feature in your document (do not use strikethrough or underline formatting).

Your submission will be maintained in active status for 21 days from the date of this letter. If we have not heard from you by Oct 13, 2022, we will assume you wish to withdraw the manuscript from further consideration.

EDITOR COMMENTS:

Please note the following:

* Please tone down findings as RRs are very modest and would not justify policy changes.

* Help us reduce the number of queries we add to your manuscript after it is revised by reading the Revision Checklist at https://journals.lww.com/greenjournal/Documents/RevisionChecklist_Authors.pdf and making the applicable edits to your manuscript.

* Figures 1-3: Please upload as figure files on Editorial Manager.

REVIEWER COMMENTS:

Reviewer #1: This well-written manuscript uses administrative data- insurance claims-- to describe the association between giving birth at periviable gestational ages and maternal mental health in the year following delivery. Modest increases in events such as psychotropic prescriptions, ED visits, etc, were observed in the women delivering periviable infants.

The main general concern is whether there is information bias of some sort. By that I mean that compared to women delivering at more advanced gestational ages, women delivering this early have extensive contact with the medical care system. In particular, most NICUs provide extensive support services for these families, and it's easy to imagine that signs of problems would be picked up and referred at milder stages than we'd see among women without as much contact. The relatively modest differences observed make this concern fairly plausible. At least, the authors should discuss the possibility of detection bias. Similarly, would there be a lower threshold for medicating, etc, knowing that the woman just experienced such an event.

Other specific concerns

1. Has the gestational age algorithm been directly validated against hand-curated medical records? If so, what's the sensitivity, specificity and predictive value.

2. How reliably can the authors identify when someone left the panel? It's not hard to imagine it might differ by pregnancy outcome, so that there's differential followup. Authors allude to knowing this, but they could be more specific.

3. Can the authors to back to the year before the pregnancy started to determine history of mental and medical diagnoses?

4. Current thinking in epidemiology is not to correct for multiple comparisons. I encourage the authors to refer to Rothman's writings on this point. Even if there's an intention to correct, Bonferroni is far too conservative. False discovery rate, or Sendak correction is preferred. And finally, all of these methods assume the comparisons are independent of each other. That seems highly unlikely here. If the outcomes are correlated, then the conservative bias is even stronger (permutation testing is needed in that case)

5. I was debating whether propensity score matching or weighting would be better than what was done here. It's main advantage is that it can eliminate individuals who are so different than they can't be meaningfully compared. In other words, for example, if there's a woman who had a periviable birth and so high a predicted probability of having one that there's no control woman whose probability was that high, the woman with that high probability would be excluded. That said, 10+ years ago there was a paper in JAMA comparing propensity score methods with conventional logistic regression and found that in large databases such as this one, it doesn't make a difference! So I'm not going to be adamant on this one.

Reviewer #2: Thank you for the opportunity to review this manuscript entitled, "Association between periviable delivery and new onset or exacerbation of existing mental health disorders." This is a retrospective cohort study of the MarketScan Commercial Research Database including all deliveries >22 weeks of gestation from 2008-2017 to examine mental health outcomes between those who had a peri-viable birth and those who did not. This manuscript is very well written and thought-out. However, I have several comments and suggestions as listed below.

Introduction: Very well-written. Would be helpful for the authors to state a hypothesis at the end of the Introduction.

Methods:
-Lines 63-64: It would be helpful for the authors to clarify why they chose 2008-2017 as the specific study period from this database, especially when mother-newborn linkage occurred after 2011. I am assuming a sample size issue but should clarify.

-Line 126: The statistical editor can weigh in, but not sure why Kruskal Wallis tests were performed for a 3-group comparison (full cohort vs. periviable vs. non-periviable) instead of just periviable vs. non-periviable 2 group-comparison as the authors described.

-Lines 128-131: Curious why parity was not included as a clinically relevant potential confounder.

-Lines 75-80: Gestational age estimation is critical since the exposure of interest is based on GA. The GA was captured and categorized based on one of three ways, but presumably GA based on newborn diagnosis codes would be the most accurate of the three methods (Tier 1 in Supplement). The authors only included liveborn gestations so presumably there was a medical record for each newborn. However, linkage was not available 2008-2011. Even outside of those years, parents in the periviable group may have been more likely to choose or have the option to NOT have their newborn resuscitated, and linkage may be even less likely in the periviable group since these data are insurance claims data. Therefore, there is the risk for selection bias since the exposure GA may be less accurate in the periviable group and it is also possible that some may have been mis-categorized. It would be helpful for the authors to provide the distribution of the method by which GA was determined in the two groups to ensure there isn't an important difference in the integrity of the data source for each comparison group. At the very least, this potential limitation needs to be addressed in the manuscript.

-Even if only liveborn deliveries are included, I would suspect that neonatal death was more frequent in the periviable group than in the non-perivable group. Neonatal death would certainly have some impact on mental health and the primary composite, and this should at least be addressed in the manuscript.

-Otherwise, study design and statistical analyses which included statistical adjustments for multiple comparisons, sensitivity analysis to only include participants with full insurance coverage, and secondary analysis to assess in 90-day intervals are thoughtful, comprehensive, and overall appropriate.

-Line 147: Would be helpful to authors to explicitly state the number of comparisons performed to get to the significant

p-value threshold of 0.002.

Discussion:
-Line 234: Based on this sentence, it seems that the authors used only the commercial insurance claims data in this data source and did not include the available but limited Medicaid and Medicare data (as described in their reference #15). This should be explicitly stated in the Methods section.

-Some additional limitations, as listed above, need to be addressed, or at least recognized, in the Discussion.

Table 1: error in count for row "urban residence". The n is same for full cohort and non-periviable group.

STATISTICAL EDITOR COMMENTS:

Table 1: Need units for age

Table 2, lines 166-168: Should indicate that the aIRR =1.16 is NS, that is, is the lower boundary CI of 1.00, exactly 1.00 or some increment above or below it. The column headings for IRR and aIRR state that there are 95%CIs. Those are not 95% CIs, but much wider than 95% CIs. They appear to be based on the Authors stated use of p < 0.002 as the threshold used (line 146). Need to clarify.

lines 189-190: Need to show this in supplemental material.

Fig 2: Need to clarify boundaries for CIs w.r.t. p < 0.002. Need to round all aIRRs and their CIs to 0.01 precision, not to 0.001

Fig 3: Need to indicate what is meant by the dashed green lines around the PVD curve and should include in figure legend a concise description of the statistical difference. Also, since the standard K-M analysis is not adjusted, why not simply call these Kaplan-Meier curves?

--

Dwight J. Rouse, MD, MSPH
Deputy Editor, Obstetrics

The Editors of Obstetrics & Gynecology

October 19, 2022

Dear Editors,

Thank you for the opportunity to revise our manuscript, "Association between periviable delivery and new onset or exacerbation of existing mental health disorders." We appreciate the review and constructive suggestions. I can confirm that I have read the 'Instructions for Authors.' Attached is the revised manuscript with tracked changes. Below are the Editor, Reviewer, and Statistical Editor comments with point-by-point responses including how and where the manuscript text was modified when able. All authors have reviewed and approve of the submitted revision.

Thank you for your ongoing consideration of our work.

Sincerely,

Ann Bruno

EDITOR COMMENTS:

Please note the following:

* Please tone down findings as RRs are very modest and would not justify policy changes.

Thank you for this feedback. The text has been modified to specifically note the findings are modest and hypothesis generating. We are open to additional language adjustments by the Editor.

Textual edits –
- Lines 176-178: "This analysis is considered exploratory and hypothesis generating with adjustments for multiple comparisons not made to avoid false negatives.[30]"
- Lines 247-248: "While the magnitude of association between our outcome and exposure was modest, our findings suggest…"
- Line 294-295: "…cohort study are modest and cannot demonstrate causality."

* Help us reduce the number of queries we add to your manuscript after it is revised by reading the Revision Checklist at https://journals.lww.com/greenjournal/Documents/RevisionChecklist_Authors.pdf and making the applicable edits to your manuscript.

Thank you for this information. The Revision Checklist has been reviewed and the manuscript follows the provided instructions.

* Figures 1-3: Please upload as figure files on Editorial Manager.

Thank you. Figures 1-3 have been uploaded as individual figure files into Editorial Manager.

REVIEWER COMMENTS:

Reviewer #1: This well-written manuscript uses administrative data- insurance claims-- to describe the association between giving birth at periviable gestational ages and maternal mental health in the year following delivery. Modest increases in events such as psychotropic prescriptions, ED visits, etc, were observed in the women delivering periviable infants.

The main general concern is whether there is information bias of some sort. By that I mean that compared to women delivering at more advanced gestational ages, women delivering this early have extensive contact with the medical care system. In particular, most NICUs provide extensive support services for these families, and it's easy to imagine that signs of problems would be picked up and referred at milder stages than we'd see among women without as much contact. The relatively modest differences observed make this concern fairly plausible. At least, the authors should discuss the possibility of detection bias. Similarly, would there be a lower threshold for medicating, etc, knowing that the woman just experienced such an event.

Thank you for these comments. We agree that patients undergoing periviable delivery may be offered additional social support services from either the maternal or neonatal health care teams. Therefore, as you state, there may be differences in detected rates of outcomes resulting from variable health care system contact. Using this data source, we were unable to

ascertain the granular details of patient use of such social support services. This limitation has been added to the Discussion.

Textual edits –
- Lines 290-292: "Additional social support through maternal and neonatal health care teams may be offered following periviable delivery; differences in detection rates for outcomes by group attributable to variable health care system contact could not be adjusted for in analyses."

Other specific concerns

1. Has the gestational age algorithm been directly validated against hand-curated medical records? If so, what's the sensitivity, specificity and predictive value.

Thank you for this question. We used a three-tiered gestational age algorithm to assign the best estimate gestational age at the time of delivery for every individual included in analysis. This algorithm has been previously described and used (lines 103-111). The three tiers including the International Classification of Diseases (ICD) 9 [h] and 10th edition codes and Current Procedural Terminology (CPT) codes are outlined in Appendix 2.

Using a dataset from Kaiser Permanente Northwest, Hornbrook et al developed and validated an algorithm for detection of pregnancy episodes and gestational age assignment. Among 32,847 individuals analyzed, medical record comparison for validation was completed in a subset of 511 individuals with 91% concordance for gestational age assignment.[1] The Hornbrook et al method reflects one tier of the algorithm used in our study. We anticipate the additional tiers used in this analysis improve accuracy of gestational age estimation. However, as MarketScan is a de-identified database, we do not have access to individual medical records to further validate the methodology.

2. How reliably can the authors identify when someone left the panel? It's not hard to imagine it might differ by pregnancy outcome, so that there's differential follow-up. Authors allude to knowing this, but they could be more specific.

Thank you for this clarification. Follow-up for individuals included in the study was censored at the first of the following: 12 months after index delivery, discontinuation of MarketScan insurance enrollment, or new pregnancy identifier present (Lines 139-141). To further address differences based on ongoing enrollment in the database, we completed a sensitivity analysis limited only to individuals with MarketScan enrollment for a complete 365 days after pregnancy outcome (Lines 173-175). Textual edits have been made for clarification.

Textual edits –
- Lines 173-175: "To address potential confounding by ongoing insurance enrollment, a sensitivity analysis was performed limited to those individuals with insurance enrollment in MarketScan for the full 365 days postpartum."

3. Can the authors go back to the year before the pregnancy started to determine history of mental and medical diagnoses?

Thank you for this question. Maternal medical and mental history was recorded for the 9 months prior to delivery (Lines 142-153). This time period for ascertainment was selected based on the 9-month gestational window for a full-term pregnancy, to reduce potential loss of sample size for

a longer lead-in period of 12+ months for MarketScan enrollment and covariate collection, and as a reasonable time period to capture codes for significant medical or mental health comorbidities. While the analysis could be repeated using a longer lead-in period of 12+ months, we do not think this would significantly change the outcome of this study.

4. Current thinking in epidemiology is not to correct for multiple comparisons. I encourage the authors to refer to Rothman's writings on this point. Even if there's an intention to correct, Bonferroni is far too conservative. False discovery rate, or Sendak correction is preferred. And finally, all of these methods assume the comparisons are independent of each other. That seems highly unlikely here. If the outcomes are correlated, then the conservative bias is even stronger (permutation testing is needed in that case)

Thank you for these comments. In our initial approach, we selected a conservative analysis approach using a Bonferroni correction for multiple comparisons. As the Reviewer outlines, Rothman, and others, provide a statistical argument for not adjusting for multiple comparisons.[2] Given the nature of our outcomes, we have decided to remove multiple testing corrections to avoid false negatives. However, we treat with caution effects with a high degree of variation. The manuscript text has been edited to reflect this change in statistical approach with the Rothman citation added. Additionally, see response to the Editor on use of language to reflect the modest effect sizes and hypothesis-generating nature of the analysis.

Textual edits –
- Deleted: "Statistical adjustments with a Bonferroni correction were made for multiple comparisons."
- Line 176-178: "A two-sided p<0.05 was considered statistically significant. This analysis is considered exploratory and hypothesis generating with adjustments for multiple comparisons not made to avoid false negatives.[30]"
- Lines 293-294: "…our findings in a retrospective cohort study are modest and cannot demonstrate causality."

5. I was debating whether propensity score matching or weighting would be better than what was done here. It's main advantage is that it can eliminate individuals who are so different that they can't be meaningfully compared. In other words, for example, if there's a woman who had a periviable birth and so high a predicted probability of having one that there's no control woman whose probability was that high, the woman with that high probability would be excluded. That said, 10+ years ago there was a paper in JAMA comparing propensity score methods with conventional logistic regression and found that in large databases such as this one, it doesn't make a difference! So I'm not going to be adamant on this one.

Thank you for this thoughtful input. We agree there are benefits to a propensity score matching or weighting approach. At this time, we have opted to not change the analysis approach.

Reviewer #2: Thank you for the opportunity to review this manuscript entitled, "Association between periviable delivery and new onset or exacerbation of existing mental health disorders." This is a retrospective cohort study of the MarketScan Commercial Research Database including all deliveries >22 weeks of gestation from 2008-2017 to examine mental health outcomes between those who had a peri-viable birth and those who did not. This manuscript is very well written and thought-out. However, I have several comments and suggestions as listed below.

Introduction: Very well-written. Would be helpful for the authors to state a hypothesis at the end of the Introduction.

Thank you. Consistent with the Reviewer request, we have added a stated hypothesis.

Textual edits –
- Lines 88-89: "We hypothesized periviable delivery would be associated with an increased risk of new onset or exacerbation of existing mental health disorders."

Methods:
-Lines 63-64: It would be helpful for the authors to clarify why they chose 2008-2017 as the specific study period from this database, especially when mother-newborn linkage occurred after 2011. I am assuming a sample size issue but should clarify.

Thank you for this clarification. We used the full timespan (2008-2017) for which maternal data was available in the MarketScan database for analysis (Lines 91-92). This provided a large, robust sample size (Lines 271-272). While maternal-newborn linkage was not available until 2011, the methodology for gestational age identification included approaches that did not require linkage (Lines 106-109, Appendix 2). Therefore, we opted to use all available data including those from 2008-2011.

-Line 126: The statistical editor can weigh in, but not sure why Kruskal Wallis tests were performed for a 3-group comparison (full cohort vs. periviable vs. non-periviable) instead of just periviable vs. non-periviable 2 group-comparison as the authors described.

Thank you for this clarification. We used Chi-squared tests for categorical variables and Kruskal-Wallis tests for continuous variables compared between individuals with and without periviable deliveries (Lines 154-155). This was a 2 group comparison and the Kruskal-Wallis test was selected as it does not require normal distribution of data. Please let us know if further clarification in the text would be beneficial.

-Lines 128-131: Curious why parity was not included as a clinically relevant potential confounder.

Thank you for this question. Multiple pregnancy complications and unexpected outcomes have been associated with postpartum depression (Lines 65-68). Based on prior literature demonstrating an association between multifetal gestation and stillbirth, we limited this analysis to liveborn singletons (Lines 112-114). Additional covariates selected *a priori* for adjustment included mode of delivery, maternal age, presence of any component of the severe maternal morbidity definition, presence of any condition within the maternal comorbidity index, and mental health history (Lines 157-160). Variables selected for adjustment were informed by prior literature (Lines 149-151). To our knowledge, parity has not been consistently associated with maternal mental health morbidity, and therefore was not selected *a priori* for adjustment.

-Lines 75-80: Gestational age estimation is critical since the exposure of interest is based on GA. The GA was captured and categorized based on one of three ways, but presumably GA based on newborn diagnosis codes would be the most accurate of the three methods (Tier 1 in Supplement). The authors only included liveborn gestations so presumably there was a medical record for each newborn. However, linkage was not available 2008-2011. Even outside of those years, parents in the periviable group may have been more likely to choose or have the option to NOT have their newborn resuscitated, and linkage may be even less likely in the periviable

group since these data are insurance claims data. Therefore, there is the risk for selection bias since the exposure GA may be less accurate in the periviable group and it is also possible that some may have been mis-categorized. It would be helpful for the authors to provide the distribution of the method by which GA was determined in the
two groups to ensure there isn't an important difference in the integrity of the data source for each comparison group. At the very least, this potential limitation needs to be addressed in the manuscript.

Thank you for this clarification. The gestational age assignment algorithm has previously been defined and used (Lines 103-111). The algorithm uses maternal-newborn linkage followed by use of International Classification of Diseases (ICD) and procedural codes (Appendix 2). The majority of gestational age assignments resulted from tier 1 in this dataset (99.9% for non-periviable deliveries vs 99.1% for periviable deliveries). This distribution has been added to the Results. This dataset does not have granular details regarding parental decision making around resuscitation following periviable delivery (Lines 288-290).

Textual edits –
  ▪ Lines 188-200: "Tier 1 of the gestational age algorithm (Appendix 2) resulted in a gestational age assignment for the majority of included deliveries (99.9% non-periviable vs 99.1% periviable). There were marginal increases in gestational age ascertainment with addition of tiers 2 and 3 (0.1% for non-periviable vs 0.9% for periviable deliveries)."

-Even if only liveborn deliveries are included, I would suspect that neonatal death was more frequent in the periviable group than in the non-perivable group. Neonatal death would certainly have some impact on mental health and the primary composite, and this should at least be addressed in the manuscript.

Thank you for this comment. Prior work has identified an association between neonatal and pediatric deaths and parental mental health diagnoses and hospitalization (Lines 77-81; 250-251). This analysis was focused on the association between periviable delivery, rather than stillbirth or neonatal death, and mental health outcomes (Lines 86-88). Granular details on neonatal outcomes and death were not available, which we acknowledge is a weakness of this analysis (Lines 288-290).

-Otherwise, study design and statistical analyses which included statistical adjustments for multiple comparisons, sensitivity analysis to only include participants with full insurance coverage, and secondary analysis to assess in 90-day intervals are thoughtful, comprehensive, and overall appropriate.

Thank you.

-Line 147: Would be helpful to authors to explicitly state the number of comparisons performed to get to the significant p-value threshold of 0.002.

Thank you for this question. Based on the feedback of Reviewer #1, we have decided to remove adjustment for multiple comparisons (see response to Reviewer #1, Question #4 above) and change the threshold for significance to <0.05.

Textual edits –

Discussion:
-Line 234: Based on this sentence, it seems that the authors used only the commercial insurance claims data in this data source and did not include the available but limited Medicaid and Medicare data (as described in their reference #15). This should be explicitly stated in the Methods section.

Thank you for this question. This analysis uses the MarketScan Commercial Research Database (Line 92). The Methods have been modified to reiterate the data source. This is also acknowledged as a limitation of this analysis in the Discussion (Lines 285-286).

Textual edits –

- Lines 94-95: "This analysis used the commercial health insurance dataset only."

-Some additional limitations, as listed above, need to be addressed, or at least recognized, in the Discussion.

Thank you for these questions and comments. Each point has been individually addressed as outlined above.

Table 1: error in count for row "urban residence". The n is same for full cohort and non-periviable group.

Thank you for identifying this error. The Table 1 row for urban residence has been updated.

Textual edits –

Characteristic	Full Cohort (N=2,300,244)	Periviable delivery (n=16,275)	Non-periviable delivery (n=2,283,969)	p
Age (years)	31 (25-34)	30 (25-34)	31 (27-34)	<0.001
Urban residence	2,012,929 (87.51)	13,209 (81.16)	1,999,720 (87.55)	<0.001

STATISTICAL EDITOR COMMENTS:

Table 1: Need units for age

Thank you for this clarification. Age units (years) have been added to Table 1.

Textual edits –

Characteristic	Full Cohort (N=2,300,244)	Periviable delivery (n=16,275)	Non-periviable delivery (n=2,283,969)	p
Age (years)	31 (25-34)	30 (25-34)	31 (27-34)	<0.001
Urban residence	2,012,929 (87.51)	13,209 (81.16)	1,999,720 (87.55)	<0.001

Table 2, lines 166-168: Should indicate that the aIRR =1.16 is NS, that is, is the lower boundary CI of 1.00, exactly 1.00 or some increment above or below it.

Thank you for this clarification. Lines 208-210 read: "The incidence of emergency department visits for mental health disorder, self-harm or suicide attempt did not differ by periviable (0.8%) versus non-periviable delivery (0.5%; aIRR 1.24, 95% CI 1.00-1.52)." The lower bound of the confidence interval for this comparison is 0.997 which is rounded and reported as 1.00. The associated non-significant p-value (0.126) is not reported in the text. The referenced aIRR 1.16 corresponds to a different outcome – new psychotropic medication (Table 2: aIRR 1.16, 95% CI 1.09-1.23) and lines 210-211 "Periviable delivery was associated with new psychotropic medication (aIRR 1.16, 95% CI 1.09-1.23)." This is statistically significant with a lower bound of the 95% confidence interval of 1.086 (rounded to 1.09) and a p-value <0.001 (not reported). No textual modifications have been made but we are open to additional input or clarification from the Statistical Editor regarding the reporting of these aIRR and associated CIs.

The column headings for IRR and aIRR state that there are 95% CIs. Those are not 95% CIs, but much wider than 95% CIs. They appear to be based on the Authors stated use of p < 0.002 as the threshold used (line 146). Need to clarify.

Thank you for this clarification. The presented columns in Table 2 include the IRR and aIRR with associated 95% confidence intervals. In response to the Reviewers (see above response to Reviewer #1), correction for multiple comparisons is no longer being presented. The Methods section has been edited consistent with this change in approach and we are now using a p<0.05 for statistical significance. The presented results with CIs are without adjustment for a lower p value threshold for multiple comparisons.

lines 189-190: Need to show this in supplemental material.

Thank you for this recommendation. The sensitivity analysis results are now included in Appendix 5.

Textual edits –
- Lines 231-232: "In sensitivity analyses limited to those who retained commercial insurance in MarketScan for a full 365 days postpartum, results were similar (Appendix 5)."
- Appendix 5 added

Appendix 5. Sensitivity analysis limited to individuals with retained commercial insurance in MarketScan for a full 365 days postpartum.

Table 1. Baseline characteristics for the analytic population with retained commercial insurance in MarketScan for a full 365 days postpartum by periviable or non-periviable delivery

Characteristic	Full Cohort (N=1,341,606)	Periviable delivery (n=9,424)	Non-periviable delivery (n=1,332,182)	p
Age (years)	31 (27-35)	30 (25-34)	31 (28-35)	<0.001
Urban residence	1,176,842 (87.72)	7,634 (81.01)	1,169,208 (87.77)	<0.001
Region				
Northeast	237,330 (17.69)	1,111 (11.79)	236,219 (17.73)	<0.001
Midwest	299,321 (22.31)	1,943 (20.62)	297,378 (22.32)	
South	520,200 (38.77)	4,259 (45.19)	515,941 (38.73)	

West	259,287 (19.33)	1,903 (20.19)	257,384 (19.32)	
Unknown	25,468 (1.90)	208 (2.21)	25,260 (1.90)	
Gestational age (delivery, in weeks)	37.7 (35.2-39.0)	24 (23.0-24.5)	37.7 (35.3-39.0)	<0.001
Mode of delivery				<0.001
Cesarean	471,050 (35.11)	4,381 (46.49)	466,669 (35.03)	
Spontaneous vaginal delivery	812,781 (60.58)	4,673 (49.59)	808,108 (60.66)	
Operative vaginal delivery	54,083 (4.03)	230 (2.44)	53,853 (4.04)	
Unknown	3,692 (0.28)	140 (1.49)	3,552 (0.27)	
Severe maternal morbidity*	15,503 (1.16)	378 (4.01)	15,125 (1.14)	<0.001
Maternal Comorbidity Index*	486,169 (36.24)	4,091 (43.41)	482,078 (36.19)	<0.001
Maternal comorbidities				
Hypertensive disorders of pregnancy	93,441 (6.96)	1,105 (11.73)	92,336 (6.93)	<0.001
Pre-existing hypertension	69,299 (5.17)	1,111 (11.79)	68,188 (5.12)	<0.001
Gestational diabetes	188,219 (14.03)	1,227 (13.02)	186,992 (14.04)	0.005
Pre-existing diabetes	52,536 (3.92)	597 (6.33)	51,939 (3.90)	<0.001
Mental health history				
Depression	42,010 (3.13)	384 (4.07)	41,626 (3.12)	<0.001
Anxiety	40,052 (2.99)	338 (3.59)	39,714 (2.98)	0.001
Psychosis	819 (0.06)	11 (0.12)	808 (0.06)	0.047
Posttraumatic stress disorder	2,639 (0.20)	19 (0.20)	2,620 (0.20)	1.0
Adjustment disorder	25,371 (1.89)	202 (2.14)	25,169 (1.89)	0.077
Pregnancy complications				
Fetal growth restriction	172,295 (12.84)	1,604 (17.02)	170,691 (12.81)	<0.001
Assisted reproductive technology (ART)	12,517 (0.93)	163 (1.73)	12,354 (0.93)	<0.001
Placental abruption	15,808 (1.18)	631 (6.70)	15,177 (1.14)	<0.001

Presented as n(%) for categorical data or median (interquartile range) for continuous data
*Presence of at least 1 component

Table 2. Primary and secondary mental health outcomes in 12 months following delivery among individuals with retained commercial insurance in MarketScan for a full 365 days postpartum with and without periviable delivery

Outcome	Periviable delivery (n=9,424)	Non-periviable delivery (n=1,332,182)	Unadjusted IRR (95% CI)	Adjusted IRR (95% CI)
Composite primary outcome	1,526 (16.19)	176,886 (13.28)	1.22 (1.07-1.38)	1.15 (1.09-1.22)
ED visit (all)	85 (0.90)	7,768 (0.58)	1.55 (0.71-2.87)	1.23 (0.93-1.58)
Suicide attempt/self-harm	1 (0.01)	123 (0.01)	†	†
Depression	42 (0.45)	3,913 (0.29)	†	†
Anxiety	49 (0.52)	4,438 (0.33)	†	†
PTSD	2 (0.02)	102 (0.01)	†	†
Psychosis	5 (0.05)	257 (0.02)	†	†
Adjustment disorder	7 (0.07)	464 (0.03)	†	†
New psychotropic medication	916 (9.72)	104,187 (7.82)	1.24 (1.05-1.46)	1.17 (1.09-1.25)
New behavioral therapy visits	900 (9.55)	107,652 (8.08)	1.18 (1.05-1.33)	1.12 (1.04-1.22)
Inpatient psychiatry services	31 (0.33)	1,975 (0.15)	2.22 (0.76-4.91)	1.68 (1.04-2.54)

Data as n(%); Unadjusted and adjusted incident rate ratio (IRR)
ED, emergency department; PTSD, posttraumatic stress disorder
†Unable to report secondary to low numbers
Adjusted for delivery mode, maternal age, severe maternal morbidity (SMM), maternal comorbidity index (MCI), and mental health history

Fig 2: Need to clarify boundaries for CIs w.r.t. p < 0.002. Need to round all aIRRs and their CIs to 0.01 precision, not to 0.001

Thank you for this clarification. Figure 2 has been updated with rounding of all IRR to 0.01 rather than 0.001 precision. The boundaries of the CIs reflect the use of a p<0.05 for statistical significance without adjustment for multiple comparisons.
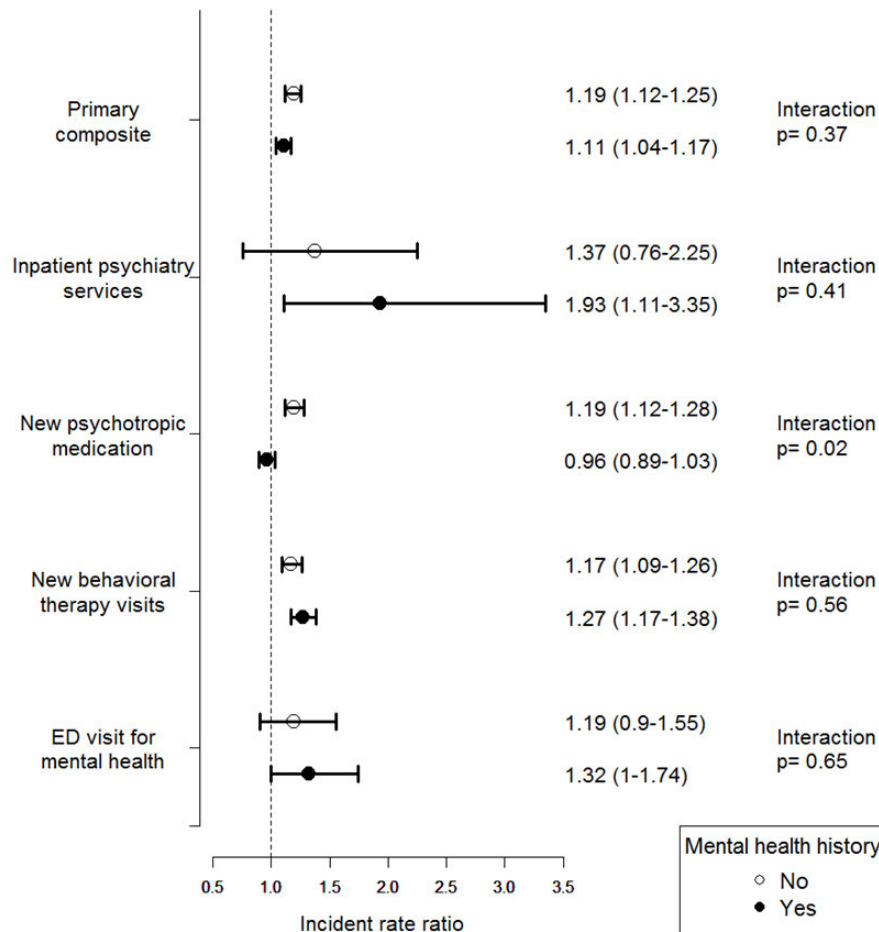
Textual edits –
- Figure 2

Fig 3: Need to indicate what is meant by the dashed green lines around the PVD curve and should include in figure legend a concise description of the statistical difference. Also, since the standard K-M analysis is not adjusted, why not simply call these Kaplan-Meier curves?

Thank you for these recommendations. The dashed lines represent the 95% confidence intervals (CIs) and this description is added to the Figure 3 legend. Consistent with the Statistical Editor recommendation, we have updated the Figure 3 Title to clearly identify this as a Kaplan-Meier curve and a statistical log-rank test result has been added.

Textual edits –

- Figure 3 title: "Kaplan-Meier Curve for primary mental health composite outcome in the 12 months following delivery by 90-day intervals by periviable delivery"
- Figure 3 legend: "Unadjusted Kaplan-Meier curves for the primary composite outcome by periviable delivery (PVD) and non-periviable delivery (Non-PVD) as time since delivery (Log rank test p<0.001). Shading represent 95% confidence intervals (CIs)."
- Figure 3
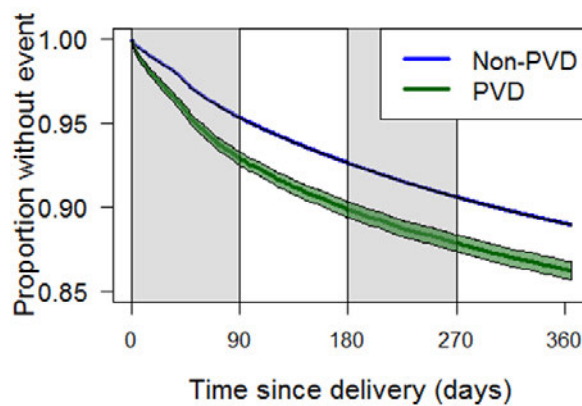
References:

1. Hornbrook MC, Whitlock EP, Berg CJ, Callaghan WM, Bachman DJ, Gold R, et al. Development of an algorithm to identify pregnancy episodes in an integrated health care delivery system. Health Serv Res 2007;42:908-27. doi: 10.1111/j.1475-6773.2006.00635.x
2. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43-6.