Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: *Maternal, Fetal, and Child Outcomes of Mental Health Treatments in Women: A Systematic Review of Perinatal Pharmacological Interventions*

Draft report available for public comment from June 15, 2020, to July 13, 2020.

Research Review Citation: Viswanathan M, Middleton JC, Stuebe A., Berkman N., Goulding AN, McLaurin-Jiang S, Dotson AB, Coker-Schwimmer M, Baker C, Voisin C, Bann C, Gaynes BN. Maternal, Fetal, and Child Outcomes of Mental Health Treatments in Women: A Systematic Review of Perinatal Pharmacologic Interventions. Comparative Effectiveness Review No. 236. (Prepared by the RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center under Contract No. 290-2015-00011-I.) AHRQ Publication No. 21-EHC001. Rockville, MD: Agency for Healthcare Research and Quality; April 2021. DOI: 10.23970/AHRQEPCCER236. Posted final reports are located on the Effective Health Care Program search page.

Comments to Draft Report

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each draft report is posted to the EHC Program website or AHRQ website for public comment for a 3- to 4-week period. Comments can be submitted via the website, mail, or email. At the conclusion of the public comment period, authors use the commentators' comments to revise the draft report.

Comments on draft reports and the authors' responses to the comments are posted for public viewing on the website approximately 3 months after the final report is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

This document includes the responses by the authors of the report to comments that were submitted for this draft report. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Summary of Peer Reviewer Comments and Author Response

This research review underwent peer review before the draft report was posted for public comment on the EHC website. Key peer review comments are summarized here.

- The initial title specified "pregnant and breastfeeding women." One reviewer noted that the title and inclusion criteria were not consistent (not all postpartum women breastfeed, Key Questions on harms pertain to women of reproductive age). In response, the EPC revised the title to include the term "perinatal" as being more descriptive of the population inclusion criteria and added more text to the review to explain the scope of the review.
- In response to comments from several reviewers, the EPC moved contextual information on the effectiveness of mental health treatments in nonpregnant populations from the results to the discussion section.
- The draft report included references to Food and Drug Administration letter categories for fetal harms. Because these labels are outdated, peer reviewers suggested removing them. The EPC added text in places to indicate that these labels were applied previously.
- Peer reviewers suggested more information for clinical context. In response, the EPC added a table to the discussion section on absolute risks of harms when available.

Public Reviewer Comments and Author Response

Commentator & Affiliation	Section	Comment	Response
American Psychological Association	General Comment	Thank you for the opportunity to comment on AHRQ's draft comparative effectiveness review *Maternal and Fetal Effects of Mental Health Treatments in Pregnant and Breastfeeding Women: A Systematic Review of Pharmacological Interventions*. We appreciate the attention given to exploring potential harms of treatment both for the mother and child and in the short and long terms. Likewise, we appreciate the explanations of potential confounding factors given the low quality of much of this information on potential harms. In the discussion, we suggest noting recent research that examined the type of information pregnant women receive from other pregnant women regarding the use of psychotropic medications, as some information received from their peers may cause harm (Denton et al., 2020). Reference: Denton, L. K., Creeley, C. E., Stavola, B., Hall, K., & Foltz, B. D. (2020). An analysis of online pregnancy message boards: Mother-to-mother advice on medication use. *Women and Birth, 33*(1), e48-58. https://doi.org/10.1016/j.wombi.2018.12.003	Added to the discussion section on clinical implications." Pregnant women often seek information outside of the clinical context on psychotropic medications; information shared on popular internet message boards may be inaccurate, contradictory, or judgmental. Clinicians and health communicators can use the findings from our review as an evidence-based source to inform and educate patients."

Commentator & Affiliation	Section	Comment	Response
American Psychological Association	Evidence Summary	Hello, We received a comment on behalf of an APA member, and we would like to add this to the comment form we submitted earlier. In the executive summary under "Main Points," could you please clarify the sixth bullet "*Lithium is more likely to be associated with overall congenital and cardiac anomalies than lamotrigine.*" While this is accurate, the mood stabilizers are used very differently in clinical practice and thus not interchangeable based solely on teratogenic potential. For example, lamotrigine does not have evidence to support its use in acute mania while lithium does. We recommend restating the sixth bullet to make it balanced. We thank you for including this addendum to our original comment submission.	This point is correct. We have clarified that this bullet refers to 1st trimester exposure and the decision to switch medications.
American Psychiatric Association	Abstract	The text should specify the point in time at which studies of brexanolone and sertraline were done (e.g., during pregnancy vs. post-partum). The specific disorder or disorders that were studied should also be specified, if relevant (e.g., major depressive disorder vs. bipolar disorder).	Added the timing to all sentences. Added the condition when possible; a lot of studies did not specify the disorder (they were based on current vs. past exposure to the drug).

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Abstract	The text that describes the observational studies is potentially misleading and may be inappropriately frightening to women and treating clinicians who do not read the remainder of the text. As written, the magnitude of these associations and the degree of confidence in these associations is not clear. People may read this sentence and assume these are significant findings and therefore choose not to use these medications when the benefits may still outweigh the risks. The statements under conclusions are more nuanced and more reflective of the uncertainties involved. It would be preferable to be more general in the results discussion and note that "observational" studies suggest possible associations between some medications and some adverse outcomes, but evidence is low quality without a clear causal relationship; rather than naming specific adverse events.	These are indeed all statistically significant associations, but statistical significance does not prove causation, so we have added a sentence describing where the evidence comes from and a caveat to interpretation. With these additions, we elect to retain the mention of these specific outcomes to add transparency to our results.
American Psychiatric Association	Evidence Summary	The new format for the executive summary does not seem as helpful as the previous format, although the icons are more appealing. Much of the text seems non-specific and the flow of the sentences and paragraphs is not as polished as the text in the main body of the report.	Thank you for your comments. This is a new format and your input will help to improve its usefulness.
American Psychiatric Association	Evidence Summary	Main Points. As with the abstract, it would be helpful to note whether these findings apply to patients treated with brexanolone, sertraline, and mood stabilizers during pregnancy or in the postpartum period and which specific disorders these findings would apply to. This is a significant problem when describing results with many of the medications throughout the document and should be addressed throughout. The statement here on mood stabilizers is slightly different from that in the abstract, which referred to discontinuation studies.	We have updated the mood stabilizer statement to match the abstract and added text when possible to indicate what the underlying disorder is. We have specified the timing of exposure, as suggested.
American Psychiatric Association	Evidence Summary	Main Points. Bullet point 5. It is not clear what associations exist with antidepressants.	Added "and adverse events."

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Evidence Summary	Main Points. Bullet point 6. The lack of mention of valproic acid in the bullet point about lithium and lamotrigine is problematic. The significant possibility of neural tube defects is often overlooked, yet many clinicians are most likely to use valproic acid when they are fearful of congenital and cardiac anomalies with lithium. Even if there is no comparative data available, there should be something noted in the main points and abstract that, although no head-to-head trials of lithium and valproic acid are available, data from other studies and patient populations suggest a significant risk of neural tube defects with valproic acid use during early pregnancy. Relative estimates of the risk of these anomalies with both lithium and valproic acid would be helpful to add as many people will not read further than the abstract or executive summary.	We have added language indicating that we did not find eligible evidence on harms of several medications, including valproate, compared with no treatment, noting that evidence is available from studies of other populations ineligible for this review.
American Psychiatric Association	Evidence Summary	Methods. It would be helpful to note the specific databases that were searched. The word “inception” presumably refers to the beginning of each database, but this could be made clearer.	Added the databases in the sentence as suggested for greater clarity.
American Psychiatric Association	Evidence Summary	KQ 1. The text refers to nine trials and six observational studies. Rather than using the vague term “trials”, it would be better to specify if this refers to randomized controlled trials or other types of trials (e.g., non-randomized, non-controlled trials). The same issue is true throughout the document.	Edited to RCT when appropriate in the evidence summary and throughout the report.
American Psychiatric Association	Evidence Summary	KQ 3. As noted above with respect to the abstract, it is important to clarify the magnitude and reliability of these associations. These statements may be taken as fact whereas the clinical significance may be relatively small. Furthermore, the presence of confounding factors is crucial to understand, but is not noted until the end of the document.	We added text directly below results to contextualize the observational evidence base.
American Psychiatric Association	Evidence Summary	KQ 4. For the discussion of lithium and lamotrigine, see comments made in reference to the abstract and ES bullet point 6.	We have modified abstract results and ES bullet point 6.
American Psychiatric Association	Evidence Summary	Limitations. 4th sentence. It is unclear whether the word “They” refers to the mental health disorders or the psychotropic medications.	Clarified to indicate that the sentence is referring to underlying mental health disorders.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Evidence Summary	Limitations. Suggest changing the word “measures” to “approaches” in terms of addressing confounding.	Edited as suggested.
American Psychiatric Association	Evidence Summary	Implications and Conclusions. It may be worth specifying that “disease severity” refers to the severity of the mental health disorder unless some other comorbid disease process is meant.	Edited as suggested.
American Psychiatric Association	Introduction	Background. 1st paragraph. The phrase “reduced use of safety and child development practices” is unclear. Also, is the punishment finding related to increased rates of harsh punishment or punishment that is of greater harshness or both?	Edited for clarity. The text now says “In addition to the negative effects on women’s health and well-being, depressive symptoms are associated with adverse parenting practices such as reduced use of safety (such as always using car seats) and child development practices (such as limiting television or video watching) and increased use of harsh punishment.”
American Psychiatric Association	Introduction	Background. 2nd paragraph. This paragraph, particularly the 2nd sentence, leads the reader to infer that pregnancy and post-partum periods are a time of substantial risk, yet the next paragraph suggests that rates of mental health disorders during this period are comparable to base rates in the population. The organization of these two paragraphs should be rearranged to emphasize the latter point. It is also not clear why any comment needs to be made about potential protective effects for bipolar disorder specifically.	We elected to retain the order of the paragraphs as is, because we think it is important to make a case for the clinical importance of the review at the start of the chapter. Instead, we revised the third paragraph for clarity and dropped the reference to the counterintuitive bipolar disorder finding.
American Psychiatric Association	Introduction	Clinical and Policy Context. It would be useful to make specific mention of the differences in clinical and policy questions that may exist for women planning to become pregnant as compared to the substantial fraction of women in the US whose pregnancies are unintended. The current wording seems to mix these two distinct scenarios.	We added a sentence to the second paragraph about unplanned pregnancies and their implications for practice.
American Psychiatric Association	Introduction	Purpose and Scope. 2nd paragraph. The sentence on KQ1 should say “consider the benefits of pharmacological treatment compared with placebo or no treatment”. The sentence on KQ3 should say KQ3 will focus on maternal and fetal/child harms of pharmacotherapy...	Revised as suggested.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	Description of Included Evidence. Table 2. From a logic perspective, the line titled “Some concerns/high risk of bias studies” should be listed just above the high risk of bias line.	Revised as suggested.
American Psychiatric Association	Results	KQ 1. The statement on depression and bipolar disorder it at the top. As written, it is expected that better evidence for the other disorders (e.g., anxiety and schizophrenia) would be listed later in the list of bullet points. To optimize clarity, it may be helpful to add “whereas for schizophrenia and anxiety, evidence was unavailable or insufficient.” to the second bullet point.	Edited as suggested.
American Psychiatric Association	Results	KQ 1. Table 3. Overall, these tables are very helpful and informative in the way they are laid out.	Thank you.
American Psychiatric Association	Results	KQ 1. Table 3. It is unclear why the anxiolytic medications are categorized as they are. Benzodiazepines are one of the commonly used classes for which there should be harms data and lumping benzodiazepines with other sedatives (e.g., barbiturates, Z-drugs) is problematic. The definition of sedative hypnotics under the table is also problematic as many clinicians inappropriately use quetiapine as a sedative/hypnotic or anxiolytic. It may be preferable to divide the anxiety related medication categories into benzodiazepines, hydroxyzine, all other anxiolytics, and other sedatives. Some mention should also be made that hydroxyzine is not a true anxiolytic and used primarily as a slightly sedating medication without significant addictive potential. The other anxiolytic and other sedatives categories should specifically exclude second generation antipsychotic medications.	We agree. We have recategorized the anxiolytics as benzodiazepines, hydroxyzine, all other anxiolytics, and other sedatives.
American Psychiatric Association	Results	KQ 1. Sertraline for Postpartum Depression: Detailed Results. 1st sentence. The sentence should say “within 1 to 3 months postpartum” rather than “or”.	Edited as suggested.
American Psychiatric Association	Results	KQ 1. Antipsychotics for Bipolar Disorder: Detailed Synthesis. 1st sentence. The sentence should read “high risk of bias” rather than “bas”.	Edited as suggested.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 2. Table 7. The item for Anxiety, All anxiolytics, Delivery mode has a light yellow box but no eligible evidence. The meaning of the shading is unclear as all other shaded boxes are with outcomes that have insufficient evidence. The Table legend gives a definition for L in terms of low evidence for benefit but none of the outcomes are rated in this fashion.	Edited as suggested, removed the legend text as suggested.
American Psychiatric Association	Results	KQ 2. Insufficient Evidence. It is not clear why "mode of delivery" would be listed as an outcome. It would seem to be better described as a characteristic of the intervention.	We were unsure if there was an association between medication use and surgical delivery/vaginal delivery, but given the lack of clarity on the relationship, we framed this broadly as mode of delivery.
American Psychiatric Association	Results	KQ 3. Overview. Readers that are unfamiliar with the way that the word "harms" is used in AHRQ reviews may find it jarring. To the average reader, "harms" tend to be major and significant. Including a definition of harms early in the document would be helpful. If other wording such as "adverse outcomes" or "adverse effects" could be used, that may be preferable to some readers.	We added this sentence to the first bullet: "Harms in the results below include any eligible adverse event; the events may not be a direct result of the exposure."
American Psychiatric Association	Results	KQ 3. Overview. Bullet points 3 and 4. The wording "one or more harm" sounds awkward. In addition, as in the abstract and executive summary, the text is non-specific and may lead readers to infer greater harms than actually exist. The word "unspecified" after SNRIs and SSRIs is also unclear as to its meaning. p. 19. KQ 3.	Revised the text to say "Interventions for which we found evidence of an association between exposure and adverse events for one or more outcome include…"
American Psychiatric Association	Results	KQ 3. Overview. Bullet point 5. Stating that there is "insufficient evidence of harms" for all of these interventions and outcomes may be similarly misleading to readers who are not used to the format and style of the AHRQ reports.	Revised to say "Interventions for which we found insufficient evidence to judge the strength of association between exposure and adverse events include…"

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Overview. Bullet point 6. As with most AHRQ reviews, the emphasis on the strength of evidence without additional information on the magnitude and likelihood of the effect is problematic, particularly for a topic such as this one. For example, the statement that "The most consistent evidence of harm is for postpartum hemorrhage" needs to be presented in the context of the likelihood that this complication will occur and the difference in the relative risk with exposure. Otherwise readers will assume that this is a significant reason to avoid treatment. These issues with writing style, phrasing, and inclusion of information are applicable throughout the review and not just to this section of the document.	We added a sentence in KQ 3 and KQ 4 overview section that the magnitude of the association varied by outcome.
American Psychiatric Association	Results	KQ 3. Evidence of Maternal Harm: Overview. Bullet point 1. Readers that are unfamiliar with the AHRQ or GRADE review methodologies may be confused by phrases such as "precise results". Since individuals may read some sections of the document but not others, it may be helpful to use clear, commonly used language whenever possible.	Revised throughout the document. Specifically, we added definitions such as "(studies large enough to detect a difference [or no difference] in effect estimates)" when talking about precision in general terms or "(i.e., the results relied on small sample sizes, few events, or had wide CIs suggestive of both benefits and harms) when talking about imprecision in general terms. We also added an explanation for the imprecision ratings in tables.
American Psychiatric Association	Results	KQ 3. Evidence of Maternal Harm: Overview. Bullet point 2. It is not clear whether these exposures are to any medication or to specific medications. Again, this is another example where the summary statement may be misleading if read or quoted separately from the more detailed text later in the document. For each of the bullet points listed in the overview sections and in the executive summary, we urge the systematic review authors to think about the implications of the bullet point as written and how it would be interpreted by clinicians and women if taken out of context. Similarly, the medicolegal implications of each statement should be considered in assuring the clarity and accuracy of each statement.	Edited to make clear that the specific exposures are detailed below.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Evidence of Maternal Harm: Overview. Bullet point 3. In discussing the potential for adverse effects with brexanolone, sedation and loss of consciousness seem like very different severities of harm. In addition to the comments on evaluating each bullet point and including information on the likelihood and magnitude of harms, and not just strength of evidence, it would be useful to consider these two harms separately rather than lumping them together.	We completely agree. This is how the FDA reported the harm and we have clarified the source in the bullet.
American Psychiatric Association	Results	KQ 3. Evidence of Maternal Harm: Overview. Bullet point 4. In addition to the comments above, the statement that association may be vary by timing of exposure warrants further explication.	Added details as suggested.
American Psychiatric Association	Results	KQ 3. Table 8. We recommend that the terms first-generation antipsychotic and second-generation antipsychotic be used rather than the terms typical and atypical antipsychotic.	We have made this change.
American Psychiatric Association	Results	KQ 3. Evidence of Maternal Harm. Detailed Synthesis. 2nd paragraph. The definition of “precise results” that is provided here is confusing. Is there a different way of saying this information that would be clearer to a typical reader?	We added text to explain precision.
American Psychiatric Association	Results	KQ 3. Dose Interruption or Reduction. If these adverse effects were dose-dependent, it would be useful to note that information.	We added this text to the results. “Rates of loss of consciousness did not appear to vary by dose and were similar among women randomized to BRX60 (2 of 38) and BRX 90 (1 of 41); in study 2, among women randomized to BRX 90, 2 of 50 experienced loss of consciousness.”
American Psychiatric Association	Results	KQ 3. Postpartum Hemorrhage. The list of medications may be confusing to non-psychiatrists as it includes categories of antidepressants (SSRIs, SNRIs) as well as listing specific medications that are within these classes. Consider whether this should be phrased differently to improve clarity.	We added the phrase “as a class” when relevant, for added clarity.
American Psychiatric Association	Results	KQ 3. Postpartum Hemorrhage. In addition to noting the absolute risk difference, it is essential to note the baseline risk of post-partum hemorrhage so that the absolute risk difference can be understood in the appropriate context.	Added as suggested.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Table 9. Although the direction of the effect is listed in parentheses in the column heading for this table, the text in the column only seems to address the strength of evidence. It would be helpful to note information on the magnitude of the effect as well, for example in terms of numbers needed to harm.	We do indicate in that column whether the effect is a benefit or a harm and indicate the magnitude of effect in terms of relative risks in the results column.
American Psychiatric Association	Results	KQ 3. Preeclampsia for SNRIs and TCAs. The inclusion of factors that contribute to preeclampsia, other than psychotropic medications, is helpful. For all of the other major complications that are discussed in the review, it would be useful to include similar information, with appropriate references, to provide context for other factors that are likely to contribute to increased risk. This will help readers interpret the data and will also help in applying the information to individual patients who may have existing risk factors that should be taken into consideration.	We believe that the subsequent sections do provide that information when relevant.
American Psychiatric Association	Results	KQ 3. Preeclampsia for SNRIs and TCAs. The connection between possible increases in serotonin and norepinephrine and possible increases in preeclampsia risk is not immediately apparent. More information should be given on this association, if it indeed exists, as it could be relevant to other medications that have yet to be studied.	We added this text: “As both serotonin and norepinephrine are vasoconstrictors, to the extent that placental ischemia contributes to the pathophysiology of preeclampsia, SNRIs could impact preeclampsia risk.”
American Psychiatric Association	Results	KQ 3. Spontaneous Abortion and SNRIs. Given the common use of SSRIs, it would be important to discuss data on spontaneous abortion and SSRIs here as well, or specifically note if no such data is available. Otherwise, there appears to be an oversight in the information that is being reported.	We covered spontaneous abortion and SSRIs in the section on insufficient evidence of maternal harms.
American Psychiatric Association	Results	KQ 3. Evidence of Fetal, Infant, or Child Harm: Overview. Bullet point 1. This bullet point is confusing and will likely be unintelligible to readers who have limited familiarity with GRADE methodology.	We added some text for clarity on the issue of plausible confounding.
American Psychiatric Association	Results	KQ 3. Evidence of Fetal, Infant, or Child Harm: Overview. Bullet point 2. If known, it would be helpful to describe the underlying reasons that there was an increased risk of NICU admission of the infant when a mother was treated with a benzodiazepine.	The reason for the association is unclear; we have explained this in greater detail in the section on this outcome. We don’t think it’s feasible to move explanatory text for each bullet back into the bullets or this will make the overview section too long.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Evidence of Fetal, Infant, or Child Harm: Overview. Bullet point 3. The current wording of this bullet point makes it seem as if women with a mental health disorder and women with prior SSRI exposure are two distinct groups of women, however, women with a prior history of SSRI exposure presumably have a history of mental health disorders in addition to a history of SSRI exposure.	Edited for clarity to cover studies where the unexposed arm had a mental health disorder but also studies where the unexposed arm was women with a prior exposure but not during pregnancy.
American Psychiatric Association	Results	KQ 3. Evidence of Fetal, Infant, or Child Harm: Overview. Bullet point 4. It is essential that risk be quantified in some fashion to avoid giving readers a misleading impression about the likelihood of harms. For example, a relatively recent review estimated a NNH of 1,000 for persistent pulmonary hypertension in the newborn (PPHN) when mothers had been treated with an SSRI (Masarwa R, Bar-Oz B, Gorelik E, Reif S, Perlman A, Matok I. Prenatal exposure to selective serotonin reuptake inhibitors and serotonin norepinephrine reuptake inhibitors and risk for persistent pulmonary hypertension of the newborn: a systematic review, meta-analysis, and network meta-analysis. Am J Obstet Gynecol. 2019;220(1):57.e1-57.e13.doi:10.1016/j.ajog.2018.08.030) and a recent expert commentary (Ornoy A, Koren G. Selective Serotonin Reuptake Inhibitors during Pregnancy: Do We Have Now More Definite Answers Related to Prenatal Exposure?. Birth Defects Res. 2017;109(12):898-908. doi:10.1002/bdr2.1078) also notes that rates are similarly high in the newborns of women with untreated depression. That impression is quite different from that which the reader would be given in reading this overview.	We added the absolute risk increase.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Evidence of Fetal, Infant, or Child Harm: Overview. Bullet point 5. This bullet point is not clearly written, particularly whether women exposed to an SSRI were more likely to have had childhood depression themselves or whether their offspring were more likely to experience childhood depression if the woman received an SSRI. This effect, if it exists, may be influenced by genetic vulnerabilities as women with more severe depression or anxiety may be more likely to be treated with an SSRI, yet also more likely to have genetic contributions to depression risk that would be passed on to their children.	We added some text regarding lack of control for severity.
American Psychiatric Association	Results	KQ 3. Evidence of Fetal, Infant, or Child Harm: Overview. Bullet point 6. As with the issue of PPHN, the issue of autism spectrum disorder and maternal SSRI treatment have been a focus of considerable interest and the nuances of the issue are not adequately addressed by this bullet point.	We added some text about residual confounding.
American Psychiatric Association	Results	KQ 3. NICU Admissions and Benzodiazepines. It would be helpful to note if the data commented on benzodiazepine dose or frequency of use (e.g., daily vs. rare use as needed).	We added this text to the section describing these results: “The study did not comment on dose or frequency.”
American Psychiatric Association	Results	KQ 3. NICU Admissions and Benzodiazepines. The point that hospital protocols may require or suggest possible NICU observation with known antenatal exposure to psychotropic medications is an important one and may warrant emphasis as a confounding factor in the overview bullet points as well as being mentioned here with the detailed synthesis of the data. It would similarly be useful to note whether other harms are more likely to be detected or suspected and then tested for in individuals with known psychotropic exposure. This is especially relevant given the paucity of data from blinded trials.	We added this point to the bullet. We also added the point about clinical protocols to the implications for research section.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. NICU Admissions and Benzodiazepines. It is not clear what would constitute non-pharmacological treatment for neonatal abstinence syndrome. In the study cited, it would be useful to know whether data were similar for neonatal abstinence due to opioid use as compared to benzodiazepine use, as the focus of this section of the document is limited to benzodiazepine use and data on neonatal opioid abstinence may be quite different. It would also be helpful to include information on neonatal adaptation syndrome, which is not discussed.	We deleted this sentence because it misinterpreted the source article.
American Psychiatric Association	Results	KQ 3. Low Apgar score and SSRIs. It is not clear why the difference in a 1.2% vs. 0.6% rates of a 5 minute Apgar less than 7 is viewed as significant. Also, the reasons that the Apgar test was developed seem less relevant than whether it does, in fact, predict need for resuscitation and/or longer-term outcomes in more recent studies. It is not clear why the information on Apgar scores is included in a separate section rather than as part of the information on respiratory difficulties where some of this information is already discussed.	We don't think it is clinically significant, but it is statistically significant, which is why we suggest that local standards may help explain score assignment. We revised the sentence about the score not being predictive somewhat and moved the text about the intent to the end of the sentence.
American Psychiatric Association	Results	KQ 3. PPHN and SSRIs. As noted previously, it would be useful to place these data in context by including NNH figures for example. Throughout the document, it would be useful if paragraphs were arranged so that the initial sentence clearly notes if there is an increased risk of harm, the confidence of evidence related to that risk, and the magnitude and severity of that risk. Spelling out this information clearly in each paragraph's topic sentence will make the key points clearer for the reader and help them in interpreting the complexities of the data.	The NNH figures are provided in a single table in the discussion where they are easy to compare against each other. We revised the paragraphs to insert the grade parenthetically when not otherwise stated.
American Psychiatric Association	Results	KQ 3. Depression in Children and SSRIs. This portion of the text seems to address the comments made about bullet point 6 but the issues should still be addressed in the overview, as noted.	We carried these edits into bullet 6 as suggested.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Insufficient Evidence of Maternal Harms. It would be more helpful to the reader if the information on maternal harms was all grouped together rather than having outcomes with insufficient evidence in a separate portion of the document. It would also be helpful to describe the reasons that the evidence was insufficient in the context of the specific psychotropic medication and outcome rather than making non-specific statements. The initial bullet under the overview is an example of text that is non-specific and not very informative as is the last bullet, which should at least be grouped together. The phrase "insufficient grades" will be confusing to those who are not well versed in GRADE methodology and alternative wording is recommended.	We elect to retain the current structure so that the most definitive evidence is presented ahead of the uncertain evidence. We removed the last bullet in the overview, because, as pointed out, it is duplicative with the first. We also added some additional text to the overview to explain the reasons for the insufficient judgment. Finally, this is a generic statement intended to address all the outcomes graded insufficient; it is then followed by detailed statements specific to exposures and outcomes.
American Psychiatric Association	Results	KQ 3. Insufficient Evidence of Maternal Harms: Detailed Synthesis. The statement that "The Results Appendix includes detailed results for these and other outcomes" could be replaced by a note to (See Appendix C). This can also be done elsewhere in the document where the same sentence is used.	Revised as suggested.
American Psychiatric Association	Results	KQ 3. Preeclampsia and SSRIs. Throughout the document, it would be helpful to minimize use of GRADE related or statistical jargon (e.g., "inclusive of the null") to enhance readability for clinicians.	Revised for clarity. This sentence explains imprecision as "(CIs for the estimate of effect span both appreciable benefit and appreciable harm)."
American Psychiatric Association	Results	KQ 3. Insufficient Evidence of Fetal, Infant, or Child Harms. As for maternal harms, it would be preferable to group all discussion of evidence on fetal, infant, or child harms together. The non-specific comments on factors contributing to insufficient evidence ratings are uninformative, as previously noted. The first and last bullet points should be grouped together, if included at all. It would be preferable to give the specific reasons for the insufficient evidence ratings with the discussion of the specific interventions.	As noted above, we elect to retain the current structure so that the most definitive evidence is presented ahead of the uncertain evidence. We removed the last bullet in the overview, because, as pointed out, it is duplicative with the first. We also added some additional text to the overview to explain the reasons for the insufficient judgment. Finally, this is a generic statement intended to address all the outcomes graded insufficient; it is then followed by detailed statements specific to exposures and outcomes.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. Insufficient Evidence of Fetal, Infant, or Child Harms: Overview. 4th bullet. As the FDA no longer uses category labelling and readers may be unfamiliar with the meaning of these categories, we suggest conveying this information without relying on the prior FDA categorization system.	We removed references to the FDA labels.
American Psychiatric Association	Results	KQ 3. Preterm Birth and SSRIs. It would be important to note here and elsewhere in the document whether "maternal severity of illness" refers to the severity of psychiatric illness or other physical health conditions.	Edited for clarity. The text now says: "Four of the six studies did not account for severity of psychiatric illness."
American Psychiatric Association	Results	KQ 3. Congenital Anomalies. This information on the prior pregnancy categories can remain here as the meaning of the categories is described. However, the FDA categorization should probably be deemphasized even here as it is an outmoded system and no longer used. The key information about treatment risks, harms, and evidence can still be provided without relying on the previous FDA designations. Throughout this section, it would be helpful to know whether the evidence looked at any type of congenital anomaly or whether the findings were related to a specific type or types of anomalies. In some portions of the text, this distinction is made, but in other portions of the text, it is not. Furthermore, given the lengthy discussion of multiple studies in this section (with 7 studies on cardiac anomalies alone) and the Forest plot of Table 2, the typical reader may find it hard to appreciate why evidence on the topic is viewed as insufficient.	Given that not all labels have been updated, we think it is still useful to have the explanation regarding the prior categories. We edited the text to use more past tense in describing these labels. The text currently clarifies that when congenital anomalies were considered overall, but we also added a parenthetical phrase indicating that this outcome included any type of anomaly. We also state that inconsistency and study limitations were reasons for the insufficient rating.
American Psychiatric Association	Results	KQ 3. Congenital Anomalies. The discussion of paroxetine and cardiac defects implies that the FDA categorization was incorrect. With the shift in the FDA use of this categorization, it would be helpful to know whether there was any shift in the FDA appraisal of paroxetine.	We have not tracked changes in FDA label status over time and so cannot add this information. We added a small note to indicate that the concern regarding cardiac malformation was at the time of labeling.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 3. No Evidence in Populations of Interest. Although this section does not seem to have a parallel section in which evidence was available, it still seems preferable to group and title the sections based on the population of interest and not on the availability of evidence or lack thereof. In bullet point one, there appears to be a typographical error “For p outcomes”. The comments above on the FDA categorization system also apply to bullet point 2.	We elected to present the evidence in order of strength of evidence (high, moderate, low) first, then insufficient, then no evidence, and we believe that this organizational structure continues to be the preferred way to focus on the most actionable evidence. All results are available in the appendix. We corrected the typographic error. We also removed references to the FDA labels.
American Psychiatric Association	Results	KQ 4. Overview. Bullet point 3. See comments about lithium, lamotrigine and valproic acid from Executive Summary comments.	We have added information on these drugs to the overview bullets.
American Psychiatric Association	Results	KQ 4. Overview. Bullet point 4. Apart from the lack of adjustment for confounding, were any of the comparison studies sufficiently large in their sample size or sufficiently reasonable in their design that they offer any useful information to clinicians and patients? If so, some qualitative mention of the studies may be useful. Particularly if the studies are often discussed or cited by other reviews or meta-analyses, it is important for clinicians and guideline developers to be able to appreciate these studies in the context of the other evidence.	We added a note that several studies draw from large databases, but the point about lack of adjustment for confounding holds, so we have not made other edits to this section.
American Psychiatric Association	Results	KQ 4. Detailed Results. As with KQ3, it would be preferable to include the evidence and insufficient evidence on a topic in the same section of the text.	As noted above, we have elected to present evidence in the order in which we think is actionable.
American Psychiatric Association	Results	KQ 4. Evidence of Fetal, Infant, or Child Harms: Overview. The second sentence of bullet point 1 is confusing as written. For bullet point 2, see prior comments related to lithium, lamotrigine and valproate that were made in reference to the abstract and Executive Summary bullet point 6.	We added some more detail to the text on precision.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Results	KQ 4. Table 13. It would be helpful to have a table or downloadable Excel workbook on the AHRQ site that includes the no treatment comparison data and the other comparator data in one table so that readers can sort by drug exposure and see all relevant comparisons at once. It would also be useful to have hyperlinked reference numbers available for each of the exposure-comparator pairs so that interested readers could quickly find the relevant studies.	We agree that this table could be useful and perhaps it can be developed as a supplemental document.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 65, 1st paragraph. See the comments made previously about the prevalence of mental health conditions among pregnant women.	We edited the introduction to address the lack of difference between pregnant and nonpregnant women and believe this statement still holds true.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 65. At the end of the first paragraph, it would be useful to be more specific about the reasons that the evidence on harms of treatment is of low quality (e.g., poorly controlled, insufficient correction for confounding variables).	Added text on confounding, for clarity.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 65, 2nd paragraph. The phrase “both were evidence bases characterized by two or more trials,” is confusing. Consider re-wording.	Edited for clarity. This now reads: “for both drugs, we found evidence bases with two or more randomized controlled trials (RCTs).”
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 65, 3rd paragraph. The “general population” implies that medications would be given to individuals regardless of indication. It would be preferable to say “the efficacy of psychotropic medications in broad groups of individuals with a mental health condition”.	Edited for clarity. The sentence now reads: “Substantial evidence exists on the efficacy of psychotropic medications across a broad spectrum of persons with mental health disorders.”
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 65, 3rd paragraph. Sentence 3. This sentence should start with “For example,”. Subsequent paragraphs highlight differences in outcomes for men as compared to women; such information would also be important to include in this paragraph.	We didn’t find any evidence of difference, but this topic does not appear to have been studied in a recent and reliable systematic review, and in the absence of a clearly citable reference, we elect to insert this language in the text.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 66, 1st full paragraph. Additional recent practice guidelines on treatment of individuals with schizophrenia include the American Psychiatric Association Practice Guideline (publication expected in September 2020; https://www.psychiatry.org/psychiatrists/practice/clinical-practice-guidelines) and the German guideline (DGPPN e.V. (ed.) for the Guideline Group: S3 Guideline for Schizophrenia. Abbreviated version (English), 2019, Version 1.0, last updated on 29 December 2019, available at: https://www.awmf.org/leitlinien/detail/ll/038-009.html)	We added the APA reference but not the German guideline because we are not able to confirm the contents.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 66, 2nd paragraph. This is a true statement but additional qualifiers may be needed to note that doses of medication may need to be adjusted to maintain comparable blood levels of medication given the physiological changes in blood volume and other pharmacokinetic considerations during pregnancy.	We added this statement “However, doses of medication may need to be adjusted to maintain comparable blood levels of medication given the physiological changes in blood volume and other pharmacokinetic considerations during pregnancy.“”
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 66, 3rd paragraph. The phrase “mode of delivery requirements” is rather cryptic; consider rephrasing.	We indicate that the mode of delivery is infusion.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 66, Last line. This statement seems inconsistent with the mixed description of the data earlier in the document.	Individual trials may not be completely consistent with one another, but the overall picture is of insufficient evidence.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 67, 1st paragraph. The potential for confounding contributions to PPHN is significant because known PPHN risk factors, such as smoking, obesity, and C-section, are all more prevalent in populations of psychiatric patients.	We added a caveat that the signals come from studies that do not fully account for confounding.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 67, 2nd paragraph. This is an important issue but is buried in the middle of this section. It may be worth moving or emphasizing in the conclusions.	Since we didn’t necessarily frame our review as an update to this narrative review, we use this citation primarily as a way of pointing out that the field hasn’t shifted; so we elect to keep it that way.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 67, 4th paragraph. “tolerance” should be “tolerability”.	Edited as suggested.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 67, Next to the last paragraph. Should read "...core illness symptoms improved more...".	Edited as suggested.
American Psychiatric Association	Discussion	Findings in Relation to the Decisional Dilemma. p. 68, middle of first full paragraph. See comments previously in document regarding p. 66, second paragraph. Medications could also be less effective in pregnancy due to physiological changes in blood volume if medication dose adjustments are not made; however, this has not been subjected to rigorous study.	We added text to implications for research on this issue.
American Psychiatric Association	Discussion	Strengths and Limitations. This section provides a nice description of the challenges of analyzing this literature.	Thank you.
American Psychiatric Association	Discussion	Applicability. The sentence "Therefore, we cannot comment on the additional effect, if any, of particular condition symptomology" may be better phrased as "Therefore, we cannot comment on the additional effect, if any, of particular symptoms of a condition".	Edited as suggested.
American Psychiatric Association	Discussion	Implications for Clinical Practice and Education. 2nd sentence. Suggest rewording to "these may be explained by residual confounding factors, rather than exposure to the drug".	Edited as suggested.
	Discussion	Implications for Clinical Practice and Education. It seems misleading and insensitive to say that harms of low Apgar scores are self-limiting, given the association with increased mortality (e.g., Chen HY, Blackwell SC, Chauhan SP. Association between Apgar score at 5 minutes and adverse outcomes among low-risk pregnancies [published online ahead of print, 2020 Apr 16]. J Matern Fetal Neonatal Med. 2020;1-8. doi:10.1080/14767058.2020.1754789, https://pubmed.ncbi.nlm.nih.gov/32609981/).	As noted in the body of the report (page 35), "The Apgar score does not predict long-term neurological outcomes and was developed to determine the immediate need for resuscitation." This language was added to the discussion of implications.
American Psychiatric Association	Discussion	Table 17. This table is very helpful.	Thank you.
American Psychiatric Association	Discussion	Implications for Health Policy. The sentence should say "Medicaid provides insurance...", not providers.	Edited as suggested.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Discussion	Implications for Health Policy and Conclusions. The sections on health policy implications and conclusions do a nice job of summarizing the challenges of clinical decision-making on this important topic with the available evidence.	Thank you.
American Psychiatric Association	Abbreviations and Acronyms	The underscores in the definition of LGA and NNNS seem to be an error.	Thank you, they have been removed.
American Psychiatric Association	Appendix A	Appendix A. The text of the document does not refer to the Appendices by their letter name. This should be checked and changed throughout. The title of Appendix A should be changed so that it matches the content. The text on this page should make clear that the FDA no longer uses their pregnancy categories. Rather than having this information in paragraph form, it may be clearer to present a table that lists the FDA category, the definition of the category and the number of drugs in the review in that category.	We have corrected the callout to the Appendices. Because Appendix A draws from undated or poorly catalogued sources, updating the table to match current labels is a challenge, so we elected to delete the table and callouts to the FDA labels in the text.
American Psychiatric Association	Appendix A	Table A-1. The title of this table should be changed to “Drug labeling related to pregnancy and nursing”. Only a small portion of this table relates to “black box warnings”. Within each type of drug, it would be helpful to list the drugs alphabetically according to the generic name. Many of the drugs do have a commonly recognized brand name that is not included in parentheses in the table.	We deleted Appendix A.
American Psychiatric Association	Appendix A	Table A-1. It is not clear why chlorpromazine is listed as discontinued as it is still available for use in the U.S.	We deleted Appendix A.
American Psychiatric Association	Appendix A	Table A-1. Trazodone has not been discontinued; it is still commonly used as a non-addicting medication to help promote sleep. The brand name Oleptro is correct for an extended release version of trazodone, but it is not a widely recognized brand name. It is not clear why it is listed when many other recognized brand names are not.	We deleted Appendix A.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Appendix A	Table A-1. It is not clear how the labeling information was chosen from the available possibilities. For example, with oxcarbazepine, the labeling information refers to Oxtellar XR, which is an extended release formulation, rather than referring to the original brand name, Trileptal, or the generic name.	We deleted Appendix A.
American Psychiatric Association	Appendix B	Details of Data Sources and Searches. The initial description notes that the searches were conducted from inception to December 11, 2018. This presumably means from the inception of each database. It would be useful to note, perhaps with the details of each search, when each database actually began.	We added some text in Appendix B detailing inception dates when available or known.
American Psychiatric Association	Appendix B	Page B-4. From the paragraph that begins “As noted in the main report...” to the end of this section, there is a lot of good information related to study strengths, limitations, confounding, etc. However, much of this information does not relate directly to the topic of this section which is “Study Selection”. Consider reorganizing this discussion so that the analytic issues are separated from study selection, per se.	Moved to the section on data synthesis, as suggested, and the section was restructured.
American Psychiatric Association	Appendix C	The appendix notes that “Additional details of the risk of bias assessments are downloadable on SRDR.” It would be helpful to have the risk of bias information available on the AHRQ website as supplementary tables (e.g., Excel workbooks, pdf). Despite the efforts invested in SRDR and the potential theoretical value of the repository, it does not yet seem to be user friendly for guideline developers or for other interested individuals. Consequently, having the detailed risk of bias information linked directly to the AHRQ site or the review appendices would be essential.	So noted, conveyed to AHRQ.
American Psychiatric Association	Appendix C	Figure C-1. The flow diagram refers to articles that were identified through text mining, yet the description of methods does not appear to give any details on how this text mining was accomplished.	Added a short paragraph as a second note under the figure.
American Psychiatric Association	Appendix C	Table C-1. The rationale for including the percentage of studies of each medication is unclear. It does not add anything beyond the number of studies. For summary purposes, it would be more useful to know the total number of individuals for whom data was available across the available studies.	Because participants may overlap across studies, we cannot generate this number with confidence.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Appendix C	Detailed Results. It would be more helpful to have the detailed results listed in a separate appendix rather than added after the excluded references list.	Although both pieces of text are still in the same appendix, we moved the list of excluded studies after the detailed results to make it easier for the reader to get the detailed results.
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-95. There should be consistency in the level of detail that is included in the overview statements. For example, the information listed in the overview for fluoxetine in depression is more elaborate than the overview statements on prior pages in this section.	Edited the fluoxetine overview bullet for consistency.
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-95. Although the study risk of bias was noted to be “some concerns”, the fact that the dose of fluoxetine wasn’t even specified seems problematic and a major study limitation.	This is an issue for nearly all the studies reporting harms. We have marked these studies as “some concerns” across the board.
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-97. In the sertraline overview and throughout the rest of this section, the phrase “a low grade of evidence” is somewhat confusing. We assume this means “strength of evidence” but the document should be consistent in wording. The GRADE terminology is confusing enough to the average reader as are the concepts of upgrading and downgrading, so using terminology consistently is important.	We revised the overview sections to call out the strength of evidence in closing parentheses.
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-97. The bullet point that begins “For reduction in depression severity...” is confusing as written. Consider whether the level of detail is excessive for an overview point and/or whether different punctuation or rewording is needed.	We split this bullet into two parts.
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-97. The Sertraline Depression overview includes a bullet point on anxiety. If this study measured anxiety symptoms in women with depression, this should be specified. If it relates to treatment of anxiety alone, it probably belongs in a different section of the document.	We specified that this result is for anxiety in women with depression.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-98. The wording of the initial sentence on this page is confusing. As written, it suggests that the treatment is trying to affect the time of onset, whereas the studies seem to have examined the effect of sertraline on response to depressive symptoms in patients with different times of depression onset post-partum.	Edited for clarity. The sentence says: "The four trials addressed postpartum depression (onset varied between 2 months and 12 months following delivery). One study looked at two intervals: onset within 3 months of delivery (the primary outcome, which is consistent with how DSM-5 defines the postpartum specifier) and onset within the more strict DSM-IV definition of 1 month postpartum."
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-101. The middle of the first paragraph seems to have a word missing in the line that begins "benefit for depression that onset…".	Edited for clarity to say "benefit for depression with onset."
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-101. The middle of the 3rd paragraph seems to have a word missing "A high risk of bias did not report…".	Edited for clarity to say: "A high risk-of-bias study did not report."
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page C-102. Because the brexanolone studies were only conducted post-partum, this should be clearly noted in the overview statements and other summaries of the findings.	We edited all overview statements to indicate the timing of exposure.
American Psychiatric Association	Appendix C	Detailed Results. KQ 1. Page. C-106. Mood Stabilizers. Detailed Results. For the publication that evaluated mood stabilizers as a class, it would be helpful to note the medications that they included. Some studies and investigators consider antipsychotic medications to have "mood-stabilizing" properties whereas others view antipsychotics as a discrete group of medications that are sometimes used in an adjunctive way for treatment of mood disorders but are not "mood stabilizers" in the way that lithium and some anticonvulsants seem to stabilize mood.	Added requested detail.
American Psychiatric Association	Appendix C	Detailed Results. KQ 2. Page C-112. Overview of Lithium vs. Paroxetine. If possible, make clear whether "specifically mood episodes" refers to fewer mood episodes, less severe mood episodes, or both.	Edited for clarity to say "benefits (specifically a mood episode)."
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-113. It would be helpful to provide examples of the cyclopyrrolones as an e.g., since most readers would not recognize these drugs by the chemical category name. Furthermore, the marketed cyclopyrrolones in the U.S. are sedative hypnotics and not anxiolytics.	Revised "cyclopyrrolones" to read "sedative-hypnotics."

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-118. As noted for the body of the document, it would be helpful if the overview statements were more specific in describing the magnitude of the risk when one is noted.	We are concerned that overview bullets will be too detailed so have elected to embed the magnitude of risk in the table and text.
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-127. There is an extra parenthesis after reference 27. The next line should read “ranging from benign…”.	Edited as suggested.
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-127. The last line gives an example of wording that is potentially confusing for individuals who don’t routinely conduct GRADE reviews. Specifically, “we rated the outcome as low for harm…”. In proofreading the document, these types of telegraphic descriptions should be written more clearly.	Edited for clarity. It now reads “Although we graded the outcome as low strength of evidence for harm, in the absence of more details on the proportion and differences between study arms for more serious outcomes, the clinical implications of this finding are unclear.”
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-128. See comments on PPHN studies related to pages 29, 33 and 67 in the body of document related to PPHN studies.	We added a note about risk factors being more common among psychiatric patients.
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-129. It is not clear from the wording of this sentence whether the musculoskeletal defects noted occurred in non-sertraline SSRIs or SSRIs including sertraline.	Edited for clarity. The text now says “One found an increased risk of craniosynostosis (RR, 2.43; 95% CI, 1.44 to 4.11; 19 exposed cases) and musculoskeletal defects (RR, 1.28; 95% CI, 1.03 to 1.58; 104 exposed cases) with nonsertraline SSRIs.”
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-134. The last sentence above Table C-19 suggesting that the incidence of autism cannot be attributed to exposure to fluoxetine seems important to emphasize in the overview of this section of the Appendix as well as in the executive summary and main body of the review (e.g., pp. 29-33).	The autism/fluoxetine exposure is not detailed in the main report because the evidence is insufficient, but in describing autism results for another drug (citalopram) that showed a signal of association, we do mention residual confounding in the evidence summary overview bullets now.
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-146. Again, as noted elsewhere, the overview statements should give estimates of risk and not just report an increased association with significant adverse outcomes. Such statements would be easy to take out of context.	As noted above, we are concerned that overview bullets will be too detailed so have elected to embed the magnitude of risk in the table and text.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-151. The last sentence just before the TCAs section is confusing and may need different punctuation or rewording.	Edited for clarity. The text now reads “Two publications with potential overlaps in the cohorts reported inconsistent and imprecise results on autism spectrum disorder. These results were rated as insufficient as a result.”
American Psychiatric Association	Appendix C	Detailed Results. KQ 3. Page C-161. The detailed summary of the studies is interesting in its conclusions as typical antipsychotics have been used in pregnant women for many years and have generally been viewed as safe.	We agree that this finding is interesting.
American Psychiatric Association	Appendix C	Detailed Results. KQ 4. Page C-206. The title SNRI Monotherapy (Venlafaxine + Desvenlafaxine) is confusing. Although venlafaxine is metabolized to desvenlafaxine, both compounds are also marketed as individual drugs and the notation “Venlafaxine + Desvenlafaxine” implies that two drugs are being given at once, precluding monotherapy.	This has been edited for clarity (“+” is replaced with “or”).
American Psychiatric Association	Appendix C	Detailed Results. KQ 4. Page C-206. The last line mentions women with epilepsy. This appears to be incorrect and is likely a cut-and-paste error, since SNRIs or SSRIs would not be used as a treatment for seizure disorder.	Actually this is not an error; the study was conducted in women with epilepsy.
American Psychiatric Association	Appendix C	Detailed Results. KQ 4. Page C-218. In the last paragraph, the results refer to a “stable dose at delivery per laboratory confirmation”. The specific aspects of the laboratory confirmation are unclear as there are few antipsychotic medications for which serum levels are routinely obtained. There can also be variations in levels even on a stable dose, so confirming a stable dose based on laboratory results may be challenging. This same approach seems to have been used for other comparisons as the wording is similar in later sections of this document. It would be helpful to be more specific about whether the laboratory tests actually confirmed that the dose was stable, or simply confirmed that the woman was taking an antipsychotic medication.	The laboratory tests did, in fact, confirm that the dose was stable. Quantification of the antipsychotic and metabolite concentrations was accomplished through high-performance liquid chromatography. Calibration curves were constructed for each antipsychotic drug assay with free human plasma by the addition of varying concentrations of the medications and their respective metabolites. We have added with each mention that inclusion criteria required receiving a stable daily dose of an antipsychotic for >5 elimination half-lives at delivery as determined by high-performance liquid chromatography.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	General Comment	Overall, the draft report is well done. It provides a detailed summary of the available information on psychotropic medications in treatment of mental health conditions in pregnancy and in the post-partum period. The information included in the report is also relevant to helping women who are planning to become pregnant make decisions about continuing or initiating medication treatment. The draft report also does a good job of highlighting gaps in our knowledge and it emphasizes the challenges in reviewing literature on medication treatment in women who are pregnant or post-partum. We very much appreciate the methodological rigor and huge amount of work that went into this systematic review.	Thank you.
American Psychiatric Association	General Comment	In general, the report does a good job of describing the importance of the topic, the research challenges involved, the pitfalls of various study designs, and the information provided by each of the included studies and the body of evidence. Throughout the report and its appendices the tables are helpful in providing an overview of the findings.	Thank you.
American Psychiatric Association	General Comment	In our detailed comments, we have pointed out specific wording in the document that was especially confusing. In addition, the requisite layout of the AHRQ reports makes them difficult to digest as there is a great deal of duplication. Dividing the evidence into key questions rather than having benefits and harms of a given treatment together is also at odds with the way in which most clinicians, patients, or guideline developers would want to review the evidence. With this document, the treatments and outcomes with insufficient evidence also seemed to be split out in separate sections from the treatments and outcomes that had some evidence (albeit limited).	Thank you for the edits. We appreciate your suggestions to organize the report in a more useful way. We agree that some of the structure can be counterproductive to a holistic understanding of the evidence, but some of our organizational decisions were intended to highlight more actionable evidence.
American Psychiatric Association	General Comment	We understand that the methodological rigor of the review relies on its use of GRADE and evidence-based practice center methodologies, and we value that rigor. However, individuals who are not steeped in GRADE methodology will have difficulty understanding many of the GRADE related terms and jargon.	We have added explanations of precision and reasons for grades in several parts of the report.

Commentator & Affiliation	Section	Comment	Response
American Psychiatric Association	General Comment	Due to the organization of the AHRQ reviews, they can be difficult to read as noted above. This contributes to substantial difficulty in finding the conclusions of the review, particularly as they would relate to clinical decision-making. Paradoxically, the detailed descriptions in Appendix C were most helpful in finding the results of the report and understanding the evidence although some sections of text and the tables in the body of the report were also helpful. As described in other comments on the draft document, the abstract and executive summary provide insufficient detail as to the findings and may actually be misleading to readers as a result. Similarly, the overview bullet points in sections in the body of the report, particularly those dealing with harms of medications, may be misleading by only emphasizing the strength of evidence and that an effect exists without providing more specifics on the magnitude of the effect and the base rates in the population. Other overview bullets provided excessive detail that was not helpful in drawing conclusions. For these reasons, we would suggest careful review of the document organization. We also suggest reading each overview bullet closely to be sure that it conveys accurate information if read as a standalone summary statement.	We appreciate the detailed review and have attempted to address the highlighted instances of insufficient detail, but also were mindful that the overview bullets cannot present all relevant details.