Author’s response to reviews

Title: Screening for depression in women during pregnancy or the first-year postpartum and in the general adult population: a protocol for two systematic reviews to update a guideline of the Canadian Task Force on Preventive Health Care

Authors:

Candyce Hamel (cahamel@ohri.ca)
Eddy Lang (Eddy.Lang@albertahealthservices.ca)
Kate Morissette (kate.morissette@canada.ca)
Andrew Beck (abeck@ohri.ca)
Adrienne Stevens (adstevens@ohri.ca)
Becky Skidmore (bskidmore@rogers.com)
Heather Colquhoun (heather.colquhoun@utoronto.ca)
John LeBlanc (John.LeBlanc@Dal.Ca)
Ainsley Moore (ainsleyemoore@gmail.com)
John Riva (rivaj@mcmaster.ca)
Brett Thombs (brett.thombs@mcgill.ca)
Ian Colman (icolman@uottawa.ca)
Sophie Grigoriadis (sophie.grigoriadis@sunnybrook.ca)
Stuart Nicholls (snicholls@ohri.ca)
Beth Potter (Beth.Potter@uottawa.ca)
Kerri Ritchie (kritchie@ottawahospital.on.ca)
Julia Robert (jrobert72@gmail.com)
Priya Vasa (VasaP@smh.ca)
Bianca Horner (hornerb@dal.ca)
Author’s response to reviews:

To the Editor:

Thank you for this valuable feedback. We have addressed each comment below.

Please also note that Appendix 3 has been removed. To provide consistent and transparent reference to working group and clinical expert selection across task force work, we have updated the response to external stakeholders and modified the text in the protocol to reflect this change.

Sincerely,

Candyce Hamel
on behalf of the authors

Reviewer report (peer-reviewers from first submission):

Review from the editor:

The protocol has been markedly improved with regards to readability and transparency of the methods for the review. A few minor points should be addressed:
1. From the authors' response regarding a cutoff for not pooling due to heterogeneity, the authors responded stating that "As stated in the protocol, we will be using both the Cochran’s Q and the I2 to determine statistical heterogeneity." While I agree that both methods are used routinely for quantifying statistical heterogeneity in a meta-analysis model, the authors still have not stated what that cutoff will be a priori. For example, will you use I2 of 50%, p <0.10 from Cochran's Q, etc. Will you calculate the uncertainty intervals around the I2 statistic, and if so, what is the cutoff that will be used to determine lack of pooling due to excessive heterogeneity.

Response: We wouldn’t pool if it was an I2 of >75%. We would use the p-value from the chi-square test as support to interpret the strength of evidence for heterogeneity. This has been added to the text for clarity.

2. In the abstract, the authors state that "there will be no language restriction and a randomized controlled trial filter will be used." While this is factually correct, it may be leading the reader to believe that you will include RCTs published in all languages (which isn't the case). Please either state the languages that will included in the abstract or remove the statement 'no language restriction'.

Response: Thank you, we have clarified this sentence and removed the ‘no language restriction’ statement.

3. In the introduction, the authors have standardized the subsections for the 'General adult population' into Prevalence, Risk factors, Consequences of depression and Current recommendations. I would suggest you clarify that 'Consequence of depression' in is the 'adult population' since you made that distinction in the pregnancy/postpartum population. Also, in the latter population, you stated 'current practice and recommendations' rather than just 'current recommendations'. Please use whichever you feel best represents that section and standardize that title for both populations.

Response: We have changed it to 'Consequence of depression' to be consistent with both populations. We have included additional information under current recommendations as there are some recommendations for the pregnant and post-partum population (usually provincially determined). We did not include this section for the adults, and there are no current recommendations.
4. In the figures, Figure 2 should come before Figure 1 as you always discussed the general adult population first.

Response: Thank you for catching this. We had changed it in the rest of the document, and in the appendices, but missed this instance. It has now been changed.

5. In the abbreviations, you elaborated with SE (Standard Error) stands for, but I don't see it used in the text. If I missed it then I apologize, else please delete this.

Response: Thank you. We did not use the abbreviation in the text, so it has been removed from the list.

Reviewer #1 (from first submission): As a reviewer of the previous version of this submission I find that the authors have satisfactorily addressed my concerns. I do note that both reviewers had the opinion that this work should perhaps be best considered not as two related, but distinct reviews. Indeed, I noted that in their own response to this issue the authors noted that "... outcomes the working group has decided on for the two separate reviews. A working group meeting was organized separately for each population ...", which implies they also consider this as two systematic reviews. I would strongly suggest the results are reported as distinct papers, given the differences in populations.

Response: Thank you for your feedback.

Reviewer #2 (new reviewer):

This review is very important, and the protocol is quite thorough. The manuscript's length is adequate, although the importance given to of each section could be revised. A few points should be considered before publication.
1. As mentioned by previous reviewers, the combination of two reviews in one protocol makes it sometimes a little difficult for the reader to follow and clearly distinguish what is specific to each review. Although the authors specified the overlap in the methods was the reason for the combined protocol, considering those are two distinct separate reviews, publication of two protocols would have been worth consideration and would have improved the clarity of each project from the readers' perspective. The combined version could be improved to further ease the reading.

Response: Thank you, we can appreciate this feedback. However, the decision was made to combine the protocols since the reviews will be used for the guideline on depression screening. We have attempted to format the document in a more intuitive way, with differences between reviews clearly stated.

Abstract

2. It is mentioned that 2 independent reviewers will screen articles. Considering the process described in the methods, the process will not be "independent".

Response: Thank you, we have clarified this in the abstract to be consistent with the methods section. For title and abstract screening, we will use the liberal accelerated method. Articles are screened at full text and ROB performed independently and in duplicate.

3. It is stated that the quality of studies will be assessed using Cochrane Risk of Bias tool. This tool does not exactly assess the quality but the risk of bias or internal validity of included studies.

Response: Agreed, this has been changed to read “assess the risk of bias of included studies…”

Introduction

4. The background/introduction section is very long (almost 6 pages) and leaves the impression of being a little disorganized or overwhelming. I think this is mainly due to the combination of the two studies. It becomes difficult to introduce both studies in an organized, clear way.
Response: Thank you, previously we tried to re-format the document as suggested by another reviewer to explain the two reviews sufficiently and improve the flow. We have aimed to follow the recommended structure from PRISMA-P (rationale, objective).

5. Moreover, the beginning of the introduction looks more like a variable definition section. I am under the impression that the introduction should be simplified and aim at explaining 1) why depression is a disease important to screen for, 2) why the pregnant and postpartum women population is particularly of interest and 3) what is currently known about the benefits and harms of screening in the general and perinatal population as well as what are the current guidelines saying. The information is already there, but there is a lot more, which might possibly undermine the clarity of the rationale for those two reviews.

Response: Some of the additional information is recommended by the CTFPHC Procedure Manual (i.e., definition, prevalence, burden) and we were following a recommendation from a previous reviewer to improve clarity. It was suggested that we streamline it for the reader by beginning with the sections on depression in the general population, then on the subpopulation of women during pregnancy and up to one-year postpartum and finalizing it with the objective of this protocol to outline the methodological process for synthesizing systematic reviews for the Task Force guidelines.

6. Line 118, I would suggest adding the words "feeling of" in the sentence: "Depression is a mood disorder characterized by states of sadness, [feeling of] worthlessness or emptiness […]"

Response: “feelings of” has been added.

7. Line 120, please consider repositioning the mention of the appendix as "poor sleep (Appendix file 1) serious" to ease the reading.

Response: The reference to the additional file has been removed. The following sentence discusses the definition, in which these symptoms are stated.

8. The MDE abbreviation is defined in 121-127, but not MDD.
Response: We have pointed to additional file 1 for further definition of MDD and added some text.


Response: This is the same reference as the following sentence. A reference has been inserted for clarity.

10. On line 165-168, it is mentioned that the annual per-capita cost among those with MDD was higher than the comparison group; could it be specified if the adjusted or non-adjusted costs are reported.

Response: These are the age- and sex-adjusted costs. We have added wording to reflect this.

11. Line 189-190 needs a reference.

Response: This is the same reference as the previous sentence. A reference to the review has been inserted.

12. Line 191, in the sentence "A recent US study in which women were interviewed, and diagnosis made using the DSM-IV criteria, found the 12-month period prevalence of MDD to be 8.4% among pregnant women, 9.3% among postpartum women, and 8.1% among non-pregnant women [26]." the 8.4% in the original paper does not refer to pregnant women but past-year pregnant women (includes currently pregnant and postpartum women).

Response: Thank you for catching this. We have updated the sentence to reflect the population.

13. Line 213, it is mentioned that postpartum depression may lead to infanticide. I could not find mention of such increased risk of infanticide in the references listed. Please clarify.
Response: We have removed this, as the reference was to a case report.

14. Lines 241-249 present eligibility criteria and seems to be repeated in the eligibility criteria tables. Please consider moving this section on the definition of a controlled trial of screening intervention in the methods section.

Response: There is a lot of misunderstanding of what exactly a screening trial is, especially among current depression screening guidelines. We felt it was important to define it explicitly in this section to frame the rest of the objectives and key questions. In the PICOs table, it is meant as a inclusion/exclusion criteria.

15. Line 265-266/272-273. Please consider adding the comparator in the research questions.

Response: Thank you for this input. We have added “versus no screening” to the key questions.

16. Line 279-299. In the objectives section, key questions 2 and 2a are mentioned. Perhaps, the criteria to undertake this subsequent review needs to be clarified. I would also suggest reducing the length of the section (and maybe refer to an appendix) as this is not part of the current reviews and will involve the development of a specific protocol. It is part of the guideline development, but not strictly related to the current protocol of systematic reviews.

Response: Thank you for this feedback. This is something that has been brought up by other reviewers, and we have provided the following response: “This systematic review is being conducted to inform a guideline on screening for depression. This protocol describes the methods for the first systematic review, which will address KQ1 and KQ1a (effectiveness of screening). We are signaling our intention to conduct a separate systematic review on KQ2 and KQ2a on patient values and preferences should the working group decide it is needed to inform the guideline. This will be decided after reviewing the evidence from the outcomes. If the Task Force working group believes that systematic review information on patient values and preferences would potentially change recommendations, beyond what is learned about values and preferences identified from focus groups conducted by the Knowledge Translation Team of St. Michaels Hospital in Toronto, Ontario supporting the development of recommendations for this guideline, then we will move forward with this additional review(s).
If we do pursue a systematic review on KQ2 and KQ2a (patient values and preferences), a separate protocol will be developed (including the relevant PICO criteria).

Methods

17. Line 301-302, the meaning of "or as methods are updated by the Task Force" is unclear.

Response: The CTFPHC Procedure Manual is a living document with the potential to update the currently accepted methods. We re-worded this section to state that the manual is a living document and if there are changes that they will be reported in the full review.

18. In line 302, It is mentioned that the Depression Working Group developed the list of outcomes and that outcomes were rated by patients. Please consider adding information on the methods used to create the list: how the group of patients that would grade the outcomes was formed, a description of those patients who contributed, were there women who experienced depression during pregnancy and/or postpartum, etc. Further details about this part of the methods could be made available through an appendix.

Response: Detailed methods used in Phase 1 projects are found in the CTFPHC’s Patient Engagement Protocol (http://canadiantaskforce.ca/methods/patient-preferences-protocol/). We have added a link within this protocol.

19. Line 325: Were the eligibility criteria identical in the previous review on the general population? If the list of criteria is (slightly) different, potentially eligible publications published prior to May 2012 could have been excluded from the previous systematic review. If so, how will this be addressed?

Response: Eligibility criteria were identical in the previous review except for the study designs included. The original review included observational studies as well. We do not feel that modification would impact study exclusion as RCTs were included in the original guideline.

20. Will screening for any diagnosis of depression (MDE, MDD, etc.? defined according to the DSM-IV or -V, etc.? be included? If all are eligible, will sensitivity/subgroup analyses be conducted?
Response: This review is focusing on the benefits and harms of screening the population for symptoms of depression, not a diagnosis of depression and those diagnosed with depression will be excluded.

21. Line 359-367. It is stated that the websites of some medical organizations will be searched. Please provide clarifications regarding the type of documents (guidelines? research articles? abstract?) searched and years of documents of interest.

Response: We will only be searching for RCTs on these websites as our study design of inclusion is specific to RCTs.

22. Line 368. The authors mention the grey literature will be confined to what can be searched in one week by one person. Considering the large list of documents that needs to be searched, I would recommend prioritizing, as also suggested by a previous reviewer. The time criteria could be problematic especially in the context of this dual review if, for example, the reviewer starts with the general population and do not have time to look at the websites specific to the pregnant and postpartum population. I would suggest targeting a few major documents/organizations to prioritize and absolutely look at, and if there is still time available, provide a list of additional websites that will potentially be consulted. Please clarify this section.

Response: As we are only including trials, we don’t believe this will be too large an undertaking. It will be fairly quick to identify RCTs compared to reports and other types of publications on these websites.

23. Line 375, a word is missing: "with the first stage [being] a broader screening" or "with [in] the first stage a broader screening".

Response: The sentence has been edited for clarity. “For each population, screening will be done in two stages. The first stage is a broad screening of the titles and abstracts.”

24. Line 394. If abstracts are excluded as part of the search strategy in Embase and Cochrane, the list of potentially relevant studies published only in abstract (line 395) will not be exhaustive.
Response: That is right. The exercise is not to identify a comprehensive list of potentially relevant studies in abstract form, but rather just provide the list of those that were excluded for that reason during full-text screening and to allow for a detailed PRISMA flow diagram. There will be an appendix in the full review which will provide the reader with a list of relevant ongoing trials.

25. Line 408-411. The data extraction process described is not independent ("Full data abstraction will be completed by one reviewer and verified by a second reviewer."). Will the data extraction be independently conducted only for the piloting of the form?

Response: Correct. It will be done independently for the pilot but will be done using the verification method for the remaining studies.

26. Line 415. "As done in other reviews" [plural] but only one reference is given. Moreover, the reference given does not provide more information about the rationale behind this method. Please provide further information to clarify this decision.

Response: We have updated the sentence to read “As done previously” with reference to the study. As we are not aware of any empirical data to be able to determine this value, in consultation with a statistician, it was decided that we would use the conservative value of 0.25.

27. Line 429-430. Please specify the customization of the risk of bias assessment planned for specific study designs. Some elements to consider in the assessment of the risk of bias in cluster RCT are provided in the Cochrane Handbook.

Response: As we are only including RCTs, we have rephrased this sentence to make it more distinct on how we will customize the ROB tool for cluster randomized trials.

28. Line 445-446. It is specified that hazard ratios will be pooled using generic inverse variance method. What will be the methods used for the pooling of other measures of association?
Response: We have added to this sentence for further detail “If it is determined to be appropriate, based on clinical similarity between studies and that the body of evidence is not at high risk of bias, data will be meta-analyzed, using random effects models for effect measures such as risk ratios and risk differences”.

29. Line 483. There is a section on "small study effects". Could the authors consider mentioning/commenting on publication bias?

Response: There is a section on small study effects under the data synthesis and statistical analysis section. We have added the following sentence to this paragraph for further detail “Funnel plot asymmetry can be used identify potential bias, as well as signal exaggeration of treatment effects in small studies [64].”

30. How will RCTs of different screening approaches (tools, definitions of depression, timing of screening, etc.) be handled? Will they all be pooled? If so, will the effect of each approach be explored separately in subgroup analyses?

Response: Yes, we will pool results from any screening tool used (whether it was a validated tool (e.g., EPDS) or an informal question). We have added the subgroup analysis for validated/non-validated screening tools. We’ve already included a subgroup analysis on timing of screening in our protocol in the table.

31. A few abbreviations need to be defined: "AHRQ SR", "IRScTn". Reference 1 needs to be corrected (P.H.A. of Canada -> Public Health Agency of Canada or PHAC).

Response: Thank you for catching these. The reference has been fixed in the reference software we are using and should now appear as Public Health Agency of Canada (PHAC). AHRQ has been expanded in the text. As it is the only instance of its use, we have not included the abbreviation.

We have added ISRCTN to the list of abbreviations and clarified in the sentence by adding ‘Registry.’ Although we did not expand the abbreviation in the text because ISRCTN is the more
common identification and, interestingly, it appears they are moving away from the expanded label because the registry scope has widened beyond randomized controlled trials.

Other items

32. Who is the guarantor of the review?

Response: Candyce Hamel is the guarantor of the protocol and review. This has been added to the contribution section.

33. Is the funding received by BT from the CIHR supporting the conduct of these review? If so, it should be listed in the Funding section.

Response: No, that funding was for other work done by Dr. Thombs, unrelated to the present work.

Figures

34. The authors mentioned that "The analytic framework depicts the structure used to address the key questions for evaluating the benefits and harms of depression screening (see Figure 1 & 2)". However, the figures do not present the comparator. Since it is meant to address the key questions, please consider adding the comparator to the figure.

Response: The purpose of the analytic framework in our sense is to explain the service we are evaluating, not to reflect the study design.

Search strategy

35. Was the adult filter tested? It is not rare to read abstracts where the word "adult" is not used to describe the population or any of the word or combination of words used in the filter. For example, a sample of adults might alternatively be described as being a sample of patients over 18 years old. The "ADULT" section of the search strategy would not allow to find such studies as it restricts the search to entries with either Adult Subheading or keywords such as "adult(s)", "adulthood", "man", "men", "woman", "women", "middle-age", "age", "elderly", "geriatric", "geriatric", "elderly", "children", "childhood", "youth", "adolescent".
"gerontology", "old-age", or "senior", or a combination of the keywords "older" with "female", "male", "patient", "person", "people" or "population".

Response: The adult filter has been used previously multiple times but is not a validated filter. As you identify, there is always a risk of missing potentially relevant items when a filter of any kind is used, should a given potentially relevant item use “alternate” terminology not used in the filter, to describe a concept. The team must weigh the benefits (i.e., lower retrieval, time/$ savings) with the risk of missing potentially relevant items. Generally, in this circumstance, the risk would be low. Additional searching (e.g., grey literature, reference lists of included studies, expert contacts, etc) is a common method used to help mitigate this risk.

36. In addition to the vocabulary used in the PREGNANCY/ANTENATAL/POSTNATAL PERIOD section of the search, gestation* could be added as the population of pregnant women can be sometimes referred to in the abstract according to gestational age or trimester of pregnancy/gestation.

Response: This synonym could certainly be added, though we think it would be fairly uncommon to find a potentially relevant study that used solely gestation* and none of the “pregnant” terminology used in the current version of the search. Our perception is that not adding this synonym would not result in a large amount of risk of missing a potentially relevant item.” In addition to this search, we are performing bibliographic and grey literature searches, and contacting experts.

Screening form

37. The authors might want to consider adding the study design in question 1 for title and abstract screening.

Response: An RCT filter has been used on the search, so that should minimize the number of non-RCTs captured. We typically do not exclude studies at title and abstract level based on study design, as authors may improperly label the study. We will however make a statement that if it clearly a cross-sectional study or case report, the reviewer should exclude the reference.
Data abstraction form

38. Items listed in Table 3 (line 477, please consider adding a title) should be listed in the data abstraction form.

Response: A title for the table has been added, and these items have been added to the draft items for data extraction.