

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Towards High Quality, Useful Practice Guidelines for Child and Youth Mental Health Disorders: Protocol for a Systematic Review and Consensus Exercise.
AUTHORS	Bennett, Kathryn; Duda, Stephanie; Brouwers, Melissa; Szatmari, Peter; Newton, Amanda; McLennan, John; Sundar, Purnima; Cleverley, Kristin; Charach, Alice; Henderson, Joanna; Courtney, Darren; Rice, Maureen

VERSION 1 – REVIEW

REVIEWER	Kimberly Hoagwood New York University School of Medicine, Department of Child and Adolescent Psychiatry
REVIEW RETURNED	03-Jul-2017

GENERAL COMMENTS	<p>This is a very comprehensive and clearly written protocol for a systematic review of practice guidelines for child and adolescent assessment, prevention and treatment. It is greatly needed and the authors do a splendid job of describing why it is important, and notably, the proposed methods for developing recommendations about the quality of practice guidelines.</p> <p>I have two questions of clarification. One is that the authors propose to include prevention guidelines, yet the focus on diagnosable conditions only may preclude their ability to actually include such PGs in this process. That would be unfortunate. I wonder if the inclusion criteria for the PGs might be broadened to include parenting skills or training, or developmental functioning (i.e., attentional control, self regulation)--in other words, areas that are potentially relevant to prevention of child/adolescent mental disorders. Secondly, there is major interest (at least in the US) on first episode psychosis and early intervention. This includes youth under 18. While very rare, inclusion of FEP, if feasible, could be important for policy decisions about quality of clinical practice. I offer these questions simply for consideration. Very important work</p>
-------------------------	--

REVIEWER	Kathryn Skivington University of Glasgow, Scotland
REVIEW RETURNED	27-Jul-2017

GENERAL COMMENTS	<p>This is a well written protocol and follows the sections in the PRISMA-P. The authors make the case for doing the review and providing a better evidence base for PGs as well as for developing recommendations for how to improve quality. There are several points, however, that should be addressed.</p> <p>Although described in the final paragraph of the introduction, to make clearer to the reader it would be helpful if the protocol had a specific section with objectives and research questions for each part of the study. Currently there is a research question for the SR but not for the user-group aspect. Is there an overall aim that takes in each of the two distinct sub-studies?</p> <p>Eligibility criteria: English language only – if the review must be limited to this then it should be acknowledged as a limitation.</p> <p>Search strategy: It is not possible to review the search strategy as full details have not been included in the protocol—the search strategy should be presented for at least one database as per the PRISMA-P guidelines. This is a major omission and should be reviewed.</p> <p>Selection process: In the initial round of screening only one methodologist will independently screen titles and abstracts. Will there be any checks on a proportion of titles/abstracts by another methodologist?</p> <p>For full-texts it is stated that two methodologists will apply inclusion criteria. It is not clear whether they will do this independently and each will review each document—please clarify. Ideally the documents will be reviewed by each reviewer. Details should be provided on how disagreements between reviewers will be handled.</p> <p>It is not clear how many ‘trained research staff’ will conduct data extraction. Although authors state that training will be included to optimize reviewer agreement for data extraction it is not clear whether any checks will be done to ensure that there is reviewer agreement e.g. having data extracted by more than one research staff for a proportion of PGs.</p> <p>The protocol justifies the use of three domains of the AGREE II to determine quality cut offs. However, does not state whether the remaining three domain scores will bear any importance. What is the point in creating the overall score (0-100%) from all six domains? Perhaps after the sentence on line 37 that states that these domains are important the authors could state how they will be used e.g. narratively in results. If the PG is calculated as being high quality on the basis of three domains, but has a low overall score will this be discussed at all?</p> <p>The data synthesis section should detail how the reviewers will synthesise and make use of the results rather than simply how the individual PGs will be summarised. It is likely that data synthesis will be narrative, and this should be detailed.</p>
-------------------------	--

	<p>RDG membership will be heterogeneous, including clinicians and service planners as well as young people and their families. Some more details on how the research team will approach the group discussions and ensure that each member feels able to voice their opinions? Those participating in a professional capacity may differ to those who are not. Some information on who will facilitate these groups (will they have expertise/training for doing this etc.?) would be useful. It is noted in the ‘potential challenges’ section that PS has expertise in engaging young people and adult family members—will she be involved in the RDGs? Additionally, it is stated that recommendations will be generated using a questionnaire, meeting interactions, and a voting system. How will the data from the groups (face-to-face meeting interactions?) be used—how will it be collected, analysed etc.?</p> <p>Although there will be training for the youth and family members I’m still not convinced that all of these RDG members will be able to fully engage with the process. This is relevant to the discussion at the RDG meetings, the online survey, and the conference call as well. Will a trial meeting be done and will there be any separate follow up, particularly with the young people involved?</p> <p>Why is ethical approval not required for the RDG meetings?</p> <p>Will the questions in the questionnaire and at the meeting be developed from the findings of the SR? Will the specific PGs identified in the SR be used for discussion? Or will the RDG be discussing ideas around PGs at a higher level? I think more information about how the specific PGs from the SR will be used would be helpful. These two aspects of the research are presented separately but it would be helpful to know how integrated they are, or whether they are stand-alone pieces of work. Is the purpose of the study to (a) produce a list of PGs that meet quality standards, and (b) produce a set of recommendations for how to improve quality of PGs for CYMH?</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Kimberly Hoagwood

Institution and Country: New York University School of Medicine, Department of Child and Adolescent Psychiatry

This is a very comprehensive and clearly written protocol for a systematic review of practice guidelines for child and adolescent assessment, prevention and treatment. It is greatly needed and the authors do a splendid job of describing why it is important, and notably, the proposed methods for developing recommendations about the quality of practice guidelines.

Thank-you.

I have two questions of clarification.

1. One is that the authors propose to include prevention guidelines, yet the focus on diagnosable conditions only may preclude their ability to actually include such PGs in this process. That would be unfortunate.

I wonder if the inclusion criteria for the PGs might be broadened to include parenting skills or training, or developmental functioning (i.e., attentional control, self regulation)--in other words, areas that are potentially relevant to prevention of child/adolescent mental disorders.

Response: This is an important point. We agree it is a potential gap in our findings re high quality CYMH guidelines.

However, after substantial reflection our view is that it is best to note this issue as a potential limitation and high priority for subsequent work, rather than incorporate it as an additional objective of the present project. Our reasons are as follows. First, if guideline developers identify prevention as one of the objectives they intend to address in their guideline, and they do a thorough job identifying all relevant prevention interventions, in theory important prevention strategies such as parenting skills or developmental function training should be identified and included. Our review could potentially document the extent to which this is happening. Second, different methods (e.g., separate search strategy and inclusion/exclusion criteria) would be required to ensure that we capture all existing relevant guidelines that focus on prevention interventions (rather than diagnoses). The result would be a parallel systematic review which unfortunately is beyond the scope of the resources currently available to support our project. Third, it is likely that only a subset of guidelines will include prevention as an objective and those PGs that do address prevention may not do a good job of identifying all relevant preventive strategies. Accordingly, the findings of this review could be used to describe how prevention is addressed in the high, minimum and poor quality guidelines and highlight important gaps in prevention intervention content in currently available guidelines. These findings would then inform the design and conduct of an adequately resourced subsequent review that rigorously addresses this very important point.

2. Secondly, there is major interest (at least in the US) on first episode psychosis and early intervention. This includes youth under 18. While very rare, inclusion of FEP, if feasible, could be important for policy decisions about quality of clinical practice. I offer these questions simply for consideration. Very important work

Response: We agree completely with the reviewers point. However, the current project does not have sufficient resources to include another topic. We think it is best to consider FEP PGs as the focus of a separate project that would need to pull in additional colleagues with relevant content expertise that is not included in our current team.

Reviewer: 2

Reviewer Name: Kathryn Skivington

Institution and Country: University of Glasgow, Scotland

This is a well written protocol and follows the sections in the PRISMA-P. The authors make the case for doing the review and providing a better evidence base for PGs as well as for developing recommendations for how to improve quality.

Thank-you.

1. There are several points, however, that should be addressed.

Although described in the final paragraph of the introduction, to make clearer to the reader it would be helpful if the protocol had a specific section with objectives and research questions for each part of the study. Currently there is a research question for the SR but not for the user-group aspect. Is there an overall aim that takes in each of the two distinct sub-studies?

Response: We agree this would be helpful to the reader. We have added a section re the research question to each part of the study.

The overall aims of this project are twofold. First we aim to document the quality of existing CYMH guidelines using systematic review methods and the AGREE-II tool (including identifying those that meet minimum and high-quality standards). Second, we aim to create recommendations for how guideline quality and usefulness can be improved in the future using the systematic review findings and the unique perspectives of four PG-user groups (clinicians, health service planners, youth and adult family members).

2. Eligibility criteria: English language only – if the review must be limited to this then it should be acknowledged as a limitation.

Response: We have added this point as a limitation.

3. Search strategy: It is not possible to review the search strategy as full details have not been included in the protocol—the search strategy should be presented for at least one database as per the PRISMA-P guidelines. This is a major omission and should be reviewed.

As noted above, we have added our search strategy for depression and anxiety guidelines as a supplementary file.

4. Selection process: In the initial round of screening only one methodologist will independently screen titles and abstracts. Will there be any checks on a proportion of titles/abstracts by another methodologist?

Response: Our methods do not call for a check at this stage of screening. As noted above in our response to the editor, only documents that clearly don't meet our inclusion criteria are excluded at this stage (e.g., title and/or abstract unambiguously indicates that the document: i) is specific to adults; or ii) does not address one or more of the target CYMH disorders; or iii) publication date is prior to 2005). All remaining documents proceed to full-text screening conducted independently by two reviewers before a final decision to include or exclude is made.

4. For full-texts it is stated that two methodologists will apply inclusion criteria. It is not clear whether they will do this independently and each will review each document—please clarify. Ideally the documents will be reviewed by each reviewer. Details should be provided on how disagreements between reviewers will be handled.

Yes, the two reviewers will work independently. The independent eligibility ratings from these reviewers will be compared and disagreements resolved through discussion with the PI. We have added full details of our methods to the text as per our response to the editor's comment.

5. It is not clear how many 'trained research staff' will conduct data extraction. Although authors state that training will be included to optimize reviewer agreement for data extraction it is not clear whether any checks will be done to ensure that there is reviewer agreement e.g. having data extracted by more than one research staff for a proportion of PGs.

Response: For this review, data extraction refers to guideline descriptive characteristics only. These variables will be used to construct 'Table 1' for resulting manuscripts. Only one reviewer is used for this data extraction. Based on our prior experience, the descriptive data required is very straightforward to extract and there is no need for checks following training and refinement of wording.

6. The protocol justifies the use of three domains of the AGREE II to determine quality cut offs. However, does not state whether the remaining three domain scores will bear any importance. What is the point in creating the overall score (0-100%) from all six domains? Perhaps after the sentence on line 37 that states that these domains are important the authors could state how they will be used e.g. narratively in results. If the PG is calculated as being high quality on the basis of three domains, but has a low overall score will this be discussed at all?

The data synthesis section should detail how the reviewers will synthesise and make use of the results rather than simply how the individual PGs will be summarised. It is likely that data synthesis will be narrative, and this should be detailed.

Response: These are important points and we have added clarification in the manuscript. First, as noted by the reviewer, guidelines will be classified as high quality or minimum quality using three AGREE-II domains (stakeholder involvement, rigor of development and editorial independence). These results will be used to answer our systematic review question: 'Among eligible PGs relevant to the assessment, prevention or treatment of common CYMH conditions, which PGs meet criteria for minimum and high quality.' Further details of our rationale are provided in the section titled 'PG High and Minimum Quality Rating Criteria'.

Second, for each guideline, the remaining individual domain scores and overall AGREE-II score will be reported for descriptive purposes. Our narrative summary will address the extent to which those guidelines deemed high quality and minimum quality also perform well on the other three domains and the overall AGREE score (i.e., achieve our score cut-offs for minimum and high quality).

Third, the individual domain scores and overall guideline quality scores will be used to describe the overall quality of CYMH guidelines. More specifically, mean domain scores and mean overall guideline quality scores for each disorder group will be calculated, and a narrative synthesis of guideline methodologic strengths and weaknesses will be conducted.

Finally, the mean guideline domain scores and overall scores will be grouped by source developer (e.g., government agency, specialty society, independent expert group or other) to explore descriptively the extent to which each of these groups are more or less likely to produce high quality guidelines than other groups.

We have revised the section of the manuscript entitled Data Synthesis and Statistical Techniques to include these additional methodologic details.

7. RDG membership will be heterogeneous, including clinicians and service planners as well as young people and their families. Some more details on how the research team will approach the group discussions and ensure that each member feels able to voice their opinions? Those participating in a professional capacity may differ to those who are not. Some information on who will facilitate these groups (will they have expertise/training for doing this etc.?) would be useful. It is noted in the 'potential challenges' section that PS has expertise in engaging young people and adult family members—will she be involved in the RDGs? Additionally, it is stated that recommendations will be generated using a questionnaire, meeting interactions, and a voting system. How will the data from the groups (face-to-face meeting interactions?) be used—how will it be collected, analysed etc.?

Response: Our methods, as described below and in the manuscript are designed to address the reviewers concerns.

We have revised the manuscript to make it clear that four separate CYMH PG PG - user groups will be formed: i) clinicians; ii) mental health service planners; iii) youth; and iv) adult family members. At the face-to-face meeting, a summary of the questionnaire results (see provisional questions in manuscript) will be presented to all participants in a large group session. Then each of the four PG user groups will work independently in small groups to formulate a set of draft recommendations

Each group will then present their draft recommendations to the full large group. The PI, with support from the research team will collate the draft recommendations presented by each of the four groups into a final provisional set for review and discussion in the final session of the day. The collated set of recommendations will include those that are common to all four groups and those that are unique to each specific PG-user group.

The level of agreement with each recommendation will be determined in an online survey conducted in two stages beginning 48 hours after the face-to-face meeting. Full details of the online survey and data summary re level of agreement can be found in the section entitled 'Generating Final Recommendations and Quantifying the Level of Consensus'. Please note that our method does not call for voting.

Regarding who will facilitate each group, we will engage experienced expert group facilitators who are not part of the research team to lead the four independent small group discussions at the face-to-face meeting. Similarly the conference call to discuss the level of agreement data from the online survey and revise the wording of each recommendation will be facilitated by a trained expert who is not part of the research team.

Preparation for and engagement in each step of the recommendation development process by youth and adult family members will be supported by 4 workshops, briefing prior to and following each workshop and as well as the provision of ongoing support for engagement. Further details are provided below in our response to question #8 and the manuscript section entitled – Training Workshops for Youth and Family Members).

We have added additional detail regarding the trained expert small group facilitators.

We have revised the manuscript to increase the clarity regarding how the questionnaire results will be used in the face-to-face meeting, how the draft recommendations will be developed and how they will be collated and reported at the end of the face-to-face meeting.

8. Although there will be training for the youth and family members I'm still not convinced that all of these RDG members will be able to fully engage with the process. This is relevant to the discussion at the RDG meetings, the online survey, and the conference call as well. Will a trial meeting be done and will there be any separate follow up, particularly with the young people involved?

Response: As noted in the manuscript, we will conduct four workshops for youth and adult family members to support capacity development and facilitate their engagement in the research process. Regarding a trial meeting, workshop four focuses specifically on preparation for the face-to-face meeting.

The first workshop will include a 1:1 orientation for each individual and a 'Terms of Reference (ToFR) document will be co-developed with the entire group. This document will capture the core engagement principals that guide our project (reciprocal relationships, co-learning, partnerships, transparency, honesty, trust) and the continuum of meaningful engagement we are utilizing (consultation, involvement, partnership and shared leadership). It will also address processes to support conflict resolution and clinical support, for example.

We will also conduct briefing prior to and following each workshop and provide ongoing support as appropriate to ensure sustained engagement.

We want to point out that our team members have substantial experience with youth and family member engagement. Sundar (in her role as Director of Knowledge Mobilization at the Ontario Centre of Excellence for Child and Youth Mental Health) has a long history of engaging with youth and families on advisory committees, co-developing a range of products and resources with youth and families (e.g., togethertolive.ca; yetoolkit.ca) and partnering with youth advisors and families. Henderson has extensive experience with youth engagement through the National Youth Advisory Committee (see for example: <https://www.ncbi.nlm.nih.gov/pubmed/28295940>). Cleverley is currently leading a Delphi study with clinicians, service providers, youth and families to develop indicators of successful transitions from child to adult mental health services (funded by the Canadian Institutes of Health Research; Bennett is a co-I).

9. Why is ethical approval not required for the RDG meetings?

Response: The research ethics office at our institution reviewed our systematic review and consensus exercise protocol and determined that it falls within quality improvement initiatives. This means that based on our national research Tri-Council Policy Statement (TCPS2, 2014), Article 2.5), we do not require ethics approval.

We have added this point to the text along with the reference for the Tri-Council Policy Statement.

10. Will the questions in the questionnaire and at the meeting be developed from the findings of the SR? Will the specific PGs identified in the SR be used for discussion? Or will the RDG be discussing ideas around PGs at a higher level?

Response: Yes - the pre-meeting questionnaire will be constructed using the results of the systematic review. The protocol text includes four provisional questions. Additional questions will be added based on the findings of the SR and input from each of the four PG-user groups.

Each question included in the questionnaire will be linked to the relevant findings from the SR. For example, for question 1 the quality ratings for each PG (i.e., domain scores) will be presented in tabular form. Each RDG participant will be asked to indicate whether each domain is relevant to defining minimum quality standards and if so what score cut-point should be used? Similarly, participants will be asked to define high quality using the AGREE-II domains. Respondents will also be asked to suggest additional criteria.

Provisional question 3 (What content and PG development processes are needed to ensure that PGs are useful to specific types of PG-users?) links to the reviewers question 'Will the specific PGs identified in the SR be used for discussion? To respond to this question on the questionnaire, each RDG participant will be provided with supplementary materials (e.g., template summary of PG content, copy of actual PG, template summary of PG development process) derived from the minimum and high quality PGs and asked to identify specific aspects of the content and developmental processes that are essential and/or missing from their user perspective.

11. I think more information about how the specific PGs from the SR will be used would be helpful. These two aspects of the research are presented separately but it would be helpful to know how integrated they are, or whether they are stand-alone pieces of work. Is the purpose of the study to (a) produce a list of PGs that meet quality standards, and (b) produce a set of recommendations for how to improve quality of PGs for CYMH?

Response: As discussed above, the specific PGs will be included in the questionnaire and as supplementary information as appropriate. As stated above and in the manuscript, the two parts of our project are fully integrated -- the findings of the SR will identify high quality PGs, inform the consensus exercise questionnaire and hence the deliberations regarding recommendations for how to improve the quality and relevance of CYMH PGs.

VERSION 2 – REVIEW

REVIEWER	Kathryn Skivington Social and Public Health Sciences Unit, University of Glasgow, Scotland
REVIEW RETURNED	06-Oct-2017
GENERAL COMMENTS	Thank you for addressing my concerns and clarifying some points about your methods. Just one outstanding point, which I think may just be confusion with wording: You state in your response "please not that our method does not call for voting", but in the Generating Recommendations section on p9 you say that "on-line voting derived from nominal group methods will be used". Is the voting you mention the 7-point scale (p10)?

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Thank you for addressing my concerns and clarifying some points about your methods.

Just one outstanding point, which I think may just be confusion with wording:

You state in your response "please not that our method does not call for voting", but in the Generating Recommendations section on p9 you say that "on-line voting derived from nominal group methods will be used". Is the voting you mention the 7-point scale (p10)?

We have edited this sentence (page 9) – voting has been removed and rating has been added as follows (thank-you to the reviewer for noticing this):

A structured questionnaire, face-to-face meeting interactions and on-line rating derived from nominal group methods will be used.

3. Kindly REMOVE FIGURES from your main document as figures should be uploaded separately.

Response: We have done this.

4. Please correct total number in the 'manuscript information' in Scholar One submission system (ex: number of tables, figures, supplementary files).

Response: Our manuscript consists of the text, references (41), 2 figures and 1 supplementary file. This is shown in the Scholar One submission system.

5. Please ensure that the authors (Szatmari, Peter and Sundar, Purnima) have unique initials in your CONTRIBUTOR SHIP statement both in your Main document and Scholar One.

Response: Szatmari is PSz and Sundar is PS in the manuscript and Scholar One.

6. December 4, 2017: Figure legends have been added at the end of the manuscript - main doc.