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)	International Clinical Practice Guidelines for Gender Minority/Trans People: Systematic Review and Quality Assessment
AUTHORS	Dahlen, Sara; Connolly, Dean; Arif, Isra; Junejo, Muhammad; Bewley, Susan; Meads, Catherine

VERSION 1 – REVIEW

REVIEWER	Gagliardi, Anna University Health Network, Toronto General Research Institute
REVIEW RETURNED	18-Feb-2021

GENERAL COMMENTS	The authors are to be commended for conducting important work, and so graciously and thoroughly responding to feedback from no less than SEVEN peer reviewers.

REVIEWER	Bachelet, Vivienne Universidad de Santiago de Chile, Facultad de Ciencias Médicas
REVIEW RETURNED	01-Mar-2021

GENERAL COMMENTS	I was one of the reviewers of your work for The BMJ. I was confused by the PDF file that was shown to me by the current journal (BMJ Open). I read the first manuscript attached in the file for download. Then there is another one further below in that same file. I was looking for the track changes but could not find them. In any event, based on the first version that comes up and which was the one that I read, I consider the manuscript sufficiently clear and well-structured to be published in its current form. I appreciate the authors did a lot of work to redo their original manuscript, which was evaluated by 7 reviewers. I don't think authors should be asked extreme efforts to get their work published, so my conclusion is that no further changes are necessary for the manuscript to be published. There is always an inherent subjectivity in reviewer assessments. Reviewer viewpoints can make manuscripts better and I think this has been achieved by the first round of peer review. I see no need to delay the publication of this manuscript by introducing further requests.

REVIEWER	Chen, Yao-Long Lanzhou University, Evidence-Based Medicine Center, School of Basic Medical Sciences
REVIEW RETURNED	03-Mar-2021

GENERAL COMMENTS	1. The background section is a bit cumbersome and it is recommended that it be simplified following this logic (first explain the importance of guidelines quality evaluation and the corresponding tools, followed by a brief introduction to the gender minority/trans people area, and then explain the necessity to evaluate the guidelines in this field). 2. The exclusion criteria of the method section are not clear, such as: how to deal with the situation of multiple versions of the same guideline published. 3. When abbreviations appear in the article for the first time, it is suggested to list their full names, such as ICD, DSM. 4. Line 209: it is suggested to explain the reasons for the supplementary search of the four journals, so as to make the results more convincing. 5. Line 214: Searching only for "clinical practice guidelines" may result in omissions, it is recommended to add "recommendations" in the title during the search procedure. 6. Line 252: Recommend more detailed reporting of AGREE II evaluation details, such as how many reviewers evaluate a guideline and how to calculate the score. 7. The report of the statistical analysis is slightly brief, so it is suggested to supplement indicators for descriptive analysis of qualitative and quantitative data and the specific statistical methods. 8. The structure of the discussion section of this version is a bit confusing. Recommend to revise this section in the following order: (1) a brief statement of the main results; (2) the corresponding analysis of the finding; for example, the score of 'applicability' domain is low, so the possible reasons should be analyzed; (3) the strengths and limitations; and (4) the research gaps. 9. The article only evaluates the methodological quality of the guidelines. Adding an evaluation of the quality of the reports is recommended to enrich the article, if possible.

REVIEWER	Feder, Gene University of Bristol, Community based medicine
REVIEW RETURNED	06-Mar-2021

GENERAL COMMENTS	Of the 7 (!) peer reviewers, I was the least critical of the original manuscript, recommending publication. The other reviewers identified problems that I had not seen. These have been comprehensively addressed by the authors, producing a higher quality systematic review. I hope you publish it.

VERSION 1 – AUTHOR RESPONSE

Comment	Our response	Line number
1. The background section is a bit cumbersome and it is recommended that it be	Thank you. We have revised and improved our introduction substantially based on the many helpful constructive suggestions made in the first round. The subheadings for our introduction now stand as: "Assessing the quality of clinical	N/A

simplified following this logic (first explain the importance of guidelines quality evaluation and the corresponding tools, followed by a brief introduction to the gender minority/trans people area, and then explain the necessity to evaluate the guidelines in this field).	practice guidelines," "Healthcare for gender minority/trans people, & "Guidelines used internationally and in the UK." We believe this is already similar enough to the new suggestion being made. The final paragraph was requested to explain our particular choice of examining international guidelines (and how they are relevant to a UK context). We are reluctant to amend the introduction further without more specifics as it might undo or go against other previous input. None of the other reviewers were dissatisfied	
2. The exclusion criteria of the method section are not clear, such as: how to deal with the situation of multiple versions of the same guideline published.	Thank you. We have added the helpful point about multiple versions.	Line 196
3. When abbreviations appear in the article for the first time, it is suggested to list their full names, such as ICD, DSM.	Many thanks. We have spelled these two out in full. If the convention of the journal is to use the abbreviations ICD and DSM, we are happy for these to be put back.	Line 198-199
4. Line 209: it is suggested to explain the reasons for the supplementary search of the four journals, so as to make the results more convincing.	Thank you. We have added a line.	Line 213-214
5. Line 214: Searching only for "clinical practice guidelines" may result in omissions, it is recommended to add "recommendations" in the title during the search procedure.	We believe this new comment that this peer reviewer did not make before is similar to one we already responded to in the first round of peer review. In the first round Peer Reviewer 1 commented: "The use of the term: "clinical practice guidelines" as only term to identify guidelines might not be enough. Many guidelines may be labeled simply as "recommendations", or "statements" "practice parameters", among others." Our reply was: "Thank you. We have not altered the manuscript but hope the following explanation satisfies: our search strategy was checked by CM (a very experienced systematic reviewer in this area) and a British Library Science information specialist. We did come across this definitional problem, which is not unique to this SR. Therefore we kept the search strategy very wide, and searched multiple places in order to find as many includable papers as possible, adding a hand-search of journals, contact with key opinion leaders for information (they did not bring up any 'statements' or 'practice parameters' that we missed) and Google searches. If the peer reviewer knows of any includable CPGs we missed we	N/A

	would be very happy to include them." Thus we believe we have already addressed this issue satisfactorily, justified not reopening the searches, and especially as there has been no suggestion of missed CPGs. Post publication, if readers identify further CPGs we would have to consider inclusion/ exclusion, scoring any missed CPGs and add a correction	
6. Line 252: Recommend more detailed reporting of AGREE II evaluation details, such as how many reviewers evaluate a guideline and how to calculate the score.	Many thanks. We have added that a minimum of two are needed for AGREE earlier in the methods. We already stated that six reviewers performed the appraisals, but have clarified this section with author initials. We have also added that the myAgree platform calculates the group scores.	170 & 254-257
7. The report of the statistical analysis is slightly brief, so it is suggested to supplement indicators for descriptive analysis of qualitative and quantitative data and the specific statistical methods.	Thank you. All the relevant data were extracted and can be found in Supplementary Tables. They do not form part of the thrust of the main paper and we did not perform a formal qualitative descriptive analysis. The report of the statistical analysis is full, even if it is brief. There are no specific statistical methods that are applicable beyond what was done by the myAgree platform and the reference is given. This is standard for other published papers using AGREE II.	N/A
8. The structure of the discussion section of this version is a bit confusing. Recommend to revise this section in the following order: (1) a brief statement of the main results; (2) the corresponding analysis of the finding; for example, the score of 'applicability' domain is low, so the possible reasons should be analyzed; (3) the strengths and limitations; and (4) the research gaps.	Many thanks. The current sub-headings of the discussion are: Statement of principal findings, Strengths and weaknesses of this study, Comparison with other studies, discussing important differences in results, Meaning of the study: possible explanations, Implications for clinicians, UK and international policymakers, and patients, Unanswered quetions and future research. The structure follows BMJ Open instructions to authors: https://bmjopen.bmj.com/pages/authors/#research "a statement of the principal findings; strengths and weaknesses of the study; strengths and weaknesses in relation to other studies, discussing important differences in results; the meaning of the study: possible explanations and implications for clinicians and policymakers; and unanswered questions and future research." We think it inappropriate to rewrite our discussion against a preferred style convention of the journal at this stage. With respect to discussion of the domain results, we thank the reviewer for highlighting that the "applicability" scores were indeed low, according to AGREE II criteria (text taken from the Applicability Section): https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf 18. The guideline describes facilitators and barriers to its application. 19. The guideline provides advice and/or tools on how the recommendations can be put into practice. 20. The potential resource implications of applying the recommendations have been considered.	N/A

	21. The guideline presents monitoring and/or auditing criteria. We chose to focus the text on the two domains selected in the stakeholder prioritisation exercise (stakeholder involvement and rigour of development). There seems less reason to concentrate on applicability or the remaining other domains without that rationale, which might be more wordy and speculative. Causality is not something we discussedfor any of the domain scores and we think this strays too far beyond the findings, rather than saying what we found and talking in more general terms about how guideline developers might approach improvements. It might be something readers comment on afterwards.	
9. The article only evaluates the methodological quality of the guidelines. Adding an evaluation of the quality of the reports is recommended to enrich the article, if possible.	Thank you. We are unclear what the reviewer means when bringing up another new point at this late stage and again feel this is beyond the scope of the paper, as the AGREE II instrument is validated to assess quality of the CPGs (in six key areas: Scope and Purpose, Stakeholder involvement, Rigour of Development, Clarity of Presentation, Applicability, Editorial Independence). We have not individually assessed the strength of each recommendation within every of the CPGs (as this was not a project aim), but have extracted and presented these in supplementary Tables, as well as Mortality and Quality of Life data. If readers are interested in this further question that the reviewer raises, we hope they can use the work and primary sources	N/A