PEER REVIEW HISTORY

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

(This paper received three reviews from its previous journal but only two reviewers agreed to published their review.)

ARTICLE DETAILS

TITLE (PROVISIONAL)	Evidence Synthesis, Practice Guidelines and Real-World Prescriptions of New Generation Antidepressants in the Treatment of Depression: A Protocol for Cumulative Network Meta-analyses and Meta-epidemiological Study.
AUTHORS	Luo, Yan; Chaimani, Anna; Kataoka, Yuki; Ostinelli, Edoardo Giuseppe; Ogawa, Yusuke; Cipriani, Andrea; Salanti, Georgia; Furukawa, Toshi A.

VERSION 1 – REVIEW

REVIEWER	Kee-Hsin Chen Post-Baccalaureate Program in Nursing, College of Nursing, Taipei Medical University, Taiwan
REVIEW RETURNED	22-May-2018

GENERAL COMMENTS	This is a manuscript proposing a study protocol in comparison of the efficacy and acceptability of new generation antidepressants by network meta-analyses, recommendation from international clinical practice guidelines, and prescription patterns of antidepressants data by representative samples in USA. The prospective study results may provide the situation between evidence and real-world clinical practices. In addition, it may give future direction to narrow the gaps from evidence to implementation. I believe the study is a good start; however, there are some comments worthwhile to mention: 1. In regard to identify and extract recommendations in guidelines, research team considers extracting recommendations for specific antidepressants from internationally representative guidelines (APA, NICE, BAP and CANMAT). I would suggest describing the details of inclusion and exclusion criteria in this protocol, for example, inclusion of guidelines developed from rigor guideline development methodology and exclusion of expert consensus guidelines. 2. The appraisal tool (or method) used to evaluate the quality of included guidelines should be stated. In addition, it is desirable to have detail items of data extraction described in the protocol. 3. For real-world prescriptions data, research team propose to search all 24 listed antidepressants in MEPS at time of data extraction. It would be desirable to describe in details about the statistics methods for data analysis. 4. Some confounders that may influence the prescription should be of concern, for example, certain antidepressants may be effective in

	some patients for years, and with agreement of the treating physicians, they might not be willing to change to newer medications. It also may be related to physician' prescription pattern, patient preferences and local regulations, etc. Multiple use of different antidepressants over years is another issue. All these may be a major challenge in data analysis and interpretation.

REVIEWER	Julian Elliott Cochrane Australia, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia I have collaborated with Prof Salanti on a Cochrane project investigating statistical approaches to the updating of meta-analyses.
REVIEW RETURNED	02-Jul-2018

GENERAL COMMENTS	The authors propose a retrospective study comparing network meta-analyses (NMA) of anti-depressants at 5 yearly intervals with major clinical practice guidelines (CPGs) and US prescription data. The authors should be commended for the objectives of this meta-epidemiological study, which aims to highlight important gaps between the accumulation of research and subsequent guideline recommendations and practice. The key challenge for this study is the complexity of guideline development and practice behaviours. Efficacy and acceptability are some, but not all the factors that must be considered when developing guideline recommendations. Similarly, there is well characterised complexity in the translation of guideline recommendations into prescription practice. The authors acknowledge these and other limitations, but I am still left with some concerns regarding the conclusions that will be possible when the planned comparisons have been completed. I have a recommendation and two additional points for consideration. First, I believe the plan to repeat the NMA at 5-yearly intervals is problematic. As I understand the aims of the project, the authors plan to compare the results of each NMA with subsequent CPGs, e.g. perform a '2005 NMA' and compare findings to CPG recommendations published between 2005 and 2009. However, for CPGs published towards the end of each 5-year period (e.g. NICE 2009, CANMAT 2009), there was substantial research available that will not be included in the '2005 NMA', further weakening the conclusions that can be made from these comparisons. With the caveat of feasibility, I suggest that to achieve the study aims, the NMAs should be performed at more frequent intervals (optimally for each year). Second, for consideration only, I wonder about an analysis that focusses not so much on 'gaps' between the NMAs and CPGs, but time from a change in NMA ranking to change in CPG recommendation. This would provide important meta-epidemiological information on delays in the evidence ecosystem over the last 20-30 years. Third, the weaknesses of the MEPS dataset acknowledged by the authors are significant, particularly that the data are not able to be disaggregated by initial versus continuing treatment. This may lead

	to erroneous conclusions about delays in prescription practice change. I wonder whether the analysis of this dataset is worth the effort, or could at least be supplemented by a more appropriate dataset with data on antidepressant initiations. Additional points: • The proposed study is not a NMA alone, but given the important role NMA plays in the study, the authors could consider including NMA in the title of the manuscript. • 'Extracting' and coding CPG recommendations is a complex task and needs further description. • The definition of trial completion year seems to suggest that for studies with a known completion year after the proposed 'NMA year', data published after the 'NMA year' will be included in the relevant NMA. Please clarify. • Please clarify the text to indicate that individuals taking mood stabilisers and antipsychotics will be excluded from denominator data. • Please clarify whether the MEPS data will be compared to APA recommendations only. • For the statement regarding regarding systematic reviews being out of date at time of publication the authors should reference Shojania, Ann Int Med 2007, which estimated that 7% of SRs were out of date when published. • A better citation is needed for the MEPS dataset. • The potential contribution of this work to living guidelines could reference recent publications on this topic. • Please describe the role of the funder in the study.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1
(Reviewer Name: Kee-Hsin Chen
Institution and Country: Post-Baccalaureate Program in Nursing, College of Nursing, Taipei Medical University, Taiwan)
This is a manuscript proposing a study protocol in comparison of the efficacy and acceptability of new generation antidepressants by network meta-analyses, recommendation from international clinical practice guidelines, and prescription patterns of antidepressants data by representative samples in USA. The prospective study results may provide the situation between evidence and real-world clinical practices. In addition, it may give future direction to narrow the gaps from evidence to implementation. I believe the study is a good start; however, there are some comments worthwhile to mention.
RESPONSE:
We thank the reviewer for these positive comments.

1. **In regard to identify and extract recommendations in guidelines, research team considers extracting recommendations for specific antidepressants from internationally representative guidelines (APA, NICE, BAP and CANMAT). I would suggest describing the details of inclusion and exclusion criteria in this protocol, for example, inclusion of guidelines developed from rigor guideline development methodology and exclusion of expert consensus guidelines.**

RESPONSE:
We thank the reviewer for this insightful suggestion. As suggested, we have added some information about the inclusion and exclusion criteria in the text as follows (page 8, paragraph 2):

English written guidelines proposed by government agencies (such as National Institute for Health and Care Excellence), or professional academic societies (such as American Psychiatric Association) will be included.

2. **The appraisal tool (or method) used to evaluate the quality of included guidelines should be stated. In addition, it is desirable to have detail items of data extraction described in the protocol.**

RESPONSE:

We appreciate these helpful suggestions. With regard to the appraisal of guidelines, we checked the items of relatively widely accepted guideline appraisal instruments including AGREE II, but found it hard to apply them because of ambiguity and subjectivity of the included items. We therefore decided not to use it but describe the development methodology of each guideline. This information was added in the text (page 8, paragraph 2):

Due to the ambiguity and subjectivity associated with the Appraisal of Guidelines Research and Evaluation II (AGREE II) instrument for quality assessment of guidelines, we will not use such an instrument in the appraisal of guidelines, but we will describe the methodology of each guideline development in supplementary materials.

Regarding the details of data extraction, we think it is a great advice and we have illustrated this issue in the "*Identification and extraction of recommendations in guidelines*" section (page 8, paragraph 3):

Basic information of the guideline (institution, year, targeted patients, diagnostic criteria, etc.), pharmacotherapy recommendations as acute phase treatment (particular drugs and categories recommended and least recommended, severity of the disease, etc.), pharmacotherapy recommendations for patients who have no satisfactory response to initial treatment will be recorded in detail, in order to be compared with results from cNMA and real world.

3. **For real-world prescriptions data, research team propose to search all 24 listed antidepressants in MEPS at time of data extraction. It would be desirable to describe in details about the statistics methods for data analysis.**

RESPONSE:

We appreciate the opportunity to clarify this point. This meta-epidemiological study is basically a descriptive study, which will visually explore the differences between evidences from cumulative network meta-analysis, guideline recommendations and real-world prescriptions. As a result, there would be no formal statistical methods that are going to be used in the study.

4. **Some confounders that may influence the prescription should be of concern, for example, certain antidepressants may be effective in some patients for years, and with agreement of the treating physicians, they might not be willing to change to newer medications. It also may be related to physician' prescription pattern, patient preferences and local regulations, etc. Multiple use of different antidepressants over years is another issue. All these may be a major challenge in data analysis and interpretation.**

RESPONSE:

We agree with the reviewer's comments, and we have acknowledged the potential limitations in the "*Discussion*" section as follows (page 12, last paragraph):

There are some limitations to our study. First, MEPS does not allow us to extract very precise information, including whether the antidepressant is being used as first-line or later treatments. This may lower the comparability between the cNMA of acute phase treatment studies of unipolar depression and the real world practices, as patients on continuation/maintenance treatment may continue using the same antidepressants after guideline recommendations for acute phase treatment change. However, it must be noted that three quarters of patients who initiate antidepressant treatment discontinue the drug within 90 days[54], suggesting that the majority of the patients in MEPS database represent initial prescriptions. (…) Lastly, the reasons behind the differences between evidence and practice may be very complicated and we will need to factor in various potential confounding factors such as the side-effects, marketing efforts of pharmaceutical companies and local regulations such as price or patent status of a particular drug.

We have also discussed the influence of potential confounder "multiple use of different antidepressants". We have therefore decided to include patients on monotherapy only. We have clarified it in the "*Real-world prescriptions data extraction*" section (page 10, last paragraph):

We will use the total number of patients with the diagnosis of depression and who are on antidepressant monotherapy but not taking mood stabilizers and antipsychotics (listed above) concomitantly at this round as denominator, and the prescription for a particular drug as monotherapy as numerator.

Reviewer 2
(Reviewer Name: Julian Elliott
Institution and Country: Cochrane Australia, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia)

The authors propose a retrospective study comparing network meta-analyses (NMA) of anti-depressants at 5 yearly intervals with major clinical practice guidelines (CPGs) and US prescription data. The authors should be commended for the objectives of this meta-epidemiological study, which aims to highlight important gaps between the accumulation of research and subsequent guideline recommendations and practice. The key challenge for this study is the complexity of guideline development and practice behaviours. Efficacy and acceptability are some, but not all the factors that must be considered when developing guideline recommendations. Similarly, there is well characterised complexity in the translation of guideline recommendations into prescription practice. The authors acknowledge these and other limitations, but I am still left with some concerns regarding the conclusions that will be possible when the planned comparisons have been completed.

RESPONSE:

We thank the reviewer for her careful review and the overall broad perspectives provided. We agree with the reviewer that these are very important limitations when interpreting the results of this study, which we have tried to cover in the Methods and Discussion sections of the protocol. Please see our response to the first reviewer above.

First, I believe the plan to repeat the NMA at 5-yearly intervals is problematic. As I understand the aims of the project, the authors plan to compare the results of each NMA with subsequent CPGs, e.g. perform a '2005 NMA' and compare findings to CPG recommendations published between 2005 and 2009. However, for CPGs published towards the end of each 5-year period (e.g. NICE 2009, CANMAT 2009), there was substantial research available that will not be included in the '2005 NMA', further weakening the conclusions that can be made from these comparisons. With the caveat of feasibility, I suggest that to achieve the study aims, the NMAs should be performed at more frequent intervals (optimally for each year).

RESPONSE:

We appreciate the reviewer's suggestions. We agree that 'CPG 2009' (guidelines published in 2009) might take updated evidences not covered in 'NMA 2005' into consideration. Please note however that we plan to run NMA every 5 years and we will use a figure to show the trend of evidence change and prescription change by connecting these 5-yearly estimates. When we make comparisons, we will consider these trends as well.

Second, for consideration only, I wonder about an analysis that focusses not so much on 'gaps' between the NMAs and CPGs, but time from a change in NMA ranking to change in CPG recommendation. This would provide important meta-epidemiological information on delays in the evidence ecosystem over the last 20-30 years.

RESPONSE:

We thank the reviewer for this very important comment. Although the exact lag time between change in NMA rankings and CPG recommendation is very critical information, such an estimate is only possible if there is straightforward association between rankings and the recommendations (e.g., the trend of rankings is approximately the same with the trend of guideline recommendations ('parallel line' in figure), only with delayed trend). However, as the reviewer herself has commented, quite a few factors may affect the prescription patterns other than NMA rankings, so that the direct association may or may not be observed. If the factors other than time contributed to the difference between rankings and recommendations, it would be hard to tell the exact time lag. Consequently, we have decided not to set the lag time as our primary aim, but to investigate this lag more qualitatively, from the view of overall change.

Third, the weaknesses of the MEPS dataset acknowledged by the authors are significant, particularly that the data are not able to be disaggregated by initial versus continuing treatment. This may lead to erroneous conclusions about delays in prescription practice change. I wonder whether the analysis of this dataset is worth the effort, or could at least be supplemented by a more appropriate dataset with data on antidepressant initiations.

RESPONSE:
We appreciate these helpful suggestions. In fact, we agree with the reviewer's point of view, that including maintenance treatment in real-world data may lead to some extra delay of change of prescription from change of recommendations. We have added this statement in the 'discussion' section (page 12, the last paragraph):

First, MEPS does not allow us to extract very precise information, including whether the antidepressant is being used as first-line or later treatments. This may lower the comparability between the cNMA of acute phase treatment studies of unipolar depression and the real world practices, as patients on continuation/maintenance treatment may continue using the same antidepressants after guideline recommendations for acute phase treatment change. However, it must be noted that three quarters of patients who initiate antidepressant treatment discontinue the drug within 90 days[54], suggesting that the majority of the patients in MEPS database represent initial prescriptions.

There indeed are not may datasets available to us that can provide nationally representative prescription data over several decades. For example, we have inquired for the UK's primary care prescription dataset from National Health Service (NHS), which is available to provide antidepressant prescription data in England from 1998 to 2016 (Curtis HJ & Goldacre B (2018) OpenPrescribing: normalised data and software tool to research trends in English NHS primary care prescribing 1998-2016. BMJ Open, 8, e019921). However, as this dataset does not include diagnostic information, antidepressants used for anxiety disorders or for pain would be included as well, which may cause additional problems in interpretation of results, we decided not to include this dataset into the study but to use MEPS database only.

Additional points:

The proposed study is not a NMA alone, but given the important role NMA plays in the study, the authors could consider including NMA in the title of the manuscript.

RESPONSE:
We thank the reviewer for this helpful suggestion. We have revised our title into (page 1, title):

Evidence Synthesis, Practice Guidelines and Real-World Prescriptions of New Generation Antidepressants in the Treatment of Depression: A Protocol for Cumulative Network Meta-analyses and Meta-epidemiological Study

'Extracting' and coding CPG recommendations is a complex task and needs further description.

RESPONSE:
We thank this advice. We have added illustration in the "*Identification and extraction of recommendations in guidelines*" section (page 8, paragraph 3):

Basic information of the guideline (institution, year, targeted patients, diagnostic criteria, etc.), pharmacotherapy recommendations as acute phase treatment (particular drugs and categories recommended and least recommended, severity of the disease, etc.), pharmacotherapy recommendations for patients who have no satisfactory response to initial treatment will be recorded in detail, in order to be compared with results from cNMA and real world.

The definition of trial completion year seems to suggest that for studies with a known completion year after the proposed 'NMA year', data published after the 'NMA year' will be included in the relevant NMA. Please clarify.

RESPONSE:
The reviewer's interpretation is correct: studies with a known completion year before the proposed 'NMA year' will be included in the relevant NMA even if the data were published after the 'NMA year'. This was because we had intended the cumulative NMA to present the best available evidence as of the proposed 5-year intervals. A counter-example would have been to include only the published studies; however, we reasoned that it would make no sense to compare the guidelines with the publication-biased results. We explained this in the Methods section as follows (page 7, paragraph 3):

Thus studies with a known completion year before the proposed cNMA will be included in the relevant cNMA even if the data were published after the that year. Only such cNMA can present the best available, publication bias-free evidence, upon which practice guidelines should ideally be founded.

Please clarify the text to indicate that individuals taking mood stabilisers and antipsychotics will be excluded from denominator data.

RESPONSE:

We appreciate the reviewer's careful review of our protocol. We have added statement concerning the denominator (page 10, the last paragraph):

We will use the total number of patients with the diagnosis of depression and who are on antidepressant monotherapy but not taking mood stabilizers and antipsychotics (listed above) concomitantly at this round as denominator, and the prescription for a particular drug as monotherapy as numerator.

Please clarify whether the MEPS data will be compared to APA recommendations only.

RESPONSE:

We thank this nice suggestion. As suggested, we have added a sentence in the '*Comparison between cNMA, CPGs, and real-world prescriptions*' section (page 10):

Moreover, as MEPS is a database from the US, we will attempt to compare results from MEPS with the US guidelines APA.

For the statement regarding systematic reviews being out of date at time of publication the authors should reference Shojania, Ann Int Med 2007, which estimated that 7% of SRs were out of date when published.

RESPONSE:

We appreciate the reviewer's kind and helpful suggestion, we have reviewed this article and added in the text as a reference (reference 16) (page 5, 3rd paragraph):

Second, it has been reported that there is approximately eight months to more than one year's lag between the last search date in the literature and the publication of an SR, [13 14] and in total 2.5-6.5 years' interval between the publication dates of the latest primary studies and the publication date of SRs.[15] It was estimated that 7% of the published SRs were already out of date at the time of publication[16].

A better citation is needed for the MEPS dataset.

RESPONSE:

We thank the reviewer's careful review. We have updated the reference information of MEPS dataset.

The potential contribution of this work to living guidelines could reference recent publications on this topic.

RESPONSE:

We thank the reviewer for this helpful suggestion. We have updated the description of clinical implications and the references we cited (page 12, '*Dicussion*' section):

For instance, the cNMA used in this study may eventually evolve into prospectively designed sequential network meta-analysis which can be used to continuously update evidence and contribute to living guidelines, which is a guideline that is updated as soon as new evidence becomes available, so that it can provide timely and trustworthy suggestions for decision-makers.

We cited 3 references:

Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. PLoS Med 2014;11(2):e1001603. doi: 10.1371/journal.pmed.1001603 [published Online First: 2014/02/22].

Elliott JH, Synnot A, Turner T, et al. Living systematic review: 1. Introduction-the why, what, when, and how. J Clin Epidemiol 2017;91:23-30. doi: 10.1016/j.jclinepi.2017.08.010 [published Online First: 2017/09/16].

Akl EA, Meerpohl JJ, Elliott J, et al. Living systematic reviews: 4. Living guideline recommendations. J Clin Epidemiol 2017;91:47-53. doi: 10.1016/j.jclinepi.2017.08.009 [published Online First: 2017/09/16].

Please describe the role of the funder in the study.

RESPONSE:

We have added a sentence in the '*Funding*' section (page 13):

The funder has no role in study design, data collection, data analysis, data interpretation, writing of the report, or in the decision to submit for publication.

Here are the comments from the editor:

Please revise your title so that it includes your study design. This is the preferred format for the journal.

RESPONSE:

We thank the reviewer for this helpful suggestion. We have revised our title into (page 1, title):

Evidence Synthesis, Practice Guidelines and Real-World Prescriptions of New Generation Antidepressants in the Treatment of Depression: A Protocol for Cumulative Network Meta-analyses and Meta-epidemiological Study

Authors must include a statement in the methods section of the manuscript under the sub-heading 'Patient and Public Involvement'.

RESPONSE:

We are thankful to this suggestion, which is becoming more and more important recent years. We have added the '*Patient and Public Involvement*' section in the '*Method*' section (page 12):

This research protocol was written without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not expected to contribute to conducting this study and to the writing of this document for readability or accuracy.

VERSION 2 – REVIEW

REVIEWER	Joyce Kee-Hsin Chen College of Nursing, Taipei Medical University, Taiwan
REVIEW RETURNED	23-Sep-2018

GENERAL COMMENTS	This study comparisons available evidence using cNMA, recommendations from guidelines, and real-world clinical practices data from MEPS database. Although there are some considerations still exist, as authors mentioned. For example, the information from MEPS are not very precise, physicians' prescription behaviors, patients' compliance, pharmaceutical companies' marketing efforts, and local regulations. Those potential confounding factors may influences final results. It should be discussed in depth after the end of the study. However, at this stage, this study Protocol still has value for understanding the gaps between evidence and the real world practice. In addition, the study methodology may provide valuable insights into future EBM studies. It's prepared to publish.