

## 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

<b>TITLE (PROVISIONAL)</b>	Characteristics of stakeholder involvement in systematic and rapid reviews: a methodological review in the area of health services research
<b>AUTHORS</b>	Feldmann, Jonas; Puhan, Milo; Mütsch, Margot

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Irina Kinchin Central Queensland University, Australia
<b>REVIEW RETURNED</b>	06-Jul-2018

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to complete a review for the manuscript ID bmjopen-2018-024587, entitled "Characteristics of stakeholder involvement in systematic and rapid reviews: a cross-sectional study in the area of health services research." The study assessed stakeholder involvement in systematic and rapid reviews. The authors focused on a best-case approach, where stakeholder involvement might be most regularly reported. The study arrived to justly conclusions that future researcher shall involve stakeholders more broadly in the process of evidence synthesis. Furthermore, stakeholder involvement should be included and evaluated in reporting checklists of all review types. I suggest the following (minor) revisions:</p> <p><b>Abstract:</b> Please confirm the number of searched databases, i.e. 4 or 3. Whether "Cochrane databases of systematic reviews and of other reviews" are indeed two distinct databases, and not the one called Cochrane Library? Methods section says the Cochrane Database (counting it as one), whereas the abstract says Cochrane databases (counting it as two).</p> <p><b>Methods:</b></p> <ul style="list-style-type: none"><li>• Please provide justification for excluding studies that focused on economic analyses or performed a narrative review.</li><li>• I noted that all included papers were funded research (as per Table 1). This made me think that papers might have been included/excluded based on the funding source. If that was one of the criteria, please explicitly say so.</li><li>• Specify keywords searched for "health service research" and "Health Technology Assessment" in Ovid Medline, Embase and the Cochrane Libarary. I suggest adding a supplementary file with an example of the Medline search to allow the study to be repeated.</li></ul>
-------------------------	---

	<ul style="list-style-type: none"> <li>• It was not clear, how the randomisation had happened? Please provide details.</li> <li>• I suggest adding description of how stakeholder involvement was ascertained. Consider providing types of stakeholder involvement in the methods section.</li> <li>• The protocol refers to an Excel file (page 6/14 in the protocol) with a summary table for all included studies. If this file is available, make it as a supplementary.</li> </ul> <p>Results:</p> <ul style="list-style-type: none"> <li>• “Where available, our rating was compared with an already existing one from the “health evidence” or “health systems evidence” platforms” – In the results, please comment on discrepancies, if any.</li> <li>• The quality of included studies was assessed using the AMSTAR tool. However, I could not find description of the rating in the results, except for listing the rating in Table 1. Have stakeholder involvement had any association with the quality of reviews in this study? Otherwise, the purpose of the quality scoring in this paper was unclear.</li> </ul>
--	--

<b>REVIEWER</b>	Lisa Hartling University of Alberta, Canada
<b>REVIEW RETURNED</b>	09-Aug-2018

<b>GENERAL COMMENTS</b>	<p>Abstract:</p> <ul style="list-style-type: none"> <li>• Objective: <ul style="list-style-type: none"> <li>○ Need to define ‘stakeholders’</li> <li>○ Add a statement about why their involvement might be important</li> <li>○ Did the authors intend to compare across the four groups of review types that they identified? If so, this should be specified. I think there are problems with how these review types have been grouped (see comments below about what authors have labelled as rapid reviews of the Cochrane Database). Did authors select these groupings in order to compare Cochrane and non-Cochrane products, and/or full systematic reviews vs. rapid reviews. The objectives need to be clarified.</li> </ul> </li> <li>• Search strategy <ul style="list-style-type: none"> <li>○ “the Cochrane databases of systematic reviews and of other reviews” – I don’t think this is the proper name – please clarify.</li> <li>○ “rapid reviews of the Cochrane Database” – to my knowledge Cochrane does not produce and publish rapid reviews; I believe these rapid reviews came from the HTA Database that has been housed in the Cochrane Library. I would clarify this as I think it is misleading to imply that these are produced within Cochrane.</li> </ul> </li> <li>• Results: <ul style="list-style-type: none"> <li>○ First sentence, did authors screen all 57,822 records to classify them, or were they classified through some other means (e.g., from the source where they were identified)? Was there a mechanism to identify and remove duplicates?</li> </ul> </li> </ul>
-------------------------	---

	<ul style="list-style-type: none"> <li>○ Second sentence, missing a word – “were based on a study protocol” (or perhaps rephrase to: reported an a priori study protocol...</li> <li>○ Second and third sentences (study protocol and conflict of interest) – these elements appear to only be reported for two of the groups of reviews – what about the other two groups?</li> <li>○ Authors assessed reviews using AMSTAR – these results should be reported. How were these used? Did they assess whether AMSTAR correlated somehow with stakeholder involvement?</li> <li>● Conclusions: <ul style="list-style-type: none"> <li>○ Most of the statements in the conclusions are actually results, i.e., information about stakeholder involvement and comparison of this across review types.</li> <li>○ Last sentence, authors suggest that reporting needs to be improved. What reporting are they referring to? Reporting of stakeholder involvement, or general reporting of review methods? How did they determine this – based on what extracted elements?</li> </ul> </li> </ul> <p>Page 3, Strengths and limitations of this study:</p> <ul style="list-style-type: none"> <li>● First point: coproduction vs. stakeholder engagement – these concepts need to be defined. Abstract only mentions engagement with no mention of coproduction.</li> <li>● Second point mentions reporting – what reporting are they referring to (general vs. specific to stakeholder engagement), and how was this assessed? What about quality as assessed by AMSTAR?</li> </ul> <p>Introduction:</p> <ul style="list-style-type: none"> <li>● This could be more succinct; some of the material seems extraneous to the objectives of the study (e.g., quality of SR reporting; variability of methods and terms for rapid reviews).</li> <li>● Page 4, first paragraph: mentions many types of evidence syntheses, however, the study focuses only on SRs and RRs – why this choice? There is some rationale on page 6, i.e., they ‘represent different types of evidence synthesis’; however, why are the other types of evidence synthesis not represented?</li> <li>● Page 4, second paragraph, I would use a more recent reference (not Jadad 1998) to describe Cochrane SRs.</li> <li>● Page 5, lines 19-23: this information is inaccurate and poorly described. These are other databases (DARE and HTA) within the Cochrane Library. They are distinct from the CDSR (Cochrane Database of Systematic Reviews) that house only Cochrane reviews. It’s important that this is clear as the groupings of reviews (and comparisons of interest) in this study is dependent on the accuracy of this information.</li> </ul>
--	---

	<ul style="list-style-type: none"> <li>• Page 5, lines 35-37: I would specify who is typically included as stakeholders.</li> <li>• Page 5, lines 43-47: need to define “health services research”</li> <li>• Pages 5-6: Study aim needs editing; some sentences are long and unclear. First mention of “coproduction” at end of “study aim” – this needs to be defined, and how this differs from stakeholder involvement (or perhaps define stakeholder involvement and the range of activities).</li> </ul> <p>Methods:</p> <ul style="list-style-type: none"> <li>• Under study design, authors state their purpose: “to assess characteristics and reporting of coproduction” – again, need to define coproduction vs. stakeholder involvement</li> <li>• Under search strategy: for non-Cochrane SRs, did you ensure that the citations identified in Medline and Embase were not also published as Cochrane SRs (this wouldn’t necessarily appear in the title or abstract). Did the authors verify that the records identified were actually “systematic reviews”? There are many citations that come up in a search for “systematic reviews” that aren’t necessarily SRs and/or don’t meet specific definitions of SRs. What definition did the authors use? Clarity needed about “Cochrane databases” and databases contained within the Cochrane Library. The authors searched for HTAs – did they consider these to be the same as rapid reviews? What was their definition of an HTA? Again, did they apply a definition or rely on the description in the primary report? How was full text screening performed and how was consensus achieved? How was ‘spot-checking’ done (random sample of reviews, sample of data elements, other)?</li> <li>• I’m not familiar with sample size calculations for a cross-sectional study (unless comparisons are planned). It does seem a bit concerning that such a small sample of full SRs were included (30 out of 56,986) compared with 30 out of 95 for “CD RRs”. It seems that such a small sample of the full SRs would not likely be generalizable across the entire population of SRs. Based on the sample size description it appears that the authors wanted to conduct comparative analyses and subgroup analyses; this should be reflected in the study aims (i.e., to compare across review types and between Cochrane and non-Cochrane – I think these are two comparisons of interest but authors chose to compare across the four groups).</li> <li>• Inclusion and exclusion criteria: need definitions of terms: SRs, HTAs, RRs. Authors indicate that “HTAs targeting effectiveness as well as meta-analyses and systematic reviews for effectiveness and utility, pharmaceutical trials under everyday conditions, basic research in care related fields...were included”. This implies that authors included primary research as well as reviews. This needs to be</li> </ul>
--	--

reworded and clarified. Last sentence on types of studies excluded also needs to be clarified, e.g., efficacy studies without assessing QOL – does that imply that efficacy studies (i.e., primary studies) that assessed QOL were included? That doesn't make sense given the focus on reviews.

- Data extraction: How was AMSTAR assessed - in duplicate, process for consensus? Authors should provide questions that were sent to contact authors of included articles. Did they also ask them about the type of stakeholders, or just the number of stakeholders?

Results:

- First sentence should read 89,798 'citations' – we can't confirm that these were reviews, until they are more closely examined. Why were so many Cochrane SRs excluded following title and abstract screening? It would seem that these would all meet the criteria for a full SR.
- Second paragraph: information about increase and annual number is not relevant to the study objectives. Also without screening them all to confirm, I am not confident that they all meet the definition of a given category of review type (e.g., SR, RR, etc). I think this paragraph should be removed.
- Table 1: I would like to see the breakdown of the AMSTAR assessments by question and study (at least in an appendix).
- Text on page 13: This could be reduced as it is not the primary objective to compare characteristics across the review types.
- Page 15, first paragraph: Authors comment on SI across years, although the numbers are so small per year and review type that this comparison does not seem valid. I would suggest that paragraph start with the results on overall SI then get into more specific details.

Discussion:

- Discussion generally needs work, and should align with study objectives.
- First sentence indicates that SI varied by review type although this comparison was not made explicit as an objective of the study.
- Third sentence indicates there was a 'significant difference' in the phases of SI; however, there were no tests of statistical significance reported. The authors need to re-word this and be careful with wording and any implications of 'significant differences'. However, the authors did report on sample size calculations, presumably with the intent to test whether there were differences in SI. Also they comment on planned subgroup analyses but none are reported, only descriptive/summary statistics.

	<ul style="list-style-type: none"> <li>• Page 19, first paragraph, the following statement needs to be clarified: “Of note, the majority of RR and SR authors reported that the involved stakeholders had positive and considerable effects on the study results, making the review more relevant for the targeted audience.” In Table 4 only 33% reported a positive effect and 67% did not mention an effect. Further, I don’t see in the results any data on relevance of the review for the target audience – was this information collected or implied by the authors? The authors also indicate that their results are “in agreement with the earlier described findings of Keown et al.[143] who concluded, that stakeholder involvement led to increased relevance and depth of review findings, more clarity in defining research questions, broader dissemination of their results and increased awareness of target groups”; however, I don’t see this level of detail collected or reported in the present study.</li> <li>• Page 19, last paragraph: I don’t think this is relevant to the present study, i.e., discussion of quality and characteristics across review types. Page 20, first paragraph: this also was not related to the aim of the study.</li> <li>• Authors should comment on the purpose and nature of RRs and SRs; they are conducted for different purposes so the different levels of SI are not necessarily surprising. Given their different purpose, is it reasonable to have the same level of SI for the different review types? This type of discussion would be valuable.</li> <li>• Data mentioned in the discussion around response rates from review authors should be presented more fully in the results.</li> <li>• The following text from page 21 is concerning: “In their emails, some authors mentioned experts or peer reviewers and listed them as stakeholders. One author listed members of the research team, who performed literature research and data extraction as stakeholders. Given all the information about who was involved in the reviews, it was sometimes difficult to conclude whom we would count as stakeholder and whom as “expert” from the same, or from a related field.” Given the varied (and confusing) responses regarding SI, it seems the emails or questions therein were not clear about what the authors were asking. I believe Cochrane recommends that the review team be comprised of clinical experts and those with technical expertise to carry out the review; this author group is different than stakeholders, and should have been clarified upfront. Likewise, I wouldn’t consider the peer-review process as stakeholder involvement in the review (or “coproduction”).</li> <li>• Discussion should include a section on “strengths and limitations” of the study, and possibly “future research needs”. A “conclusions” section would be helpful as well – what are the key messages, and new findings this study provides?</li> </ul>
--	---

	<p>Supplementary File S1 (study protocol):</p> <ul style="list-style-type: none"> <li>• The objective stated in the study protocol is different than the manuscript.</li> <li>• The protocol lists primary and secondary outcomes - I think these are actually primary and secondary objectives or aims; however, it would be beneficial to have these in the manuscript and to present the results according to these primary and secondary points.</li> <li>• The protocol indicates: “We will compare the proportion of stakeholders involved between groups using the non-parametric Wilcoxon rank sumtest.” This was not described in the methods of the manuscript, and was not reported on. The protocol indicates that differences between protocol and review (i.e., study) will be listed if applicable; however, no differences and reasons are listed in the manuscript.</li> </ul> <p>Page 45: the authors list their search strategy. I would have recommended using an established search filter for SRs.</p> <p>Page 50: Not all STROBE checklist items are accurately reported, e.g., item 9 bias, authors report using AMSTAR. This is a tool to assess the methodological quality of SRs. It does not assess potential bias in the conduct of the present study. Item 21, generalizability: I don't really see a discussion about generalizability and I think this is a major concern given the small sample, particular with respect to the huge number of SRs identified.</p> <p>Page 52: there are a number of citations in the excluded list that have “systematic review” in the title; it's unclear why these were excluded.</p>
--	--

<b>REVIEWER</b>	Adrian Barnett Queensland University of Technology Australia
<b>REVIEW RETURNED</b>	08-Nov-2018

<b>GENERAL COMMENTS</b>	<p>This study concerns the important topic of engaging end users in research. The ideas and design were clearly explained. The methods were appropriate and the results were useful and well presented. I have only minor comments.</p> <p>There was no mention of the James Lind Alliance who do excellent work in ensuring that end-users are involved in research. They also have a good idea about the extra time and resources needed, and the benefits.</p> <p>Would it be worth mentioning the recent controversy around the Cochrane systematic review of myalgic encephalomyelitis? This would not get into the issues of the complaints, but it is a very topical example of how end users involvement (or lack of) in systematic reviews is a topical issue.</p>
-------------------------	---

	<p>Better engagement of stakeholders might reduce research waste, and systematic reviews are a great source of waste, see for example DOI: 10.1111/1468-0009.12210.</p> <p>Minor comments</p> <ul style="list-style-type: none"> <li>- the use of SR and RR in the abstract made it hard to read. I would also avoid these acronyms in the main text.</li> <li>- page 4, "concluded that the reporting of Cochrane SRs was the most complete one", compared with what? Also repeated in the next sentence.</li> <li>- the authors use both 'stakeholders' and 'end-users', would it be better to use just one, and/or explain the difference between the two?</li> <li>- page 7, lines 42 to 53, this is a long passive sentence. I suggest moving 'were included' towards the start of the sentence. The last sentence in this section is also a relatively long passive sentence.</li> <li>- page 8, line 38 "checklist was reporting", should be 'reported' or 'recorded'. I would also start a new paragraph at the end of this sentence.</li> <li>- Table 1 footnote, might be worth mentioning that higher AMSTAR scores indicate higher quality</li> <li>- Table 2, decimal places are not needed in the population size results</li> <li>- page 19, "We, therefore, suggest that the contribution of involved stakeholders in RRs seemed to be more substantial than in SRs in general and still needs improvement" This sentence is not clear, does it mean improvements are needed for systematic reviews?</li> <li>- page 21, "One author listed members of the research team, who performed literature research and data extraction as stakeholders." shouldn't this be 'erroneously mentioned'?</li> </ul>
--	---

<b>REVIEWER</b>	Agnes Dechartres Sorbonne University, Paris, France
<b>REVIEW RETURNED</b>	21-Nov-2018

<b>GENERAL COMMENTS</b>	<p>Review for BMJ open, manuscript entitled « Characteristics of stakeholder involvement in systematic and rapid reviews : a cross sectional study in the era of health services research »</p> <p>This manuscript reports the results of a methodological review focusing on a random sample of systematic and rapid reviews (including those conducted by the Cochrane Collaboration) that were published between Jan 2011 and Oct 2015. The total number of included reviews is 120 : 30 for each of the following categories : Cochrane systematic reviews, other systematic reviews, Cochrane rapid reviews, other rapid reviews. The authors found stakeholder involvement in 13% of Cochrane reviews, 20% of other systematic reviews, 40% of rapid reviews and 93% of Cochrane rapid reviews.</p> <p>The topic is interesting but I found some parts of the manuscript long and I think it could be more straightforward as it is difficult to see what is the most important to retain and what this study adds. Here are my detailed comments :</p> <ul style="list-style-type: none"> <li>- Some wording frequently used are not clear <ul style="list-style-type: none"> <li>o General and reporting characteristics. What are 'reporting characteristics' ? general characteristics can also be considered as reported characteristics.</li> <li>o Coproduction</li> </ul> </li> </ul>
-------------------------	---

	<p>For example : in the study aim section : « and to specifically point out reporting and study characteristics as well as measures and identified challenges of coproduction ».</p> <p>What is the difference between reporting and study characteristics ? What do the authors mean by challenges of coproduction ?</p> <ul style="list-style-type: none"> <li>- The introduction is much too long.</li> <li>- The authors conducted a specific search for each of the 4 categories (Cochrane systematic reviews, other systematic reviews, Cochrane rapid reviews, other rapid reviews). They give the results of the search as if all records retrieved corresponded to what they searched, which is not the case. Therefore, the authors cannot say « with 57,822 articles remaining after exclusion of duplicates... 56,986 were categorized as SR. Not all the records retrieved by the search actually correspond to a systematic review. The authors cannot base parts of their results page 9 lines 5 to 38 on that.</li> <li>- The search is not really recent : date of last search is October 2015 so 3 years ago. The sample of reviews used can be outdated, which limits the generalisability of findings in the current context.</li> <li>- The sample size calculation does not correspond to what the authors did. In sample size calculation, the authors planned to compare rapid reviews and systematic reviews but they constructed the study and presented the results for the 4 categories.</li> <li>- The concept of health services research as defined by the authors is very wide and the interest of involving stakeholders may vary depending on the topic.</li> <li>- This methodological review aims to evaluate stakeholder involvement in reviews but they did not involve stakeholder involvement themselves, which would have been of interest.</li> <li>- The first part of the results including description of characteristics of the reviews is very long and the authors could come more quickly to the main results of this study.</li> <li>- The authors probably did not go sufficiently in depth on the interest of involving stakeholders.</li> <li>- I think it would be useful to present some key examples in an appendix</li> <li>- The authors should also be more careful on their interpretation of results in the discussion. For example : « the higher median population as well as the less specific age range or study setting in non Cochrane reviews illustrates that the studies they included had a broader spectrum of possibly included patients ».</li> </ul> <p>Minor comments</p> <ul style="list-style-type: none"> <li>- Abstract : the objective could be more clearly formulated</li> <li>- Abstract : I would not present the number of variables targeting stakeholder involvement as it is not really important and it is difficult in the methods of the manuscript to know what are these 9 variables</li> <li>- The sentence « 33% of the responding contact authors mentioned unspecified positive effect of stakeholder involvement » is unclear</li> <li>- Abstract conclusion should be in line with results presented in abstract</li> <li>- The study aim mentions challenges of stakeholder involvement but it is not what they evaluate.</li> </ul>
--	---

	<ul style="list-style-type: none"> <li>- The Strobe statement is not really appropriate. This study is more a methodological review so Prisma could be more appropriate</li> <li>- In data extraction, the effect on the review outcome is not really clear</li> <li>- In discussion, the authors should avoid 'significantly' as they did not conduct any formal comparison</li> <li>- In Discussion, Prisma is not a guideline aiming to improve methodological quality but to improve reporting</li> </ul>
--	---

## VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Irina Kinchin

Institution and Country: Central Queensland University, Australia

Thank you for the opportunity to complete a review for the manuscript ID bmjopen-2018-024587, entitled "Characteristics of stakeholder involvement in systematic and rapid reviews: a cross-sectional study in the area of health services research." The study assessed stakeholder involvement in systematic and rapid reviews. The authors focused on a best-case approach, where stakeholder involvement might be most regularly reported. The study arrived to justly conclusions that future researcher shall involve stakeholders more broadly in the process of evidence synthesis. Furthermore, stakeholder involvement should be included and evaluated in reporting checklists of all review types. I suggest the following (minor) revisions:

Abstract:

Please confirm the number of searched databases, i.e. 4 or 3. Whether "Cochrane databases of systematic reviews and of other reviews" are indeed two distinct databases, and not the one called Cochrane Library? Methods section says the Cochrane Database (counting it as one), whereas the abstract says Cochrane databases (counting it as two).

Response: We agree and thank you for pointing to this disagreement.

It is corrected in the manuscript's abstract and method section (search strategy and screening) to four databases including the denominations. The Cochrane database of systematic reviews and the databases of the University of York, Centre for Reviews and Dissemination ([www.crd.york.ac.uk/CRDWeb](http://www.crd.york.ac.uk/CRDWeb)) were included separately. This change happened when we already had completed our search.

Methods:

- Please provide justification for excluding studies that focused on economic analyses or performed a narrative review.

Response: We aimed to perform a comparison of systematic reviews and as the approaches are different for narrative reviews and economic analyses, we excluded them as not to mix so different methodologies.

We have added a sentence to justify this in the methods section (inclusion and exclusion criteria): “We excluded studies that focused exclusively on economic or cost analyses, performed a narrative review, an overview of reviews or a protocol of a systematic review in order not to include so different methodological approaches.”

- I noted that all included papers were funded research (as per Table 1). This made me think that papers might have been included/excluded based on the funding source. If that was one of the criteria, please explicitly say so.

Response: We did not use the funding neither the funding source as an inclusion or exclusion criteria. However, we were interested in to assess this variable to explore whether funding or the funding source were different between the types of reviews included and with respect to the involvement of stakeholders (no evaluation possible due to the low numbers of SI).

- Specify keywords searched for “health service research” and “Health Technology Assessment” in Ovid Medline, Embase and the Cochrane Library. I suggest adding a supplementary file with an example of the Medline search to allow the study to be repeated.

Response: Thank you to raise this important point. The search strategies did only include the review type. We did not search for “health services research” or “Health Technology Assessment” but included it in the screening steps. To assign a study to the “health services research” field we used the list of inclusion and exclusion criteria as provided in the Methods section, inclusion and exclusion criteria. This screening was done for titles and abstracts and included the verification of the review type. In the course of the full text screening the appropriateness of the classification as “health services research” was rechecked and re-assessed where necessary.

- It was not clear, how the randomisation had happened? Please provide details.

Response: The following sentences were added to the method section, study design:

“For each of the four study types, each reference was given a unique random number using the sample() function available in R [1]. Each list was then sorted by the random number, and the articles were screened in order.” Reference: R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- I suggest adding description of how stakeholder involvement was ascertained. Consider providing types of stakeholder involvement in the methods section.

Response: Stakeholder involvement was assessed in the article if it was mentioned as such or if the engagement of one or multiple representatives of stakeholder groups were described. This included patients/consumers, professional organisations, health care providers, hospital representatives and/or community services, researchers (not members of the primary research team, see below), policymakers at Federal, State and local levels.

To ascertain stakeholder involvement and to assess missing information we sent an email to all contact persons of the included articles (and one reminder to non-respondents).

The term “stakeholder” was described in the method section (data extraction): “Stakeholders were defined as people with a particular interest in the research topic (but are not members of the primary research team). “

Stages of stakeholder engagement in the production of the corresponding review were either unspecified or included one or multiple of the following steps: the formulation of the research question(s), determining study characteristics, contributing to the writing of the protocol, participating within the review process with searching, screening, data extraction, synthesizing, interpreting the study results and/or establishing recommendations.

We added potential stakeholders to be involved, added a sentence about the ascertainment of stakeholder involvement as well as about the stages of stakeholder involvement in the method section (data extraction).

- The protocol refers to an Excel file (page 6/14 in the protocol) with a summary table for all included studies. If this file is available, make it as a supplementary.

Response: All included studies were cited: refs 22 to 141. There is the full data extraction table from which we have included AMSTAR data as Supplementary file S4 and study characteristics as Supplementary File S5. There is no summary table for all included studies.

Results:

- “Where available, our rating was compared with an already existing one from the “health evidence” (he) or “health systems evidence” (hse) platforms” – In the results, please comment on discrepancies, if any.

Response: AMSTAR ratings of the “he or hse platforms” were available only for a minority of reviews (19 studies). Ratings were similar for 68.4% (13/19) of studies with respect to the rating as strong, moderate or weak, respectively. We did not perform any validation studies with the teams of the hse or he-platforms. Therefore, and keeping in mind the low number of comparative data, we did not perform any detailed analyses.

Our initial sentence might be prone to misunderstandings. Therefore, we adapted the wording and made the following addition in the results section as requested:

“Where available, we also collected existing AMSTAR ratings from the “health evidence” or “health systems evidence” platforms. For 68% (13/19 study ratings) we found the same classification as strong ( $\geq 8$ ), moderate (4-7) or weak ( $< 4$ ), respectively.”

- The quality of included studies was assessed using the AMSTAR tool. However, I could not find description of the rating in the results, except for listing the rating in Table 1. Has stakeholder involvement had any association with the quality of reviews in this study? Otherwise, the purpose of the quality scoring in this paper was unclear.

Response: We have provided the AMSTAR ratings as Supplementary file S4 and have added a sentence to the manuscript (see question above). Due to the low numbers of stakeholder involvement we could not assess an association with the quality of reviews (with the exception of CRD rapid

reviews). Furthermore, there is no rating instrument for rapid reviews yet and the AMSTAR tool is not specifically designed for rapid reviews as some questions are not applicable for rapid reviews.

Reviewer: 2

Reviewer Name: Lisa Hartling

Institution and Country: University of Alberta, Canada

Abstract:

- Objective<sup>[SEP]</sup>

- o Need to define 'stakeholders'<sup>[SEP]</sup>

Response: Stakeholders were defined as “people with a particular interest in the research topic (but are not members of the primary research team)”.

This definition was provided in the method section, sub-section “data extraction”. It was also included in the abstract.

- o Add a statement about why their involvement might be important<sup>[SEP]</sup>

Response: The following statement was added to the abstract: “Engaging stakeholders in reviews is considered to generate more relevant evidence and to facilitate reviews’ dissemination and use.”

- o Did the authors intend to compare across the four groups of review types that they identified? If so, this should be specified. I think there are problems with how these review types have been grouped (see comments below about what authors have labelled as rapid reviews of the Cochrane Database). Did authors select these groupings in order to compare Cochrane and non-Cochrane products, and/or full systematic reviews vs. rapid reviews. The objectives need to be clarified.

Response: As described in the method section, first, amount and characteristics of stakeholder involvement regarding full systematic reviews and rapid review products were assessed and second, based upon, full systematic reviews and rapid reviews were compared to each other. This was described in the method section, sample size calculation. Considering differences between Cochrane and non-Cochrane systematic reviews and between rapid reviews and CRD rapid reviews and therefore, to include the full potential (methodological) spectrum of rapid and systematic reviews we included all four types as described. In practice, the reported extent of stakeholder engagement was too low to make statistical comparisons feasible and meaningful.

- Search strategy<sup>[SEP]</sup>

- o “the Cochrane databases of systematic reviews and of other reviews” – I don’t think this is the proper name – please clarify<sup>[SEP]</sup>

- o “rapid reviews of the Cochrane Database” – to my knowledge Cochrane does not produce and publish rapid reviews; I believe these rapid reviews came from the HTA Database that has been housed in the Cochrane Library. I would clarify this as I think it is misleading to imply that these are produced within Cochrane.

Response: Thank you very much, this is an important point and the situation has changed in the meantime. It was included and corrected in the abstract and the method section. The DARE and the

HTA databases were transferred to the University of York, Centre for Reviews and Dissemination and can be accessed there ([www.crd.york.ac.uk/CRDWeb](http://www.crd.york.ac.uk/CRDWeb)).

- Results:<sup>[1]</sup><sub>[SEP]</sub>

- o First sentence, did authors screen all 57,822 records to classify them, or, were they classified through some other means (e.g., from the source where they were identified)? Was there a mechanism to identify and remove duplicates?<sup>[1]</sup><sub>[SEP]</sub>

Response: Through the search strategy, the four review types were defined. We then combined all identified references per review type and identified duplicates by standard measures (references compared by authors, year, journal in endnote).

This was described in the method section. We then performed the screening steps for each review type separately. Therewith, the review type was checked again and the inclusion and exclusion criteria were screened.

The following wording is used in the abstract:

“Four types of reviews (Cochrane and non-Cochrane systematic reviews, rapid and CRD rapid reviews were retrieved between January, 2011 and October, 2015, pooled by review type and duplicates excluded.”

- o Second sentence, missing a word – “were based on a study protocol” (or perhaps rephrase to: reported an a priori study protocol...<sup>[1]</sup><sub>[SEP]</sub>

Response: We have rephrased the sentence according to your suggestion.

- o Second and third sentences (study protocol and conflict of interest) – these elements appear to only be reported for two of the groups of reviews – what about the other two groups?<sup>[1]</sup><sub>[SEP]</sub>

Response: Due to word count restrictions, we have eliminated the sentence about the study protocol from the abstract. The sentence on potential conflicts of interests was rephrased:

“A conflict of interest statement remained unmentioned in 40% (12/30) of non-Cochrane systematic reviews and in 27% (8/30) of rapid reviews, but not in Cochrane or CRD reviews.”

- o Authors assessed reviews using AMSTAR – these results should be reported. How were these used?

Did they assess whether AMSTAR correlated somehow with stakeholder involvement?

Response: We intended to do so. Due to the low amount of stakeholder involvement this was not meaningful and feasible.

- Conclusions:

- o Most of the statements in the conclusions are actually results, i.e., information about stakeholder involvement and comparison of this across review types.

Response: The Conclusion section was rephrased:

“Stakeholder engagement was not general practice in either review type (except for CRD rapid reviews), although it was more common in rapid reviews. Especially for systematic reviews, there is still a considerable potential to improve reporting e.g. by including an a priori study protocol and a conflict of interest statement or to engage stakeholders.”

o Last sentence, authors suggest that reporting needs to be improved. What reporting are they referring to? Reporting of stakeholder involvement, or general reporting of review methods? How did they determine this – based on what extracted elements?

Response: It was specified, see rephrased sentence above.

Page 3, Strengths and limitations of this study:

- First point: coproduction vs. stakeholder engagement – these concepts need to be defined. Abstract only mentions engagement with no mention of coproduction. [SEP]

Response: There are no unambiguous definitions for these terms. They both stand for the concept of establishing interaction, such as relationships and collaborative processes with stakeholders to improve the mutual understanding, enhance relevance, dissemination and use of reviews.

We have deleted the term “coproduction” and use stakeholder involvement or engagement.

- Second point mentions reporting – what reporting are they referring to (general vs. specific to stakeholder engagement), and how was this assessed? What about quality as assessed by AMSTAR? [SEP]

Response: This sentence was re-formulated:

- “Different review types were able to inform on stakeholder involvement and pointed to potential improvements.”

Introduction: [SEP]

- This could be more succinct; some of the material seems extraneous to the objectives of the study (e.g., quality of SR reporting; variability of methods and terms for rapid reviews). [SEP]

Response: The Introduction section was revised and shortened.

- Page 4, first paragraph: mentions many types of evidence syntheses, however, the study focuses only on SRs and RRs – why this choice? There is some rationale on page 6, i.e., they ‘represent different types of evidence synthesis’; however, why are the other types of evidence synthesis not represented?

Response: The following sentence was added to the Introduction section: “We will focus on full systematic reviews as well established review type in healthcare and on rapid reviews as an emerging type.”

It was often suggested that rapid reviews might replace systematic reviews to some extent as they would be a more convenient format for stakeholders. Therefore, we were interested in exploring the situation regarding stakeholder involvement. However, we did not intent to investigate reasons for potentially different stakeholder involvement fostered by review type as this would have asked for a different study design.

- Page 4, second paragraph, I would use a more recent reference (not Jadad 1998) to describe Cochrane SRs. [SEP]

- Response: The intention of this paragraph was to summarize issues that make Cochrane systematic reviews different from non-Cochrane systematic reviews. We have deleted the reference of Jadad and have re-structured the paragraph to:

- “Cochrane full systematic reviews (Cochrane SRs) were established as “gold standard” in knowledge synthesis. A couple of studies compared Cochrane and non-Cochrane systematic reviews and found that the reporting of Cochrane systematic reviews were the most complete one [5], that they were less prone to bias due to greater transparency in reporting as well as due to the quality criteria used, such as the risk of bias assessment [4]. Furthermore, Hopewell et al. described that the inclusion of grey literature routinely performed in Cochrane SRs limited publication bias and provided more conservative treatment effects compared to SR without including grey literature.[6]“

- Page 5, lines 19-23: this information is inaccurate and poorly described. These are other databases (DARE and HTA) within the Cochrane Library. They are distinct from the CDSR (Cochrane Database of Systematic Reviews) that house only Cochrane reviews. It’s important that this is clear as the groupings of reviews (and comparisons of interest) in this study is dependent on the accuracy of this information. [SEP]

- Response: Yes, these databases were named Cochrane databases of other reviews. However, in the meantime these databases have been referred (back) to the UK University of York, Center of Reviews and Dissemination. Therefore, we have adopted it across the manuscript. In the Introduction:

- “To cover the variety of rapid reviews we included rapid reviews listed in medical databases as well as other rapid reviews from the databases of the University of York, Center for Reviews and Dissemination ([www.crd.york.ac.uk/CRDWeb](http://www.crd.york.ac.uk/CRDWeb)), such as the Dare reviews (Database of Abstracts of Reviews of Effects) or rapid reviews as part of health technology assessments (CRD RRs).”

- Page 5, lines 35-37: I would specify who is typically included as stakeholders. [SEP]

Response: We have added this suggestion to the article:

“Hence, this targets not only decision makers, health professionals and their organisations but also citizen or patients, other researchers and the media.”

- Page 5, lines 43-47: need to define “health services research” [SEP] Response: This was done as:

“We especially focused on stakeholder involvement among health services research. This embraced the multidisciplinary field of scientific studies regarding social, financing and personal factors, organizational structures and processes, health technologies, and how these factors affect access to, the quality and cost of health care, and ultimately, our health and well-being”.[15]”

- Pages 5-6: Study aim needs editing; some sentences are long and unclear. First mention of

[SEP]“coproduction” at end of “study aim” – this needs to be defined, and how this differs from

stakeholder involvement (or perhaps define stakeholder involvement and the range of

activities). [SEP]

Response: This section was edited and re-phrased. Coproduction was deleted as it was done in the

Introduction section:

“To date, little is known about the extent of stakeholder engagement in systematic and rapid reviews and there have been few efforts to directly report the specific effects regarding stakeholder involvement in a review. We consider this information to be relevant for integrated knowledge translation, the dissemination and acceptance of systematic and rapid reviews in policy and practice. We focused on systematic and rapid reviews as they represent different types of evidence synthesis and are well-established in research and practice. In addition, reviews needed to belong to the area of health services research as an established field for systematic and rapid reviews as well as a best-case sample with a considerable extent of stakeholder involvement. We aimed to assess the extent and characteristics of stakeholder engagement in recently published systematic and rapid reviews and to specifically determine reporting characteristics.”

Methods:

- Under study design, authors state their purpose: “to assess characteristics and reporting of coproduction” – again, need to define coproduction vs. stakeholder involvement <sup>[1]</sup><sub>[SEP]</sub>

Response: We have deleted the term “coproduction”.

There are no unambiguous definitions for these terms, coproduction and stakeholder involvement. They both stand for the concept of establishing interaction, such as relationships and collaborative processes between researchers and stakeholders, to improve the mutual understanding, enhance relevance, dissemination and use of reviews.

Under search strategy: for non-Cochrane SRs, did you ensure that the citations identified in Medline and Embase were not also published as Cochrane SRs (this wouldn't necessarily appear in the title or abstract). Did the authors verify that the records identified were actually “systematic reviews”?

There are many citations that come up in a search for “systematic reviews” that aren't necessarily SRs and/or don't meet specific definitions of SRs. What definition did the authors use? Clarity needed about “Cochrane databases” and databases contained within the Cochrane Library. The authors searched for HTAs – did they consider these to be the same as rapid reviews? What was their definition of an HTA? Again, did they apply a definition or rely on the description in the primary report? How was full text screening performed and how was consensus achieved? How was ‘spot-checking’ done (random sample of reviews, sample of data elements, other)? <sup>[1]</sup><sub>[SEP]</sub>

Response: We had the search strategy and a two-step screening process as described in the Method section. Following the search for systematic reviews as reported by authors in title or abstract, we checked its presence in title and abstract and checked it again during the full text screening.

As described above, the databases were previously named Cochrane databases of other reviews. However, in the meantime these databases have been referred (back) to the UK University of York, Center of Reviews and Dissemination. Therefore, we have adopted it across the manuscript (CRD rapid reviews).

Regarding HTAs: If a rapid (or systematic) review was identified by the search and was part of a HTA it was included.

The full text screening was performed by one reviewer with checking the article regarding study type and the mentioned inclusion and exclusion criteria (methods section). Spot-checking was done for every sixth review per review type and this was not communicated to the first reviewer.

I'm not familiar with sample size calculations for a cross-sectional study (unless comparisons are planned). It does seem a bit concerning that such a small sample of full SRs were included (30 out of 56,986) compared with 30 out of 95 for "CD RRs". It seems that such a small sample of the full SRs would not likely be generalizable across the entire population of SRs. Based on the sample size description it appears that the authors wanted to conduct comparative analyses and subgroup analyses; this should be reflected in the study aims (i.e., to compare across review types and between Cochrane and non-Cochrane – I think these are two comparisons of interest but authors chose to compare across the four groups). [1] [SEP]

Response: If we intend to assess stakeholder engagement and to potentially compare it between systematic and rapid reviews we need a pre-defined sample size. This sample was randomly selected. It cannot be representative for the whole sample as we would then need to define criteria for representativity. We would then need to screen the whole sample for inclusion criteria as "health services research". This is not feasible for thousands of reviews.

Inclusion and exclusion criteria: need definitions of terms: SRs, HTAs, RRs. Authors indicate that "HTAs targeting effectiveness as well as meta-analyses and systematic reviews for effectiveness [1] [SEP]

and utility, pharmaceutical trials under everyday conditions, basic research in care related

fields...were included". This implies that authors included primary research as well as reviews. This needs to be reworded and clarified. Last sentence on types of studies excluded also needs to be clarified, e.g., efficacy studies without assessing QOL – does that imply that efficacy studies (i.e., primary studies) that assessed QOL were included? That doesn't make sense given the focus on reviews.

Response: SRs and RRs were previously defined in the methods section and responses based upon were included above. The sentence in the method section, inclusion and exclusion criteria, was clarified as follows:

"Therefore, we included Health Technology Assessments (e.g. rapid reviews as part of HTAs) targeting effectiveness as well as meta-analyses and systematic reviews for effectiveness and utility, reviews of pharmaceutical trials under everyday conditions, reviews of basic research in care-related fields, reviews of quality research or of methodological developments in the field of health services research, reviews about the development and application of new technologies (e-health) and the implementation of knowledge into the clinical practice."

- Data extraction: How was AMSTAR assessed - in duplicate, process for consensus?

Response: Following instruction the AMSTAR assessment was done by one reviewer and every sixth review per review type was spot-checked. In case of disagreements arguments for each decision were collected, evaluated and discussed.

- Authors should provide questions that were sent to contact authors of included articles. Did they also ask them about the type of stakeholders, or just the number of stakeholders? [1] [SEP]

Response: As described in the Method section, last paragraph of data extraction, we sent the following questions to the contact authors by email:

- Did you involve any stakeholders at any time during the process of your review (yes/no)?
- if yes: at which stages and for which purpose?
- How many stakeholders were involved in total?
- Did you observe or detect any influence/effect of involving stakeholders?

Results: <sup>[1]</sup><sub>[SEP]</sub>

- First sentence should read 89,798 ‘citations’ – we can’t confirm that these were reviews, until they are more closely examined. Why were so many Cochrane SRs excluded following title and abstract screening? It would seem that these would all meet the criteria for a full SR.

Response: Yes, we agree. There were two question sets for the screening: one to check the assessment of the review type (Cochrane SR, non-Cochrane SR, RR, CRD RR) and the next one to check the inclusion and exclusion criteria for the review belonging to the field of health services research or not (as depicted in the method section).

- Second paragraph: information about increase and annual number is not relevant to the study objectives. Also without screening them all to confirm, I am not confident that they all meet the definition of a given category of review type (e.g., SR, RR, etc). I think this paragraph should be removed. <sup>[1]</sup><sub>[SEP]</sub>

- Response: We have removed this paragraph and the corresponding Supplementary File, although some readers might be interested in a, at least approximately, it puts these review types in perspective with respect to their amounts.

- Table 1: I would like to see the breakdown of the AMSTAR assessments by question and study (at least in an appendix). <sup>[1]</sup><sub>[SEP]</sub>

Response: We have added the AMSTAR assessments per study per question as Supplementary File S4.

- Text on page 13: This could be reduced as it is not the primary objective to compare characteristics across the review types. <sup>[1]</sup><sub>[SEP]</sub>

Response: They were assessed as secondary objective. To have an overview of selected study characteristics including the methodological issues, such as the presence of an a priori protocol, a conflict of interest statement and the AMSTAR ratings as well as the thematic focus of the study, we have included Table 1 and some text referring to it. We have reduced the text, have removed the previous Table 2 and have added it as Supplementary Table S2.

- Page 15, first paragraph: Authors comment on SI across years, although the numbers are so small per year and review type that this comparison does not seem valid. I would suggest that paragraph start with the results on overall SI then get into more specific details. <sup>[1]</sup><sub>[SEP]</sub>

Response: We have re-arranged this paragraph according to your suggestion.

Discussion: <sup>[1]</sup><sub>[SEP]</sub>

- Discussion generally needs work, and should align with study objectives. <sup>[1]</sup><sub>[SEP]</sub>

Response: We have aligned the Discussion section with the study objectives, have shortened and re-arranged it according to the BMJ Open recommendations.

- First sentence indicates that SI varied by review type although this comparison was not made explicit as an objective of the study.

Response: In the Method section, sample size determination, we have included the (rough) comparison of the extent of SI according to systematic and rapid review types. It is clear that in the Discussion section you put it into perspective, otherwise it would seem to be fragmented.

- Third sentence indicates there was a 'significant difference' in the phases of SI; however, there were no tests of statistical significance reported. The authors need to re-word this and be careful with wording and any implications of 'significant differences'. However, the authors did report on sample size calculations, presumably with the intent to test whether there were differences in SI. Also they comment on planned subgroup analyses but none are reported, only descriptive/summary statistics.

Response: Due to the unexpectedly low amounts of SI, subgroup analyses did not make sense.

- Page 19, first paragraph, the following statement needs to be clarified: "Of note, the majority of RR and SR authors reported that the involved stakeholders had positive and considerable effects on the study results, making the review more relevant for the targeted audience." In Table 4 only 33% reported a positive effect and 67% did not mention an effect.

Response: This was a misunderstanding. In the article, we refer to rapid reviews and non-Cochrane systematic reviews and not to the whole sample of reviews. Then, we have e.g. 11 (85%) of contact authors of rapid reviews reporting a substantial or positive effect.

- Further, I don't see in the results any data on relevance of the review for the target audience – was this information collected or implied by the authors? The authors also indicate that their results are "in agreement with the earlier described findings of Keown et al.[143] who concluded, that stakeholder involvement led to increased relevance and depth of review findings, more clarity in defining research questions, broader dissemination of their results and increased awareness of target groups"; however, I don't see this level of detail collected or reported in the present study.

Response: This sentence was rephrased to:

"The reported involvement of stakeholders corresponded with Cottrell et al. who mentioned that stakeholders might contribute to different study types and evidence phases.[142] Based upon, Keown et al.[143] concluded, that stakeholder involvement led to an increased relevance and depth of review findings, more clarity in defining research questions, broader dissemination of their results and increased awareness of target groups. Although they mentioned that this engagement process required flexibility and might be resource- and time-intensive, they nevertheless concluded that involving stakeholders facilitated implementation and should be indispensable for future research."

- Page 19, last paragraph: I don't think this is relevant to the present study, i.e., discussion of quality and characteristics across review types. Page 20, first paragraph: this also was not related to the aim of the study.

Response: We have shortened this.

- Authors should comment on the purpose and nature of RRs and SRs; they are conducted for different purposes so the different levels of SI are not necessarily surprising. Given their different purpose, is it reasonable to have the same level of SI for the different review types? This type of discussion would be valuable. [SEP]

Response: Related to the purpose of a review different stakeholders or stakeholder groups might be included. This does not mean to abstain from the engagement of stakeholders. We do not have investigated which stakeholder groups might be relevant to be included.

- Data mentioned in the discussion around response rates from review authors should be presented more fully in the results. [SEP]

Response: This was re-located.

- The following text from page 21 is concerning: “In their emails, some authors mentioned experts

or peer reviewers and listed them as stakeholders. One author listed members of the research team, who performed literature research and data extraction as stakeholders. Given all the information about who was involved in the reviews, it was sometimes difficult to conclude whom we would count as stakeholder and whom as “expert” from the same, or from a related field.” Given the varied (and confusing) responses regarding SI, it seems the emails or questions therein were not clear about what the authors were asking. I believe Cochrane recommends that the review team be comprised of clinical experts and those with technical expertise to carry out the review; this author group is different than stakeholders, and should have been clarified upfront. Likewise, I wouldn’t consider the peer-review process as stakeholder involvement in the review (or “coproduction”). [SEP]

Response: Yes, we agree and these groups were not classified as stakeholders. But it shows the difficulties of the stakeholder definition and was therefore mentioned as a limitation of the stakeholder assessment.

- Discussion should include a section on “strengths and limitations” of the study, and possibly “future research needs”. A “conclusions” section would be helpful as well – what are the key messages, and new findings this study provides? [SEP]

Response: Yes, we have re-structured the Discussion section according to the BMJ Open recommendations..

Supplementary File S1 (study protocol): [SEP]

- The objective stated in the study protocol is different than the manuscript. The protocol lists primary and secondary outcomes - I think these are actually primary and [SEP]secondary objectives or aims; however, it would be beneficial to have these in the manuscript [SEP]and to present the results according to these primary and secondary points. [SEP]

Response: We have done it accordingly.

- The protocol indicates: “We will compare the proportion of stakeholders involved between

[SEP]groups using the non-parametric Wilcoxon rank sumtest.” This was not described in the

methods of the manuscript, and was not reported on. The protocol indicates that differences

between protocol and review (i.e., study) will be listed if applicable; however, no differences

and reasons are listed in the manuscript. <sup>[1]</sup><sub>SEP</sub>

Response: Due to the unexpectedly low amounts of SI subgroup analyses did not make sense. We have added this to the protocol.

Page 45: the authors list their search strategy. I would have recommended using an established search filter for SRs.

Response: According to our specialized librarian we used the self-declaration of the authors for the search strategy in order to perform it consistently for RRS and SRs as well as to avoid classification differences between the search filters for reviews of the different databases.

Page 50: Not all STROBE checklist items are accurately reported, e.g., item 9 bias, authors report using AMSTAR. This is a tool to assess the methodological quality of SRs. It does not assess potential bias in

Response: The STROBE assessment was replaced by the PRISMA checklist as requested by other reviewers.

the conduct of the present study. Item 21, generalizability: I don't really see a discussion about generalizability and I think this is a major concern given the small sample, particular with respect to the huge number of SRs identified.

Page 52: there are a number of citations in the excluded list that have "systematic review" in the title; it's unclear why these were excluded.

Response: An exclusion criteria was also when they did not fulfill the criteria for belonging to the area of "health services research".

Reviewer: 3

Reviewer Name: Adrian Barnett

Institution and Country: Queensland University of Technology, Australia

This study concerns the important topic of engaging end users in research. The ideas and design were clearly explained. The methods were appropriate and the results were useful and well presented. I have only minor comments.

There was no mention of the James Lind Alliance who do excellent work in ensuring that end-users are involved in research. They also have a good idea about the extra time and resources needed, and the benefits.

<http://www.jla.nihr.ac.uk/>

Response: Thank you for this good example. We have added the James Lind Alliance as an example organization ensuring stakeholder engagement to the Discussion section.

Would it be worth mentioning the recent controversy around the Cochrane systematic review of myalgic encephalomyelitis? This would not get into the issues of the complaints, but it is a very topical example of how end users involvement (or lack of) in systematic reviews is a topical issue.

Response: This is an important discussion but would go beyond the scope of this study as it explicitly calls on concerns from the perspective of endusers. We do not have assessed the perspective of endusers.

Better engagement of stakeholders might reduce research waste, and systematic reviews are a great source of waste, see for example DOI: 10.1111/1468-0009.12210.

Minor comments

- the use of SR and RR in the abstract made it hard to read. I would also avoid these acronyms in the main text.

Response: I have replaced these acronyms overall the manuscript. The four types of reviews were named as follows: Cochrane systematic review, non-Cochrane systematic review, rapid review, CRD rapid review.

- page 4, "concluded that the reporting of Cochrane SRs was the most complete one", compared with what? Also repeated in the next sentence.

Response: This sentence was reformulated as follows:

"A couple of studies compared Cochrane and non-Cochrane systematic reviews and found that the reporting of Cochrane systematic reviews were the most complete one [5], that they were less prone to bias due to greater transparency in reporting as well as due to the quality criteria used, such as the risk of bias assessment [4]."

- the authors use both 'stakeholders' and 'end-users', would it be better to use just one, and/or explain the difference between the two?

Response: We agree we used the term "end-users" only once in the Introduction. We replaced it by stakeholders.

- page 7, lines 42 to 53, this is a long passive sentence. I suggest moving 'were included' towards the start of the sentence. The last sentence in this section is also a relatively long passive sentence.

Response: We have changed these two sentences according to your suggestion.

- page 8, line 38 "checklist was reporting", should be 'reported' or 'recorded'. I would also start a new paragraph at the end of this sentence.

Response: We have changed these two points according to your suggestion.

- Table 1 footnote, might be worth mentioning that higher AMSTAR scores indicate higher quality Response: This is an important point, thank you, and we have added it to the footnote of Table 1.

- Table 2, decimal places are not needed in the population size results

Response: That is correct and we have deleted these decimal places.

- page 19, "We, therefore, suggest that the contribution of involved stakeholders in RRs seemed to be more substantial than in SRs in general and still needs improvement" This sentence is not clear, does it mean improvements are needed for systematic reviews?

Response: That is correct and we have deleted these decimal places.

- page 21, "One author listed members of the research team, who performed literature research and data extraction as stakeholders." shouldn't this be 'erroneously mentioned'?

Response: Our intention was just to mention this fact as an example revealing the challenges of assessing stakeholder involvement and not to judge it. Based on your suggestion we changed this sentence to "erroneously mentioned".

Reviewer: 4

Reviewer Name: Agnes Dechartres

Institution and Country: Sorbonne University, Paris, France

Review for BMJ open, manuscript entitled « Characteristics of stakeholder involvement in systematic and rapid reviews : a cross sectional study in the era of health services research »

This manuscript reports the results of a methodological review focusing on a random sample of systematic and rapid reviews (including those conducted by the Cochrane Collaboration) that were published between Jan 2011 and Oct 2015. The total number of included reviews is 120 : 30 for each of the following categories : Cochrane systematic reviews, other systematic reviews, Cochrane rapid reviews, other rapid reviews. The authors found stakeholder involvement in 13% of Cochrane reviews, 20% of other systematic reviews, 40% of rapid reviews and 93% of Cochrane rapid reviews. The topic is interesting but I found some parts of the manuscript long and I think it could be more straightforward as it is difficult to see what is the most important to retain and what this study adds.

Response: The manuscript was revised and shortened accordingly.

Here are my detailed comments :

- Some wording frequently used are not clear o General and reporting characteristics. What are 'reporting characteristics' ? general characteristics can also be considered as reported characteristics.

Response: "General" was deleted.

o Coproduction

For example: in the study aim section: « and to specifically point out reporting and study characteristics as well as measures and identified challenges of coproduction ».

What is the difference between reporting and study characteristics?

Response: "Study characteristics" were deleted.

What do the authors mean by challenges of coproduction?

Response: We agree that this formulation might be misleading. We did not mean that we assess reasons why stakeholder involvement might be challenging. We, rather, intended to point to the extent of stakeholder involvement being a challenge. We have changed this formulation.

- The introduction is much too long.

Response: The Introduction was shortened and has been kept more straightforwardly.

- The authors conducted a specific search for each of the 4 categories (Cochrane systematic reviews, other systematic reviews, Cochrane rapid reviews, other rapid reviews). They give the results of the search as if all records retrieved corresponded to what they searched, which is not the case.

Therefore, the authors cannot say « with 57,822 articles remaining after exclusion of duplicates... 56,986 were categorized as SR. Not all the records retrieved by the search actually correspond to a systematic review. The authors cannot base parts of their results page 9 lines 5 to 38 on that.

- The search is not really recent : date of last search is October 2015 so 3 years ago. The sample of reviews used can be outdated, which limits the generalisability of findings in the current context.

- The sample size calculation does not correspond to what the authors did. In sample size calculation, the authors planned to compare rapid reviews and systematic reviews but they constructed the study and presented the results for the 4 categories.

- The concept of health services research as defined by the authors is very wide and the interest of involving stakeholders may vary depending on the topic.

Response: This must be investigated. If stakeholders are affected by the topic (as definition) one could argue that they could be interested in being involved.

- This methodological review aims to evaluate stakeholder involvement in reviews but they did not involve stakeholder involvement themselves, which would have been of interest.

Response: There is evidence that stakeholder involvement would be beneficial for reviews as was shortly mentioned in the introduction. In case we would have involved stakeholders in this methodological study it would look like “a self-fulfilling prophecy” and this would not be unexpected.

- The first part of the results including description of characteristics of the reviews is very long and the authors could come more quickly to the main results of this study.

- The authors probably did not go sufficiently in depth on the interest of involving stakeholders.

- I think it would be useful to present some key examples in an appendix

- The authors should also be more careful on their interpretation of results in the discussion. For example : « the higher median population as well as the less specific age range or study setting in non Cochrane reviews illustrates that the studies they included had a broader spectrum of possibly included patients ».

Response: We went throughout the manuscript, shortened and rephrased the results and discussion section more straightforwardly.

Minor comments

- Abstract : the objective could be more clearly formulated Response: We have revised this accordingly:

“Objective Engaging stakeholders in reviews is considered to generate more relevant evidence and to facilitate its dissemination and use. As little is known about stakeholder involvement, we assessed the extent and the characteristics of their engagement in systematic and rapid reviews. Stakeholders were defined as people with a particular interest in the research topic (but were not members of the primary research team). “

- Abstract: I would not present the number of variables targeting stakeholder involvement as it is not really important and it is difficult in the methods of the manuscript to know what are these 9 variables

Response: we have deleted the number of variables (9).

- The sentence « 33% of the responding contact authors mentioned unspecified positive effect of stakeholder involvement » is unclear

Response: This sentence was rephrased to:

“Overall, 33% (17/51) of the responding contact authors mentioned, but did not specify, positive effects of stakeholder involvement.”

- Abstract conclusion should be in line with results presented in abstract

Response: This was done accordingly:

Conclusion Stakeholder engagement was not general practice in either review type (except for CRD rapid reviews), although it was more common in rapid reviews. Especially for systematic reviews, there is still a considerable potential to improve reporting, e.g. by including an a priori study protocol and a conflict of interest statement or to engage stakeholders.”

- The study aim mentions challenges of stakeholder involvement but it is not what they evaluate.

Response: We agree that this formulation might be misleading. We did not mean that we assess reasons why stakeholder involvement might be challenging. We, rather, intended to point to the extent of stakeholder involvement being a challenge.

- The Strobe statement is not really appropriate. This study is more a methodological review so Prisma could be more appropriate

Response: It is correct that stakeholder engagement is a methodological issue. We have assessed it across a random sample of reviews. The question is whether you classify this approach as a review or as a cross-sectional design. We have the impression that in the context of reviews the term “methodological review” is more accepted and therefore, we have changed it throughout the manuscript accordingly. Furthermore, we have added the PRISMA checklist instead of the Strobe statement, although not all of the steps described were applicable.

- In data extraction, the effect on the review outcome is not really clear Response: We have rephrased this sentence at the end of the methods section to:

“We asked them, if there had been any stakeholder involvement and if so, to specify the number of stakeholders, the stages of involvement and their self-reported estimation of the effect of stakeholder involvement on the review outcome.”

- In discussion, the authors should avoid ‘significantly’ as they did not conduct any formal comparison Response: We have checked this throughout the manuscript and have replaced it.

- In Discussion, Prisma is not a guideline aiming to improve methodological quality but to improve reporting  
 Response: Yes, we agree PRISMA is aiming to improve the quality of reporting and the cited references are based upon. However, reporting can be considered as one criterion when assessing the methodological quality because it helps to identify what was done and how the procedure was and therefore, contributes to transparency.

We have replaced “methodological quality” by “reporting quality”.

### VERSION 2 – REVIEW

<b>REVIEWER</b>	Irina Kinchin CQUniversity, Australia
<b>REVIEW RETURNED</b>	19-Feb-2019

<b>GENERAL COMMENTS</b>	I would like to thank the authors for addressing comments and substantially revising their manuscript. Happy to support its publication in current form.
-------------------------	--

<b>REVIEWER</b>	Adrian Barnett Queensland University of Technology Australia
<b>REVIEW RETURNED</b>	15-Feb-2019

<b>GENERAL COMMENTS</b>	<p>The authors have answered my questions and I have only minor comments.</p> <p>Page 8, line 31. How often were reviews with an "unclear design" excluded? Removing lots of these reviews would likely bias the sample towards a higher quality population.</p> <p>Page 12, lines 35 to 45. The wording switches confusingly between talking about studies that did and did not mention the protocol. I would be consistent.</p> <p>Page 15, line 11. An "exponential" increase has a very specific meaning and is not a synonym for a very large increase. Was the increase really exponential?</p>
-------------------------	---

<b>REVIEWER</b>	Agnes Dechartres Sorbonne University
<b>REVIEW RETURNED</b>	16-Feb-2019

<b>GENERAL COMMENTS</b>	<p>Review for BMJ open, manuscript entitled « Characteristics of stakeholder involvement in systematic and rapid reviews : a cross sectional study in the era of health services research »</p> <p>The manuscript is improved. I still have comments as the authors did not answer some of my previous points :</p> <ul style="list-style-type: none"> <li>- The authors conducted a specific search for each of the 4 categories (Cochrane systematic reviews, other systematic reviews, Cochrane rapid reviews, other rapid reviews). They give the results of the search as if all records retrieved corresponded to what they searched, which is not the case. Therefore, the authors cannot say « with 57,822 articles remaining after exclusion of</li> </ul>
-------------------------	---

	<p>duplicates... 56,986 were categorized as SR. Not all the records retrieved by the search actually correspond to a systematic review. The authors cannot base parts of their results page 9 lines 5 to 38 on that.</p> <p>The authors did not answer that point. I think it is just a problem of formulation in both abstract and results but they need to change that. I would also remove these numbers from Table 1</p> <ul style="list-style-type: none"> <li>- The search is not really recent : date of last search is October 2015 so 3 years ago. The sample of reviews used can be outdated, which limits the generalisability of findings in the current context.</li> <li>- The sample size calculation does not correspond to what the authors did. In sample size calculation, the authors planned to compare rapid reviews and systematic reviews but they constructed the study and presented the results for the 4 categories.</li> <li>- The concept of health services research as defined by the authors is very wide and the interest of involving stakeholders may vary depending on the topic.</li> </ul> <p>The authors partly respond this comment but this should be discussed and perhaps investigated.</p> <p>Other comments based on the revised version</p> <ul style="list-style-type: none"> <li>- Abstract : paragraph search strategy : 1) AMSTAR is a tool for assessing methodological quality of systematic reviews not quality of reporting. 2) I would say « characteristics of stakeholder involvement »</li> <li>- Abstract : results. The last sentence is unclear. We wonder what is the link with stakeholder involvement. The authors also evaluated other characteristics of the identified systematic reviews such as presence of a study protocol or reporting of a conflict of interest statement but it is not related to their main objective. Therefore, description of these characteristics makes the reading a little bit confusing.</li> <li>- Abstract : conclusions. Same comment as above. I think the authors should conclude on stakeholder involvement and not on including an a priori protocol</li> <li>- The introduction is still very long but more straightforward</li> <li>- Methods : the sentence « JF following introduction and spot checking » is unclear</li> </ul>
--	---

## VERSION 2 – AUTHOR RESPONSE

**Reviewer: 3**

Reviewer Name: Adrian Barnett

Institution and Country: Queensland University of Technology, Australia

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

The authors have answered my questions and I have only minor comments.

Page 8, line 31. How often were reviews with an "unclear design" excluded? Removing lots of these reviews would likely bias the sample towards a higher quality population.

*Response: We had included this exclusion criteria to be prepared for such a case. But in practice, we were not aware of such a case. We agree, otherwise, the sample might have been biased towards a higher quality population of articles.*

Page 12, lines 35 to 45. The wording switches confusingly between talking about studies that did and did not mention the protocol. I would be consistent.

*Response: Thank you for this comment. We have re-arranged these sentences to be more consistent.*

*The following changes were made:*

*„As an example, all 30 Cochrane systematic reviews (100%) mentioned a pre-existing review protocol, whereas a high amount of non-Cochrane systematic reviews (59%) and of rapid reviews (43%) did not clarify whether there had been a protocol or not.“*

Page 15, line 11. An "exponential" increase has a very specific meaning and is not a synonym for a very large increase. Was the increase really exponential?

*Response: We have re-worded this sentence to be less confusing.*

*This sentence was changed as follows:*

Interestingly, the proportion of stakeholder involvement has been quite constant between the years 2012 (5/15, 33%) and 2014 (12/34, 34%) even though the amount of rapid reviews has increased *substantially* (Table 2).

#### **Reviewer: 4**

Reviewer Name: Agnes Dechartres

Institution and Country: Sorbonne University

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

Review for BMJ open, manuscript entitled « Characteristics of stakeholder involvement in systematic and rapid reviews : a cross sectional study in the era of health services research»

The manuscript is improved. I still have comments as the authors did not answer some of my previous points :

- The authors conducted a specific search for each of the 4 categories (Cochrane systematic reviews, other systematic reviews, Cochrane rapid reviews, other rapid reviews). They give the results of the search as if all records retrieved corresponded to what they searched, which is not the case. Therefore, the authors cannot say « with 57,822 articles remaining after exclusion of duplicates... 56,986 were categorized as SR. Not all the records retrieved by the search actually correspond to a systematic review. The authors cannot base parts of their results page 9 lines 5 to 38 on that. The

authors did not answer that point. I think it is just a problem of formulation in both abstract and results but they need to change that. I would also remove these numbers from Table 1

*Response: We apologize not having sufficiently responded to your comments and we give our best to satisfy your expectations.*

*As described in the Method sections, we searched for the references by potential review category there were supposed to belong to. We agree that not all records retrieved corresponded to what we had searched for and therefore, we had screened for eligibility. We, therefore, presented it more clearly in Figure 1 and moved the initial potential review groups more upwards to the identification stage and labelled them as e.g. "potential RR" etc..*

*To be more clear and consistent we adapted Figure 1 and the text of the Result section as follows:*

*„Our search identified 57,822 citations remaining after exclusion of duplicates. Although, not all of these records will fulfil the inclusion criteria, such as e.g. being designed as a systematic review, we assigned them as potential review group. For the screening step, the counts of articles that needed to be screened to reach the final set of studies are depicted in Figure 1.[22-141] The excluded studies are presented in Supplementary Table S1.“*

*The corresponding information was removed from Table 1 as requested.*

- The search is not really recent : date of last search is October 2015 so 3 years ago. The sample of reviews used can be outdated, which limits the generalisability of findings in the current context.

*Response: We have transparently reported the method and framing conditions of our study. The chosen search date of our study sample does not restrict the generalizability of our findings. Importantly, it shows the findings within the defined time period and potential changes over time might be presented in a follow-up study.*

- The sample size calculation does not correspond to what the authors did. In sample size calculation, the authors planned to compare rapid reviews and systematic reviews but they constructed the study and presented the results for the 4 categories.

*Response: Yes, our intention was to compare systematic review and rapid reviews which we also report in the article. For each of the two general review groups we included a benchmark group as a potential gold standard regarding conduct and reporting (Cochrane SR and CRD RR). We could not base the sample size calculation on them as this would have resulted in a biased estimate (inclusion of high-quality reviews only).*

- The concept of health services research as defined by the authors is very wide and the interest of involving stakeholders may vary depending on the topic.

The authors partly respond this comment but this should be discussed and perhaps investigated.

*Response: Our approach was to include reviews targeting the field of health services research. Therefore, we were aware of the broad definition. But this was considered essential as there is no rationale why stakeholder involvement should not be relevant for every considerable subtype of health services research. Knowledge transfer and dissemination is even an important topic of all health research.*

Other comments based on the revised version

- Abstract : paragraph search strategy : 1) AMSTAR is a tool for assessing methodological quality of systematic reviews not quality of reporting. 2) I would say « characteristics of stakeholder involvement »

*Response: We have changed the two sentences as requested:*

*“Their methodological quality was assessed using AMSTAR and characteristics of stakeholder involvement were collected.”*

- Abstract : results. The last sentence is unclear. We wonder what is the link with stakeholder involvement. The authors also evaluated other characteristics of the identified systematic reviews such as presence of a study protocol or reporting of a conflict of interest statement but it is not related to their main objective. Therefore, description of these characteristics makes the reading a little bit confusing.

*Response: Stakeholder involvement should not be separated from an attempt to conduct a review methodologically sound and appropriate. This would also contribute to reduce research waste. Therefore, we favored an integrated approach. This is then also one part of the study characteristics being present or not when evaluating stakeholder involvement. An a priori protocol is important as it offers the opportunity to describe the stakeholder involvement or its absence, respectively. Following that, the conduct and the stages of stakeholder involvement might be described a priori and allow for a methodologically sound procedure.*

*Furthermore, there might be a situation where involved stakeholders might lead to a conflict of interest. This should be assessed and mentioned to reach transparency*

- Abstract : conclusions. Same comment as above. I think the authors should conclude on stakeholder involvement and not on including an a priori protocol

*Response: We would like to favor an integrated approach and not separating stakeholder involvement from a transparent study conduct. See comments above.*

*An a priori protocol is important as it offers the opportunity to describe the stakeholder involvement or its absence, respectively. Following that, the conduct and the stages of stakeholder involvement might be described a priori and allow for a methodologically sound process.*

- The introduction is still very long but more straightforward

*Response: Our aim was to shortly describe the different review types and their potentially contributing stakeholders as well as to define the terms “stakeholder” and “health services research” in order to provide this basis as a background of our study.*

- Methods : the sentence « JF following introduction and spot checking » is unclear.

*Response:*

*This sentence was re-worded to:*

*“JF and MM tested their consistency in assessing the inclusion and exclusion criteria for about ten different reviews and MM checked the assignment for about half of the study sample.”*

**Reviewer: 1**

Reviewer Name: Irina Kinchin

Institution and Country: CQUniversity, Australia

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

I would like to thank the authors for addressing comments and substantially revising their manuscript. Happy to support its publication in current form.

**VERSION 3 – REVIEW**

<b>REVIEWER</b>	Agnes Dechartres Sorbonne University
<b>REVIEW RETURNED</b>	18-Apr-2019

<b>GENERAL COMMENTS</b>	<p>Review R2 Characteristics of stakeholder involvement in systematic and rapid reviews : a methodological review in the area of health services research</p> <p>I would like to thank the authors for answering my comments. I understand their point about the need for integrated approach accounting for both stakeholder involvement and evaluation of methodological quality. However, the manuscript focuses on two points: stakeholder involvement and evaluation of methodological quality with no real interconnection between them. I just need some consistency in the abstract as in the objective it is clearly focused on stakeholder involvement. There is only the reporting of conflict of interest in the results and part of the conclusion is on the need to include an a priori study protocol. This lack of consistency makes the reading more difficult.</p> <p>I would add in the limitations that the search is not recent.</p>
-------------------------	---

**VERSION 3 – AUTHOR RESPONSE****Reviewer: 4**

Agnes Dechartres

Institution and Country

Sorbonne University

Review R2 Characteristics of stakeholder involvement in systematic and rapid reviews : a methodological review in the area of health services research I would like to thank the authors for answering my comments.

I understand their point about the need for integrated approach accounting for both stakeholder involvement and evaluation of methodological quality. However, the manuscript focuses on two

points: stakeholder involvement and evaluation of methodological quality with no real interconnection between them. I just need some consistency in the abstract as in the objective it is clearly focused on stakeholder involvement. There is only the reporting of conflict of interest in the results and part of the conclusion is on the need to include an a priori study protocol. This lack of consistency makes the reading more difficult.

**Response:** *We have added the mentioned missing information to the abstract to make it more consistent:*

**Objective** Engaging stakeholders in reviews is considered to generate more relevant evidence and to facilitate dissemination and use. As little is known about stakeholder involvement, we assessed the characteristics of their engagement in systematic and rapid reviews and the methodological quality of included studies. Stakeholders were people with a

particular interest in the research topic.

**Design** Methodological review

**Search strategy** Four databases (Medline, Embase, Cochrane database of systematic reviews, databases of the University of York, Center for Reviews and Dissemination (CRD)) were searched based on an a priori protocol. Four types of reviews (Cochrane and non-Cochrane systematic reviews, rapid and CRD rapid reviews) were retrieved between

January, 2011 and October, 2015, pooled by potential review type and duplicates excluded. Articles were randomly ordered and screened for inclusion and exclusion criteria until 30 reviews per group were reached. Their methodological quality was assessed using AMSTAR and stakeholder characteristics were collected.

**Results** In total, 57,822 deduplicated citations were detected with potential non-Cochrane systematic reviews being the biggest group (56,986 records). We found stakeholder involvement in 13% (4/30) of Cochrane, 20% (6/30) of non-Cochrane, 40% (13/30) of rapid and 93% (28/30) of CRD reviews. Overall, 33% (17/51) of the responding contact authors mentioned positive effects of stakeholder involvement. A conflict of interest statement remained unmentioned in 40% (12/30) of non-Cochrane and in 27% (8/30) of rapid reviews, but not in Cochrane or CRD reviews. At maximum half of non-Cochrane and rapid reviews mentioned an a priori study protocol in contrast to approximately all Cochrane and CRD reviews.

**Conclusion** Stakeholder engagement was not general practice, except for CRD reviews, although it was more common in rapid reviews. Reporting factors, such as including an a priori study protocol and a conflict of interest statement should be considered in conjunction with involving stakeholders, especially for non-Cochrane systematic reviews.

I would add in the limitations that the search is not recent.

**Response:** *We have added the following sentences in the Discussion section as a limitation (5<sup>th</sup> paragraph):*

“Although our search is not very recent this article highlights the current situation and there is a call for action. Of course, an updated assessment might be needed in the years following.”