

PEER REVIEW HISTORY

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

ARTICLE DETAILS

TITLE (PROVISIONAL)	Reproducible research practices, openness and transparency in health economic evaluations: study protocol for a cross-sectional comparative analysis
AUTHORS	Catalá-López, Ferrán; Caulley, Lisa; Ridao, Manuel; Hutton, Brian; Husereau, Don; Drummond, Michael; Alonso-Arroyo, Adolfo; Pardo-Fernández, Manuel; Bernal-Delgado, Enrique; Meneu, Ricard; Tabarés-Seisdedos, Rafael; Repullo, José; Moher, David

VERSION 1 - REVIEW

REVIEWER	Barbara Claus Ghent University, Ghent, Belgium Belgian Reimbursement Committee
REVIEW RETURNED	10-Oct-2019

GENERAL COMMENTS	<p>Meaningful research! Two additional comments: Did you consider to study associations between for example the general characteristics of your extracted studies and your enablers of reproducibility, transparency and openness? Then this should also be mentioned.</p> <p>Second, from the perspective of a healthcare payer, it could be interesting to know if more transparency, reproducibility...(or some elements of them) are associated with other (higher?) base case ICERs. Although your methodology is not designed for this, and this might be highly speculative and will depend on the nature of your extracted studies such as disease area studied (which is not predictable at this point), this can be an element to keep in mind for a subsequent follow-up or ad hoc analysis.</p>
-------------------------	---

REVIEWER	Joanna Thorn University of Bristol
REVIEW RETURNED	28-Nov-2019

GENERAL COMMENTS	The authors describe a protocol for a literature review aiming to address issues of reproducibility, openness and transparency in the conduct and reporting of economic evaluations. This is an important and timely study; however, I'm not convinced of the additional value of waiting until 2023 to be able to include 2022 studies. There is a clear rationale for anticipating a change
-------------------------	---

between 2012 and 2019 (publication of CHEERs), but the rationale for expecting any particular change in reporting standards between now and 2022 is less convincing. The research question (the extent to which ideals of reproducibility, openness and transparency are met in economic evaluations) could usefully be answered with 2019 data. If the data show that there is still a problem, it is better addressed earlier than later.

Specific comments

- Introduction: at the end of the paragraph beginning Jefferson et al, the results are quoted rather baldly and without context. Presumably, the very short timescale meant that there wasn't really time to see a difference? Is there really no further follow-up showing the effectiveness or otherwise of the guideline?
- Screening: If liberal acceleration is being used for title/abstract screening, presumably it is only discrepancies in the full-text screening that will be resolved via discussion?
- Data extraction: what is the strategy if fewer than 200 articles are identified in a given year?
- General characteristics: this would be easier to read in list format
- Enablers...: Again, this would be easier to read as a list. It would also be helpful to separate it out into the three different types of indicator (although there may be some overlap).
- Enablers...: I would query whether the citation of the CHEERs statement by itself is adequate as an indicator – I think there needs to be some assessment of how well the report adheres to the statement. It would be quite easy to mis-specify pages in the CHEERs checklist – eg iterative responses to referees could lead to some elements falling out of the report.
- Enablers...: The list of data to extract is extensive, but there is an omission that the authors might like to consider including as an indicator. There has been growing interest recently in the use of Health Economics Analysis Plans (HEAPs) that set out the proposed economic analysis in advance. [I have recently conducted a Delphi survey to identify the key items that should appear in a HEAP.] HEAPs aim to reduce un-specified post hoc analyses and therefore contribute to transparency. The pre-publication of a HEAP (whether peer-reviewed or deposited in a repository) is a potentially useful indicator of transparency.
- Patient and public involvement: Although it is appropriate that patients/public are not involved in designing or interpreting the study, I think it would be a mistake not to disseminate the results to the public. It's a bit of a jarring statement for a study about transparency. Hopefully, the study will inspire confidence in research but, even if the results are concerning, the public deserves to know.
- Ethics and dissemination: could the authors expand a little on how the long list of potentially interested audiences might use the results?
- Search strategy: Why isn't "economic evaluation" included as a search term?
- One of the indicators of transparency is whether the economic evaluation is published at all, and it is not uncommon for the clinical effectiveness results to be published some time ahead of the economic evaluation. The issue is out of the scope of this study, which focuses on studies that have made it to the publication process, but the authors might like to discuss it in any final report.
- The standard of English is reasonable, but the manuscript would benefit from some tidying up.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer 1: Barbara Claus

Meaningful research!

Two additional comments:

Did you consider to study associations between for example the general characteristics of your extracted studies and your enablers of reproducibility, transparency and openness? Then this should also be mentioned.

Response to Reviewer 1 Comments - Question #1:

Thank you very much for your kind and constructive comments, and for your time to review our manuscript. In “Data analysis” (page 9 on the manuscript), we report that: “The proportion of general, methodological and reproducibility indicators will be reported, stratified by year, citation use of the CHEERS statement, and journal (e.g. according to whether it is an original CHEERS endorsed journal or not).”

Following your suggestions, we have included a new subsection: “Updates and additional analysis”: (...) Any (new) additional analysis examining potential associations between general characteristics from extracted studies (e.g. results including index ICER, or funding source) and enablers of reproducibility, transparency and openness (e.g. mention of CHEERS statement, open access, protocol registration, or mention of raw data) will be prospectively reported in a new specific (sub-study) protocol, following standard methods described in this paper. (...)”

Second, from the perspective of a healthcare payer, it could be interesting to know if more transparency, reproducibility...(or some elements of them) are associated with other (higher?) base case ICERs. Although your methodology is not designed for this, and this might be highly speculative and will depend on the nature of your extracted studies such as disease area studied (which is not predictable at this point), this can be an element to keep in mind for a subsequent follow-up or ad hoc analysis.

Response to Reviewer 1 Comments - Question #2:

Thank you very much for your constructive comments. See our comments above.

Reviewer 2: Joanna Thorn

The authors describe a protocol for a literature review aiming to address issues of reproducibility, openness and transparency in the conduct and reporting of economic evaluations. This is an important and timely study; however, I'm not convinced of the additional value of waiting until 2023 to be able to include 2022 studies. There is a clear rationale for anticipating a change between 2012 and 2019 (publication of CHEERS), but the rationale for expecting any particular change in reporting standards between now and 2022 is less convincing. The research question (the extent to which ideals of reproducibility, openness and transparency are met in economic evaluations) could usefully be answered with 2019 data. If the data show that there is still a problem, it is better addressed earlier than later.

Response to Reviewer 2 Comments - Question #3:

Thank you very much for your kind and constructive comments, and for your time to review our manuscript.

To provide a reliable summary of the literature, we will search MEDLINE® through PubMed candidate studies throughout three cross-sectional, comparative time periods. First, we will search MEDLINE®-indexed articles in 2019 (“reference year”) as it is the year closest to when the protocol for this study was drafted. In part two, we will search for articles indexed in 2012 and 2022, respectively, in order to further assess whether the transparency and reproducibility practices improved between 2012 (as it is one year before the publication of the CHEERS statement in 2013), and 2022 (10 years after).

We would like to clarify that we plan to conduct a continual surveillance of the health economic literature, keeping evidence as up-to-date as possible (2019, 2022, 2025, etc). Accurate reanalysis of the proposed reproducibility and transparency metrics might offer some evidence for whether design, conduct, and analysis of health economic evaluations are improving with time. Then iterations of the searches and review process will be repeated at regular intervals (e.g. 3 year intervals after 2022) to keep our analyses up-to-date over time. (...)

In addition, it is important to note that our experienced team (including research methodologists, health economists, and clinicians) does not have the the technical resources (and/or human capacity) to review all the indicators proposed for more than 400-600 studies in less than 2 years (e.g. before 2023).

Following the Reviewer's comments, we have updated/acknowledged the following in methods section:

(...) We plan to conduct a continual surveillance of the health economic literature, keeping evidence as up-to-date as possible. Iterations of the searches and review process will be repeated at regular intervals (e.g. 3 year intervals after 2022) to continue to present timely and accurate findings. Reanalysis of the proposed reproducibility and transparency metrics may offer insight into progressive improvements in design, conduct, and analysis of health economic evaluations over time. (...) Any (new) additional analysis examining potential associations between general characteristics from extracted studies (e.g. results including index ICER, or funding source) and enablers of reproducibility, transparency and openness (e.g. mention of CHEERS statement, open access, protocol registration, or mention of raw data) will be prospectively reported in a new specific (sub-study) protocol, following standard methods described in this paper. (...)

Specific comments

- Introduction: at the end of the paragraph beginning Jefferson et al, . Presumably, the very short timescale meant that there wasn't really time to see a difference? Is there really no further follow-up showing the effectiveness or otherwise of the guideline?

Response to Reviewer 2 Comments - Question #4:

Thank you. In their discussion, authors acknowledged they "collected few manuscripts relating to the "after" period from BMJ, as manuscripts submitted prior to June 1997 had been shredded and manuscripts submitted from August 1997 onward were still undergoing editorial assessment. Missing manuscripts and exclusive focus on management of economic submissions in only 2 journals may introduce a bias into the study results, hence the absence of statistical analysis of the data".

To our knowledge, and after confirming with the authors, there is no further follow-up showing the effectiveness of the BMJ guideline.

- Screening: If liberal acceleration is being used for title/abstract screening, presumably it is only discrepancies in the full-text screening that will be resolved via discussion?

Response to Reviewer 2 Comments - Question #5:

Thank you for this comment. We have revised the main text as follows: "Any discrepancies in screening full-text articles will be resolved via discussion or adjudication by a third reviewer if necessary."

- Data extraction: what is the strategy if fewer than 200 articles are identified in a given year?

Response to Reviewer 2 Comments - Question #6:

Thank you for this comment.

We have included the following in main text: "If fewer than 200 articles are identified in a given year (e.g. 2012), we will randomly select the sufficient number of studies published from the preceding year (e.g. October-December 2011) to match the number used in the study sample."

Note: to avoid including articles published after March 2013 (e.g. when CHEERS was published).

In our opinion, it is highly improbable that fewer than 200 articles are identified in a given year. For example, Neumann et al recently reported a prevalence of 6981 cost-effectiveness analyses reporting QALYs/DALYs published in 2016, with an incidence roughly 700 and rising (through PubMed search).

Ref. Neumann PJ, Anderson JE, Panzer AD, Pope EF, D'Cruz BN, Kim DD, Cohen JT. Comparing the cost-per-QALYs gained and cost-per-DALYs averted literatures. Version 2. Gates Open Res. 2018 Mar 5 [revised 2018 Jan 1];2:5. doi: 10.12688/gatesopenres.12786.2. eCollection 2018. PubMed PMID: 29431169; PubMed Central PMCID: PMC5801595.2.

- General characteristics: this would be easier to read in list format

Response to Reviewer 2 Comments - Question #7:

Thank you, we agree that this would make sense. We have revised the format as you have suggested.

- Enablers...: Again, this would be easier to read as a list. It would also be helpful to separate it out into the three different types of indicator (although there may be some overlap).

Response to Reviewer 2 Comments - Question #8:

Thank you. We have revised the format as you have suggested.

- Enablers...: I would query whether the citation of the CHEERs statement by itself is adequate as an indicator – I think there needs to be some assessment of how well the report adheres to the statement. It would be quite easy to mis-specify pages in the CHEERs checklist – eg iterative responses to referees could lead to some elements falling out of the report.

Response to Reviewer 2 Comments - Question #9:

Thank you for this important comment.

The selection and wording of general, methodological and reproducibility indicators has been influenced by recommendations in relevant articles on research transparency and reproducibility, including CHEERS. We have revised the indicators, to include some additional assessment of how well the reports the statement. The standardized data extraction form will include also the following CHEERS reporting items: Study perspective (e.g. society, healthcare system/provider) and relate this to the costs being evaluated; Time horizon over which costs and outcomes are being evaluated; Discount rate used for costs and outcomes (when applicable); Health outcomes used as the measure of benefit in the evaluation (e.g. life years gained, quality-adjusted life years or disability-adjusted life years); Measurement of effectiveness (e.g. for single-study based estimates: describe fully the design features of the single effectiveness study, and why the single study was a sufficient source of clinical effectiveness; and for synthesis-based estimates: describe fully the methods used for identification of included studies and meta-analysis of clinical effectiveness data); Estimating resources and costs (describe approaches used to estimate resource use associated with the alternative interventions; and describe methods for valuing each resource item in terms of its unit costs); Discussed all analytical methods supporting the evaluation; and model calibration and validation (when applicable); (...).

- Enablers...: The list of data to extract is extensive, but there is an omission that the authors might like to consider including as an indicator. There has been growing interest recently in the use of Health Economics Analysis Plans (HEAPs) that set out the proposed economic analysis in advance. [I have recently conducted a Delphi survey to identify the key items that should appear in a HEAP.] HEAPs aim to reduce un-specified post hoc analyses and therefore contribute to transparency. The pre-publication of a HEAP (whether peer-reviewed or deposited in a repository) is a potentially useful indicator of transparency.

Response to Reviewer 2 Comments - Question #10:

Thank you for bringing these to our attention. Indeed, we included as an indicator “protocol/registration mentioned (no protocol, full protocol publicly available, full protocol publicly available and preregistered)”, on page 8 of our manuscript.

We now mention health economics analysis plans, as follows: “Health economics analysis plan mentioned (no analysis plan, indicated that analysis plan was available on request, full access to analysis plan along with research protocol)”

In addition, we have included the following new references:

Dritsaki M, Gray A, Petrou S, Dutton S, Lamb SE, Thorn JC. Current UK Practices on Health Economics Analysis Plans (HEAPs): Are We Using Heaps of Them? *Pharmacoeconomics*. 2018 Feb;36(2):253-257. doi: 10.1007/s40273-017-0598-x. PubMed PMID: 29214388.

Aczel B, Szaszi B, Sarafoglou A, Kekecs Z, Kucharský Š, Benjamin D, et al. A consensus-based transparency checklist. *Nat Hum Behav*. 2019 Dec 2. doi: 10.1038/s41562-019-0772-6. [Epub ahead of print] PubMed PMID: 31792401.

- Patient and public involvement: Although it is appropriate that patients/public are not involved in designing or interpreting the study, I think it would be a mistake not to disseminate the results to the public. It's a bit of a jarring statement for a study about transparency. Hopefully, the study will inspire confidence in research but, even if the results are concerning, the public deserves to know.

Response to Reviewer 2 Comments - Question #11:

Thank you. We have revised the text as suggested.

- Ethics and dissemination: could the authors expand a little on how the long list of potentially interested audiences might use the results?

Response to Reviewer 2 Comments - Question #12:

Thank you for this comment. We have revised the text as follows:

“Without complete and transparent reporting of how a health economic evaluation is being designed and conducted, it is difficult for readers and potential knowledge users to assess its conduct and validity. Strengthening the reproducibility and reporting of methods and results can maximize the impact of health economic evaluations by allowing more accurate interpretation and use of their findings. We anticipate the study could be relevant to a variety of audiences including journal editors, peer reviewers, research authors, health technology assessment agencies, research funders, educators and other potential key stakeholders. Moreover, the study findings could further be used in discussions to strengthen Open Science in order to increase value and reduce waste from incomplete or unusable reports of health economic evaluations.

- Search strategy: Why isn't “economic evaluation” included as a search term?

Response to Reviewer 2 Comments - Question #13:

Thank you for bringing this to our attention. The term “economic evaluation” term is typically captured by the term “cost-benefit analysis” [MeSH], that we had incorporated into our search strategy. However, he have modified the strategy to include “economic evaluation”[title].

As you see, the results increased by 148 records (from 103,356 to 103,502), which is 0.14% of the total.

Predefined search strategy:

1.	"cost-benefit analysis"[mh] OR "costs and cost analysis"[mh] OR "cost-effective*"[ti] OR "cost-utility"[ti]	232861
2.	Journal Article[pt] AND hasabstract[text] AND English[lang] AND ("humans"[mh] OR "humans"[All Fields])	10198821
3.	Editorial[pt] OR Letter[pt] OR Historical Article[pt] OR Meta-Analysis[pt] OR Retracted Publication[sb] OR Review[pt] OR systematic[sb]	4566117
4.	#1 AND #2	123453
5.	#4 NOT #3	103356

Modified search strategy:

1.	"cost-benefit analysis"[mh] OR "costs and cost analysis"[mh] OR "cost-effective*" [ti] OR "cost-utility"[ti] OR "economic evaluation" [ti]	233805
2.	Journal Article[pt] AND hasabstract[text] AND English[lang] AND ("humans"[mh] OR "humans"[All Fields])	10198821
3.	Editorial[pt] OR Letter[pt] OR Historical Article[pt] OR Meta-Analysis[pt] OR Retracted Publication[sb] OR Review[pt] OR systematic[sb]	4566117
4.	#1 AND #2	123646
5.	#4 NOT #3	103502

• One of the indicators of transparency is whether the economic evaluation is published at all, and it is not uncommon for the clinical effectiveness results to be published some time ahead of the economic evaluation. The issue is out of the scope of this study, which focuses on studies that have made it to the publication process, but the authors might like to discuss it in any final report.

Response to Reviewer 2 Comments - Question #14:

Thank you for bringing this to our attention. We will consider discussing this in any final report, as suggested.

• The standard of English is reasonable, but the manuscript would benefit from some tidying up.

Response to Reviewer 2 Comments - Question #15:

Thank you, we will be sure to improve the use and flow of language in our final version.