



BRIEF REPORT

REVISÉ Learning from Cochrane systematic reviews: what improvements do these suggest for the design of trials?

[version 2; peer review: 2 approved]

Stefania Piroasca¹, Mike Clarke ², Shaun Treweek ³

¹School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK

²Northern Ireland Methodology Hub, Queen's University Belfast, Belfast, UK

³Health Services Research Unit, University of Aberdeen, Aberdeen, UK

v2 First published: 11 Mar 2020, 9:178
<https://doi.org/10.12688/f1000research.22635.1>

Latest published: 25 Jun 2020, 9:178
<https://doi.org/10.12688/f1000research.22635.2>

Abstract

Background: Many randomised trials have serious methodological flaws that fatally undermine their results, which makes the research wasteful. This is of concern for many, including those doing systematic reviews that include trials. Cochrane systematic reviews have a section called 'Implications for research', which allows authors of the review to present their conclusions on how future research might be improved. Looking at these conclusions might highlight priority areas for improvement.

Methods: We focused on the Cochrane Schizophrenia Review Group and the Multiple sclerosis and rare diseases of the central nervous system Review Group (the MS Review Group). Reviews with citation dates between 2009 and 2019 were identified and the recommendations of review authors in 'Implications for research' were put into categories.

Results: Between 2009 and 2019 we identified 162 reviews for the Schizophrenia Review Group and 43 reviews for the MS Review Group. We created 22 categories of recommendations in total, of which 12 were common to both groups. The five most used categories were the same for both: better choice of outcomes; better choice of intervention/comparator; longer follow-up; larger sample size; use of validated scales. Better choice of outcomes and/or intervention/comparator was recommended in over 50% of reviews. Longer follow-up and larger sample size were recommended in over a third, with use of validated scales being suggested in around a fifth of reviews. There was no obvious pattern of improvement over time for trials included in systematic reviews published by both groups.

Conclusions: We suggest that trialists working in these and other areas ask themselves, or are compelled to do so by others (e.g. funders), why they have chosen their outcomes, intervention and comparator, whether follow-up is long enough, if the sample size is

Open Peer Review

Approval Status

	1	2
version 2 (revision) 25 Jun 2020	 view	
	↑	
version 1 11 Mar 2020	 view	 view

1. **Livia Puljak** , Catholic University of Croatia, Zagreb, Croatia
2. **Christian Gluud** , Copenhagen University Hospital, Copenhagen, Denmark

Any reports and responses or comments on the article can be found at the end of the article.

big enough and whether the scales they choose to measure their outcomes are appropriate.

Keywords

Randomised controlled trial, Trial methodology, Systematic reviews, Schizophrenia, Multiple sclerosis.



This article is included in the [Research on Research, Policy & Culture gateway](#).

Corresponding author: Shaun Treweek (streweek@mac.com)

Author roles: **Pirosca S:** Formal Analysis, Investigation, Methodology, Project Administration, Writing – Review & Editing; **Clarke M:** Conceptualization, Formal Analysis, Methodology, Writing – Review & Editing; **Treweek S:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: MC is the Co-ordinating Editor of the Cochrane Methodology Review Group. ST is an editor (similar to being on an editorial board of a journal) for the Cochrane Incontinence Review Group.

Grant information: The study was funded by the Health Research Board (grant 'Capacity Building for Evidence Synthesis – Evidence Synthesis Ireland 2018') as part of Trial Forge (<http://trialforge.org>), a non-commercial research initiative led by ST. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2020 Pirosca S *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Pirosca S, Clarke M and Treweek S. **Learning from Cochrane systematic reviews: what improvements do these suggest for the design of trials? [version 2; peer review: 2 approved]** F1000Research 2020, 9:178 <https://doi.org/10.12688/f1000research.22635.2>

First published: 11 Mar 2020, 9:178 <https://doi.org/10.12688/f1000research.22635.1>

REVISED Amendments from Version 1

1. We have changed the title to 'Learning from Cochrane systematic reviews: what improvements do these suggest for the design of trials?'

2. The addition of two sections of text to the Conclusions section: 'Previous research found that small, underpowered studies make up the entirety of evidence in most meta-analyses reported by Cochrane reviews'. Only 35% of the reviews in our sample mentioned sample size in the Implication for research, which suggests that reviewers may be underestimating the size of this problem. The most frequent issue raised in our work was the choice of outcome, a persistent problem that led to the COMET initiative to facilitate the development of core outcome sets^{5,8}. Cochrane reviewers need to continue to highlight trial design problems; indeed they could perhaps do so more often, especially around sample sizes.'

'There are resources that can help with the above. For example, COMET (<http://www.comet-initiative.org>) can help with outcome choice and PRECIS-2 can help with design decisions around comparators and follow-up⁹. Discussions with staff at clinical trial units and other research support centres are also likely to improve designs. Change is slow at present and initiatives to encourage, or force, trialists to consider these questions would be welcome, especially from funders.'

3. The addition of three new references (Refs 7–9) to support the above.

4. Provided two additional files containing all the data we extracted. These are available at <https://doi.org/10.17605/OSF.IO/P9CW6>

Any further responses from the reviewers can be found at the end of the article

Introduction

Randomised trials are at the heart of evidence-informed health-care systems. Not all trials, however, are created equal. Some are excellent, others have serious methodological flaws that fatally undermine their results and make the research wasteful^{1–3}. This is of particular importance for the systematic reviews that should be used to underpin decision making in health care, and systematic reviewers often identify relevant issues and suggest ways to improve the quality of future trials. Cochrane is an international organisation that aims to provide high-quality information to support health decisions by systematically reviewing research, especially from randomised trials to investigate the effects of healthcare interventions. It is organized across more than 50 Cochrane Review Groups, each of which looks after a particular area of health.

Cochrane pays great attention to the methodological quality of both the reviews and the studies they include⁴. Cochrane systematic reviews all have a section called 'Implications for research', which allows the authors of the review to present their conclusions on how future research might be improved, for example by discussing the types of interventions that need more evaluation, or the outcomes which should, or should not, be measured and reported⁵.

The work described here looked at the 'Implications for research' sections of reviews published by two Cochrane Review Groups

between 2009 and 2019 (2019 was incomplete at the time of the work). Our aims were to:

1. Categorise the recommendations made in 'Implications for research'.
2. Explore whether there have been changes over time and between the two groups.

Methods

We focused on the Schizophrenia Review Group and the Multiple sclerosis and rare diseases of the central nervous system Review Group (which we call the MS Review Group hereafter). The Schizophrenia Review Group was chosen because it has a long-standing interest in what is written in the 'Implications for research' section and we expected it to have good, consistent reporting. The MS group was selected because one of us (SP) had a special interest in this clinical area.

The work was split into two stages:

1. Identify and extract the 'Implications for research' sections of all reviews in both groups with a citation date of 2009 to 2019. All reviews are available at the following URLs:

1. Schizophrenia Review Group: <https://schizophrenia.cochrane.org/topic-tree>
2. MS Review Group: <https://msrdcns.cochrane.org/our-review>

2. Categorise the recommendations on trial design and methodology made by the review authors.

Stage 1 was straightforward because the 'Implications for research' sections of Cochrane reviews are easily identified. Stage 2 involved reading through the text and putting the issues raised by review authors into categories. We used a simple Excel spreadsheet data extraction form for this (see Data availability). We did not try to fit the data into pre-existing categories but were guided by what came from the data. The initial categorisation was done by SP before being discussed with ST until agreement was reached. Additionally we extracted information on the number of included participants and studies, the risk of bias etc. (See Data Availability for a file containing the full list). No statistical analysis was necessary beyond counting.

Results

For the period 2009 to 2019, we identified 162 reviews for the Schizophrenia Review Group and 43 reviews for the MS Review Group. A wide variety of interventions were covered by these reviews, including drugs, educational and behavioural interventions, and other therapies such as physical exercise and acupuncture. The median number of included studies in reviews from the Schizophrenia Group was seven (range 0 to 174); for the MS Group it was five (range 0 to 45).

We created 22 categories in total, of which 12 were common to both review groups (Table 1). Six categories were unique to the Schizophrenia Review Group and four were unique to the MS Review Group. The five most used categories were the same for both groups (better choice of outcomes, better choice of

Table 1. Recommendation categories identified in review ‘Implications for research’ sections for two Cochrane Review Groups. A total of 22 categories were identified, 12 of which were common to both the Schizophrenia and Multiple Sclerosis (MS) Review Groups. The five most frequently occurring implications for research (marked in grey) were the same for both review groups.

Categories	Overall use of the category (both review groups, only categories common to both, 205 reviews) (n/%)	Use of the category in the Schizophrenia Review Group (162 reviews)	Use of the category in the MS Review Group (43 reviews)
Better outcome choice	118 (58%)	89 (55%)	29 (68%)
Better choice of future intervention and/or choice of comparator	111 (54%)	89 (55%)	22 (51%)
Longer follow-up	73 (36%)	55 (34%)	18 (42%)
Larger/sufficient sample size that has statistical power	71 (35%)	55 (34%)	16 (37%)
Use validated rating scales	39 (19%)	28 (17%)	11 (26%)
Ensure blinding (or mitigate lack of blinding where the nature of intervention makes blinding difficult)	27 (13%)	24 (15%)	3 (7%)
Better choice of eligibility criteria	19 (9%)	15 (9%)	4 (9%)
Use internationally agreed set of standardised outcomes	15 (7%)	12 (7%)	3 (7%)
Use real-world effectiveness/pragmatic design approach	10 (5%)	9 (6%)	1 (2%)
Ensure allocation is concealed	9 (4%)	7 (4%)	2 (5%)
Provide cost-effectiveness analysis	-	5 (3%)	2 (5%)
Better design	-	Not used	7 (16%)
Better methodology	-	Not used	4 (9%)
Better link to current clinical pathways	-	Not used	4 (9%)
Intention-to-treat analysis for all outcomes	-	14 (9%)	Not used
More head-to-head trials	-	Not used	3 (7%)
Do not use cross-over methodology	-	7 (4%)	Not used
More appropriate choice of trial setting	-	6 (4%)	Not used
More complete collection of participant data	-	3 (2%)	Not used
Predefine outcomes	-	3 (2%)	Not used
Include randomisation from a waiting list	-	3 (2%)	Not used

intervention/comparator, longer follow-up, larger sample size and use of validated scales). However, the ranking of each category in that top five list varied from one year to the next; no category was consistently most used (Figure 1 and Figure 2). There was no obvious pattern of improvement over time for trials included in systematic reviews published by both groups.

Conclusions

There is substantial overlap in the types of recommendation made in the ‘Implications for research’ sections of systematic reviews done by the Cochrane Schizophrenia and MS Review Groups. These were easier to identify in the Schizophrenia Review

Group’s reviews because of their consistent approach to presenting implications for research in accordance with published guidance⁶. Many of their reviews also routinely include a suggested design for a future trial in this section.

The five most frequently made recommendations are the same for both groups with better choice of outcomes being top of the list and used in over half the 205 reviews included in our study. Looking across the decade to 2019, there is no obvious pattern of decrease in the areas of methodology that need to be improved in trials included in systematic reviews published by both groups.

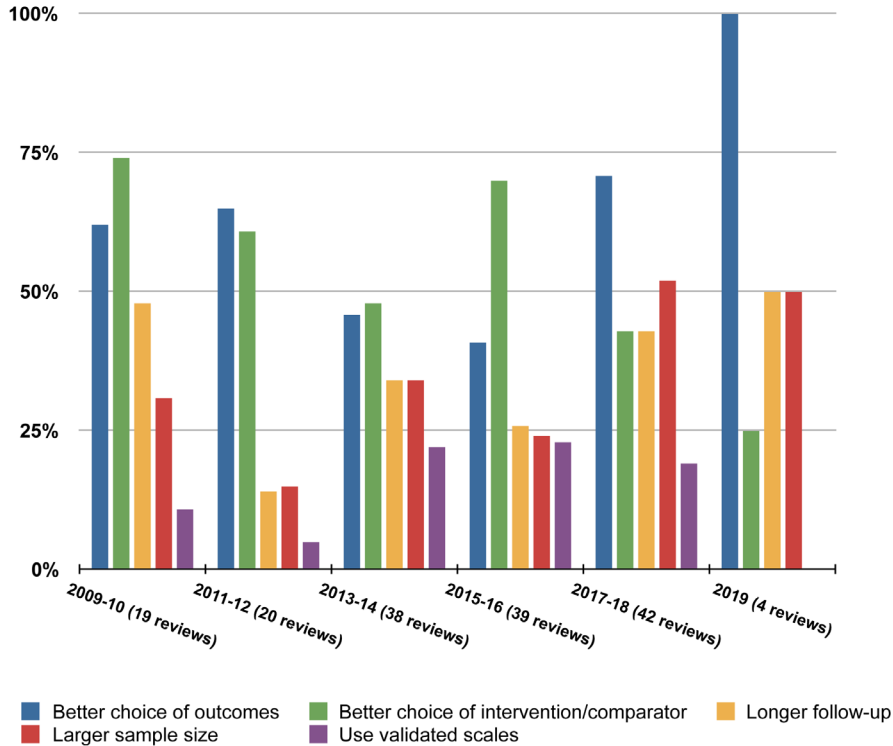


Figure 1. The five most used ‘Implications for research’ recommendations for the Cochrane Schizophrenia Review Group 2009–2019. The percentages shown are the average over a two-year period, apart from 2019 which was incomplete at the time the work was done. While some categories are used in well over half of all review published in a given year, there is no clear pattern over time.

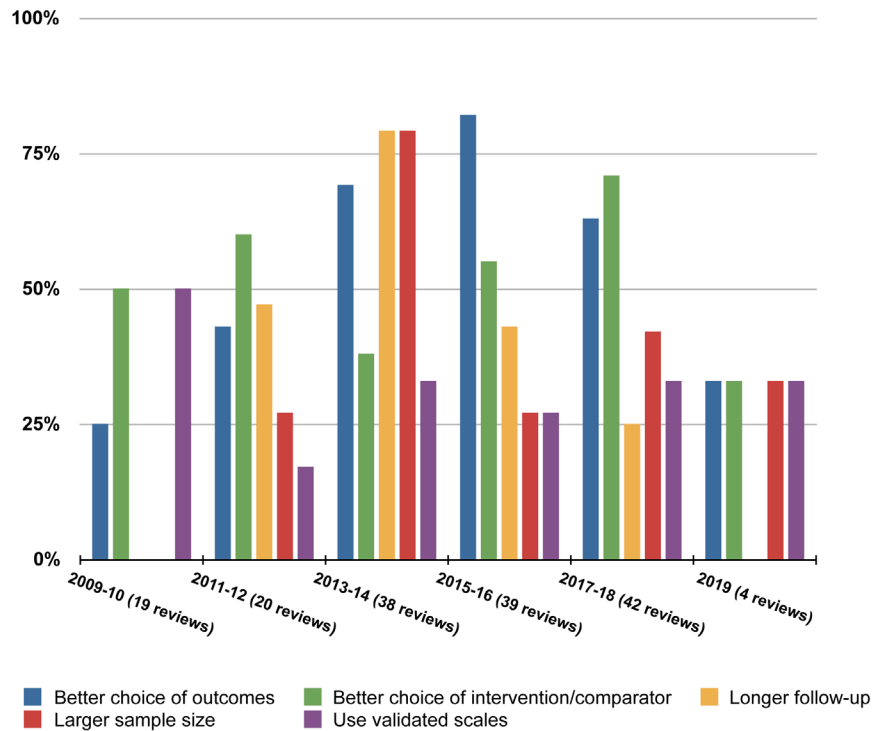


Figure 2. The five most used ‘Implications for research’ recommendations for the Cochrane MS Review Group 2009–2019. MS=Multiple sclerosis and rare diseases of the central nervous system. The percentages shown are the average over a two-year period, apart from 2019 which was incomplete at the time the work was done. While some categories are used in well over half of all reviews published in a given year, there is no clear pattern over time. Use of ‘Better choice of outcomes’ may be increasing but review numbers are small so caution is needed in interpreting this increase.

Previous research found that small, underpowered studies make up the entirety of evidence in most meta-analyses reported by Cochrane reviews⁷. Only 35% of the reviews in our sample mentioned sample size in the Implication for research, which suggests that reviewers may be underestimating the size of this problem. The most frequent issue raised in our work was the choice of outcome, a persistent problem that led to the COMET initiative to facilitate the development of core outcome sets^{3,8}. Cochrane reviewers need to continue to highlight trial design problems; indeed they could perhaps do so more often, especially around sample sizes.

Limitations

It is possible that different researchers would have categorised the implications for research differently, although we did use two reviewers working independently and there was little disagreement. Selecting other Cochrane review groups may have led us to conclude that different recommendations were most common but without doing that work we have no way of knowing this.

Despite looking at just two Cochrane review groups, we believe that our findings are likely to be generalisable to other areas of health and health care but, at a minimum, a good start for the 2020s would be for researchers planning trials in schizophrenia and multiple sclerosis to ask themselves (or be compelled to do so by others, e.g. funders) the following questions:

1. Why did we choose these outcomes?
2. Why is this the right intervention and comparator?
3. Is follow-up long enough?
4. Is the sample size large enough?
5. Will validated scales be used to measure outcomes?

There are resources that can help with the above. For example COMET (<http://www.comet-initiative.org>) can help guide outcome choice and PRECIS-2 can help with design decisions around comparators and follow-up⁹. Discussions with staff at

clinical trial units and other research support centres are also likely to improve designs. Change is slow at present and initiatives to encourage, or force, trialists to consider these questions would be welcome, especially from funders.

Data availability

Underlying data

Open Science Framework: Data extracted from the Implications for Research sections of reviews from two Cochrane Review Groups. <https://doi.org/10.17605/OSF.IO/XCJ7R>

This project contains the following underlying data:

- Data from Implications for Research section of Cochrane MS Review Group 2009–2019.csv (Data extracted from Cochrane MS Review Group)
- Data from Implications for Research section of Cochrane Schizophrenia Review Group 2009–2019.csv (Data extracted from Cochrane Schizophrenia Review Group)

Extended data

Open Science Framework: Data extraction form used in ‘What improvements do Cochrane systematic reviewers suggest for the design of trials?’. <https://doi.org/10.17605/OSF.IO/FNJVZ>

This project contains the following extended data:

- Data extraction form.csv (Study data extraction form)

Open Science Framework: Full data extraction from the Implications for Research sections of reviews from two Cochrane Review Groups. <https://doi.org/10.17605/OSF.IO/P9CW6>

This project contains the following data:

1. Data extraction MS 12-6-2020.csv
2. Data extraction Schizophrenia 12-6-2020.csv

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

References

1. Glasziou P, Chalmers IC: **Research waste is still a scandal—an essay by Paul Glasziou and Iain Chalmers**. *BMJ*. 2018; **363**: k4645. [PubMed Abstract](#) | [Publisher Full Text](#)
2. Yordanov Y, Dechartres A, Porcher R, et al.: **Avoidable waste of research related to inadequate methods in clinical trials**. *BMJ*. 2015; **350**: h809. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Ioannidis JP: **Clinical trials: what a waste**. *BMJ*. 2014; **349**: g7089. [PubMed Abstract](#) | [Publisher Full Text](#)
4. Higgins JPT, Thomas J, Chandler J, et al. (editors): **Cochrane Handbook for Systematic Reviews of Interventions version 6.0**. (updated July 2019). Cochrane. 2019. [Reference Source](#)
5. Clarke L, Clarke M, Clarke T: **How useful are Cochrane reviews in identifying research needs?** *J Health Serv Res Policy*. 2007; **12**(2): 101–3. [PubMed Abstract](#) | [Publisher Full Text](#)
6. Brown P, Brunnhuber K, Chalkidou K, et al.: **How to formulate research recommendations**. *BMJ*. 2006; **333**(7572): 804–6. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Turner RM, Bird SM, Higgins JPT: **The Impact of Study Size on Meta-analyses: Examination of Underpowered Studies in Cochrane Reviews**. *PLoS One*. 2013; **8**(3): e59202. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Williamson P, Clarke M: **The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: its role in improving Cochrane Reviews**. *Cochrane Database Syst Rev*. 2012; (5): ED000041. [PubMed Abstract](#) | [Publisher Full Text](#)
9. Loudon K, Treweek S, Sullivan F, et al.: **The PRECIS-2 tool: Designing trials that are fit for purpose**. *BMJ*. 2015; **350**: h2147. [PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 2

Reviewer Report 17 July 2020

<https://doi.org/10.5256/f1000research.27618.r65530>

© 2020 Puljak L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Livia Puljak 

Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia

Version 2 is great.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Methodological research, clinical epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 29 April 2020

<https://doi.org/10.5256/f1000research.24991.r62266>

© 2020 Gluud C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Christian Gluud 

The Copenhagen Trial Unit, Copenhagen University Hospital, Copenhagen, Denmark

Pirosca and colleagues have examined the 'Implications for research' section of 205 Cochrane reviews published between 2009 to 2019 from two editorial groups. The five most used categories were the same for both editorial groups: better choice of outcomes; better choice of

intervention/comparator; longer follow-up; larger sample size; use of validated scales. Better choice of outcomes and/or better choice of intervention/comparator was recommended in over half of reviews; longer follow-up and larger sample size were recommended in over a third; and use of validated scales being suggested in around a fifth of reviews.

I think the authors should also discuss their interesting findings from the point of view: how good are the Cochrane authors? E.g. we know from numerous studies (e.g., Rebecca M Turner *et al.*) that 80% to 90% of all Cochrane reviews do not reach a fair meta-analytic sample size providing the necessary power to assess medium large and small (but still clinically relevant) intervention effects. How come then that almost 66% of reviews do not mention larger sample sizes?

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Evidence based clinical practice.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 22 Jun 2020

Shaun P. Treweek, University of Aberdeen, Aberdeen, UK

**Response to comments on F1KR00CDE F1R-VER24991-A
22/6/2020**

Thanks for the reviewers' comments on our manuscript. Our responses are below. We have made the necessary changes to the manuscript using track changes so that they can be easily seen. We have also submitted a clean version with the changes accepted.

Reviewer #2

- I think the authors should also discuss their interesting findings from the point of view: how good are the Cochrane authors? E.g. we know from numerous studies (e.g., Rebecca M Turner.) that 80% to 90% et al of all Cochrane reviews do not reach a fair meta analytic sample size providing the necessary power to assess medium large and small (but still clinically relevant) intervention effect. How come then that almost 66% of reviews do not mention larger sample sizes?

We have added the following text to the 'Conclusions' section:

Previous research found that small, underpowered studies make up the entirety of evidence in most meta-analyses reported by Cochrane reviews⁷. Only 35% of the reviews in our sample mentioned sample size in the Implication for research, which suggests that reviewers may be underestimating the size of this problem. The most frequent issue raised in our work was the choice of outcome, a persistent problem that led to the COMET initiative to facilitate the development of core outcome sets^{5,8}. Cochrane reviewers need to continue to highlight trial design problems; indeed they could perhaps do so more often, especially around sample sizes.

Competing Interests: No competing interests were disclosed.

Reviewer Report 16 March 2020

<https://doi.org/10.5256/f1000research.24991.r61192>

© 2020 Puljak L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Livia Puljak 

Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia

I have reviewed with interest the manuscript titled "What improvements do Cochrane systematic reviewers suggest for the design of trials? [version 1; peer review: awaiting peer review]". The authors report a study based on a very elegant idea, analysing content of the "Implications for research" section in selected Cochrane reviews. I would like to suggest the following minor revisions:

Title: I would suggest revising in a way to reflect the fact that these recommendations came from Cochrane reviews. When I first read the title, I initially thought that this manuscript reports results of a survey among Cochrane authors.

Abstract

- "Reviews with citations between 2009 and 2019 were identified" – should this be "reviews published between 2009 and 2019"?

- “We created 22 categories” – I would suggest revising “categories” into “categories of recommendations”.
- Conclusion: I am not sure that it is sufficient to recommend to trialists to “ask themselves”. The problems that were mentioned (i.e. recommendations) were consistent over the analysed period. I think it would be more meaningful to conclude that we need to analyse the effects of interventions that will force/motivate trialists to change their actions when designing clinical trials.

Methods

- It would be useful to have a section called Data extraction, to describe all the data that were extracted. For example, in the beginning of the Results, the authors have mentioned type of interventions, and number of studies included in analysed reviews, but this was not mentioned as extracted in the Methods.

Results

- Results section is rather short. I would appreciate to read more text in this section, i.e. about categories specific only to each review group analyzed. However, I understand that this is brief report, so I would not insist on this, if this is not feasible for this type of article.
- In the Discussion, the authors mentioned that some categories were “clearer”, but I am not really sure what this means. I also do not see anything mentioned about that in the Methods, and Results.
- In the Abstract, the authors wrote “ There was no obvious pattern of improvement over time for trials included in systematic reviews published by both groups”, and in the Results, the authors wrote “However, the use of these five categories varied over time for each group”. Is this supposed to be the same? The expression “varied” is very non-specific.

Discussion

- I would prefer to see this explained in more detail; I am not sure I understand completely what it means: “These were clearer in the Schizophrenia Review Group’s reviews because of their structured approach to presenting implications for research in accordance with published guidance”. What kind of “structured approach” is this?
- Is this a formal requirement, is there a reference to be used for this sentence: “Their reviews also routinely include a suggested design for a future trial in this section.”
- “cited in over half” – I would suggest to revise “mentioned (or used) in over half”.
- Conclusion statement at the end of the manuscript: the same comment as for the conclusion statement at the end of the Abstract.
- Perhaps the authors should mention some other resources that the trialists should use to improve some of these aspects, such as core outcome sets.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Methodological research, clinical epidemiology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 22 Jun 2020

Shaun P. Treweek, University of Aberdeen, Aberdeen, UK

**Response to comments on F1KR00CDE F1R-VER24991-A
22/6/2020**

Thanks for the reviewers' comments on our manuscript. Our responses are below. We have made the necessary changes to the manuscript using track changes so that they can be easily seen. We have also submitted a clean version with the changes accepted.

Reviewer #1

- Title: I would suggest revising in a way to reflect the fact that these recommendations came from Cochrane reviews. When I first read the title, I initially thought that this manuscript reports results of a survey among Cochrane authors.

We have changed the title to "Learning from Cochrane systematic reviews: what improvements do these suggest for the design of trials?"

- "Reviews with citations between 2009 and 2019 were identified" – should this be "reviews published between 2009 and 2019"?

In the first draft of the report we had 'published' but we have decided to change to 'citations' because of the way Cochrane makes its reviews available. Through 'citation' we mean that we examined the most recent version of any review that was first published or was updated in 2009 to 2019. This is important because reviews from before 2009 were still "published" during the decade, but were not updated in that time period.

- "We created 22 categories" – I would suggest revising "categories" into "categories of recommendations".

Done.

- Conclusion: I am not sure that it is sufficient to recommend to trialists to “ask themselves”. The problems that were mentioned (i.e. recommendations) were consistent over the analysed period. I think it would be more meaningful to conclude that we need to analyse the effects of interventions that will force/motivate trialists to change their actions when designing clinical trials.

We have added a few sentences including the available resources for researchers to change their actions in the *‘Conclusions’* section:

There are resources that can help with the above. For example, COMET (<http://www.comet-initiative.org>) can help with outcome choice and PRECIS-2 can help with design decisions around comparators and follow-up⁹. Discussions with staff at clinical trial units and other research support centres are also likely to improve designs. Change is slow at present and initiatives to encourage, or force, trialists to consider these questions would be welcome, especially from funders.

In the abstract, we have also added a few words to address this comment (addition in italics, which is also included in the main Conclusions):

We suggest that trialists working in these and other areas ask themselves, *or are compelled to do so by others (e.g. funders),...*

- It would be useful to have a section called Data extraction, to describe all the data that were extracted. For example, in the beginning of the Results, the authors have mentioned type of interventions, and number of studies included in analysed reviews, but this was not mentioned as extracted in the Methods.

We have added a new file in the repository including all the data we extracted so that readers can see the data we extracted. We have added the following text to the *‘Methods’* section:

Additionally we extracted information on the number of included participants and studies, the risk of bias etc. (See Data Availability for a file containing the full list).

- Results section is rather short. I would appreciate to read more text in this section, i.e. about categories specific only to each review group analyzed. However, I understand that this is brief report, so I would not insist on this, if this is not feasible for this type of article.

As this is a Brief Report, we have decided to keep the *‘Results’* section to the point and we have not repeated in words information that is in the table. We have added some extra detail though, see response to Comment #8.

- In the Discussion, the authors mentioned that some categories were “clearer”, but I am not really sure what this means. I also do not see anything mentioned about that in the Methods, and Results.

We have replaced this text with:

However, the ranking of each category in that top five list varied from one year to the next; no category was consistently most used (Figures 1 and 2). There was no obvious pattern of improvement over time for trials included in systematic reviews published by both groups.

- In the Abstract, the authors wrote “ There was no obvious pattern of improvement over time for trials included in systematic reviews published by both groups”, and in

the Results, the authors wrote “However, the use of these five categories varied over time for each group”. Is this supposed to be the same? The expression “varied” is very non-specific.

The two sentences refer to different things: the one in the abstract to general improvement over time, the one in Results to the position of a given category in the top five from one year to the next. We have modified the text, see response to Comment #7.

- I would prefer to see this explained in more detail; I am not sure I understand completely what it means: “These were clearer in the Schizophrenia Review Group’s reviews because of their structured approach to presenting implications for research in accordance with published guidance”. What kind of “structured approach” is this?

Consistency was the key advantage and we have changed the wording of this sentence accordingly:

These were easier to identify in the Schizophrenia Review Group’s reviews because of their consistent approach to presenting implications for research in accordance with published guidance⁶

- Is this a formal requirement, is there a reference to be used for this sentence: “Their reviews also routinely include a suggested design for a future trial in this section.”

We are not aware of this being a formal requirement, or of a reference we can point readers to. We have modified the sentence so that it now comes across as an observation rather than review group policy:

Many of their reviews also include a suggested design for a future trial in this section.

- “cited in over half” – I would suggest to revise “mentioned (or used) in over half”.

We have changed to ‘used in over half’.

- Conclusion statement at the end of the manuscript: the same comment as for the conclusion statement at the end of the Abstract.

We have addressed this issue. See response to comment 4.

- Perhaps the authors should mention some other resources that the trialists should use to improve some of these aspects, such as core outcome sets.

See response to Comment #4.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research