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](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

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>What’s in a name? The challenge of describing interventions in systematic reviews: analysis of a random sample of reviews of non-pharmacological stroke interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Hoffmann, Tammy; Walker, Marion; Langhorne, Peter; Eames, Sally; Thomas, Emma; Glassziou, Paul</td>
</tr>
</tbody>
</table>

VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Lisa Douet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wessex Institute, University of Southampton, UK</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>22-Jul-2015</td>
</tr>
</tbody>
</table>

GENERAL COMMENTS

The authors have chosen to use the TIDieR check list in their study, it would be helpful to have further clarification around why they chose this check list, as it is primarily aimed at supporting the description of intervention in primary research evaluate intervention studies as opposed to systematic reviews.

Was the control arm of studies rated - this becomes most relevant if the studies included in the review are pragmatic and compare active treatments as opposed to placebo.

I note that disagreements between raters was discussed and resolved (if necessary a third author was involved) - it would be interesting to know the level of disagreements between the raters.

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Nicola McCleary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University of Aberdeen, Scotland, UK</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>28-Aug-2015</td>
</tr>
</tbody>
</table>

GENERAL COMMENTS

Thank you for the opportunity to review this manuscript, which I read with great interest. This manuscript provides an interesting overview of the completeness of reporting of non-pharmacological stroke interventions in a sample of systematic reviews. It highlights the intervention components that are frequently not specified adequately, thereby illuminating areas where reporting practice could be improved in the future. The manuscript rightly notes that improvements require the action and collaboration of all those involved in the publishing process, including researchers, peer reviewers, and journal editors. Suggestions for improvement are included below.

1. The authors note in the discussion section on page 19 that ‘This study highlights an even greater problem in systematic reviews, where the poor reporting of interventions in trials is compounded.'
This reinforces the importance of trialists providing sufficient intervention details.' However, this study focussed on reporting of interventions in systematic reviews, and did not consider reporting in the original trials themselves. Because reporting in the original trials was not reviewed, it is unclear the extent to which the reporting issues highlighted can be attributed to the reviews or to the original trials. Indeed, in paragraph 2 on page 16 of the discussion, the authors discuss previous work that found issues with intervention reporting in subsequent reviews even when details were present in the original trials. Taken together, these results would seem to highlight that improvements in intervention reporting are the responsibility of both trialists and reviewers. The manuscript would benefit from clarification of this, and it would also be useful to include as a study limitation the fact that reporting in the original trials was not assessed.

2. On page 5 at the beginning of the methods section, the manuscript outlines that the study included 30 Cochrane systematic reviews and 30 non-Cochrane systematic reviews. The manuscript would benefit from a rationale for arriving at this specific number of reviews. For example, was it because the goal was to highlight key reporting issues, rather than to be exhaustive, in line with previous research into reporting of non-pharmacological interventions (Boutron I, Tubach F, Giraudeau B, Ravaud P: Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. JAMA 2003, 290:1062–1070)?

3. Page 7 outlines how the completeness of intervention reporting was assessed. This section would benefit from further detail to facilitate study evaluation and replication. For example, it is not currently clear how interventions were scored according to the TIDieR criteria: I assume after reading the data analysis section that each intervention was scored for each criterion as either adequately reported, inadequately reported, or not at all reported? It may also be informative to provide details of the coding disagreements that did occur: for example, were there many, did they relate to any of the TIDieR items in particular?

4. Linked to point 3, this section would also benefit from a clear summary of the TIDieR criteria. On page 7 line 16, readers are referred to Figure 1, which appears later in the paper and includes only the labels of the criteria, not their descriptions. The inclusion of a box summarising the criteria labels and descriptions may assist readers with interpreting the meaning of the criteria. This may be especially important given that the authors presumably want to encourage uptake of TIDieR for investigating reporting practices and specifying interventions in both trials and reviews.

5. On page 6 in the methods section it is stated that 'The full text of the selected reviews was examined to identify reviews that met the inclusion criteria'. Can the authors give more details about this process? Was this also performed by the two independent coders? If not, the authors may want to highlight this as a study limitation.

6. In the ‘search strategy and selection of reports of trials’ section in the methods, it may be beneficial to move the final paragraph, which clarifies the types of reviews that were included, to the beginning of this section. This may help the reader follow the search strategy development and the search process, since the reasons for inclusion
of certain search terms are clarified beforehand.

7. There does not seem to have been any quality assessment of the systematic reviews included in this study, and so the overall quality of the reviews included remains unclear. This is an important limitation that should be discussed, since it is not clear whether or not high quality reviews in this area are lacking in intervention detail.

8. On page 8, the first paragraph in the results section states that interventions were coded into broad categories according to a relevant guideline: this would be better placed in the methods section, along with details of the coding process.

9. The authors also note in this section that two of the 30 Cochrane reviews were excluded, as no interventions reviewed were eligible for inclusion. Can the authors provide more detail with regard to why this occurred, and why they did not ensure at the review sampling stage that all reviews included eligible interventions? Considering that some of the reviews included many interventions, this may have resulted in data relating to a considerable number of interventions being lost, which is an important limitation to highlight.

10. In the ‘completeness of intervention descriptions in systematic reviews’ section in the results, the authors note that intervention fidelity, materials, procedures, and tailoring were the least adequately reported TIDieR items. However, upon reviewing Figure 1, it appears that modifications were the least well reported item. This should be included amongst these key results, and also in the abstract and discussion.

11. In the final paragraph on page 16 of the discussion section, the authors note that ‘Elements that are easier and more succinct to report, such as the ‘when and how much’ of the intervention and its mode of delivery, were better reported.’ The authors could elaborate on this by discussing the correspondence between these results and the wider literature on behavioural interventions in particular, which notes that there is often a greater focus on reporting intervention objectives, settings, methods of administration, and providers (Davidson KW, Goldstein M, Kaplan RM, Kaufmann PG, Knatterud GL, Orleans CT, Spring B, Trudeau KJ, Whitlock EP: Evidence-based behavioural medicine: what is it and how do we achieve it? Ann Behav Med 2003, 26:161–171. Abraham C, Michie S: A taxonomy of behavior change techniques used in interventions. Health Psychol 2008, 27:379–387).

12. As mentioned in various previous comments, the manuscript would benefit from elaboration of the study limitations in the discussion section, including that reporting in the original trials was not assessed, that two Cochrane reviews had to be excluded, and that included reviews were not quality-assessed. This section should also highlight that it is unclear the extent to which the random sample reflects systematic reviews in this particular area.

13. The design section of the abstract on page 2 should specify that the study also involved an online survey of review authors.


15. Page 3, line 39 would benefit from re-wording, such as ‘our sample contained only reviews of non-pharmacological and non-
16. On page 3, there are full stops for some bullet points, but not others.

17. Page 5, line 53: I assume that the two independent screeners were SE and ET, who are specified later in the methods section as conducting key tasks, but please specify here.

18. Page 8 line 45: ‘non-Cochrane reviews’ should be singular.

19. Page 11, Figure 2 title: please replace ‘and’ with ‘or’.

REVIEWER
Andrew Cook
Wessex Institute
University of Southampton
England
I’m currently involved with on a similar, but non overlapping, project.

REVIEW RETURNED
01-Sep-2015

GENERAL COMMENTS
Ethics: In my institution I’d consult with my ethics committee chair before approaching a set of review authors with these kinds of questions. I’d expect the chair would confirm no formal ethics approval would be required, but I would ask. No interaction with an ethics committee is reported here. Was action along these lines considered? (Editor: I’ve checked “no” above, if I had the option I’d have checked "don't know" - I don't think it's strictly necessary, and would come down to local institutional policies, and BMJ Open Policies)

Why was the sample size of 60 selected? And within that 60, why 30 Cochrane and 30 non-Cochrane? (I'm not querying why Cochrane reviews were examined as a special case, this is set out - rather why these particular counts?)

I didn’t find a justification for collapsing the ‘adequate for all’ and ‘adequate for some’ categories when comparing cochrane and non-cochrane reviews. This should be provided.

In figure 2, why use a line chart? The items on the x axis are not directly related to each other. The ordering is arbitrarily determined by the TIDieR check list. I’d prefer a format which didn’t imply any particular relationship, such as some form of column chart.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

1. The authors have chosen to use the TIDieR check list in their study, it would be helpful to have further clarification around why they chose this check list, as it is primarily aimed at supporting the description of intervention in primary research evaluate intervention studies as opposed to systematic reviews.

Authors’ response: While the TIDieR checklist is primarily designed to guide the reporting of intervention in primary studies, in the TIDieR explanation and elaboration paper (BMJ 2014;324:g1687) it does note that TIDieR’s use in systematic reviews is also appropriate (Discussion, paragraph 3). As the focus of this study was the reporting of interventions, the components contained...
in the TIDieR checklist are appropriate, regardless of whether the interventions are reported in a randomized trial or in a systematic review. There are no other appropriate or more comprehensive intervention reporting checklists that could have been used in place of TIDieR for this sample of studies.

2. Was the control arm of studies rated - this becomes most relevant if the studies included in the review are pragmatic and compare active treatments as opposed to placebo.

Authors’ response: No, the completeness of reporting of the control interventions was not rated in this study. We agree that is an important issue and can affect interpretation of effect sizes of the intervention. It certainly would be appropriate to perform such ratings in future studies of intervention reporting, in both randomised trials and systematic reviews.

3. Comment: I note that disagreements between raters was discussed and resolved (if necessary a third author was involved) - it would be interesting to know the level of disagreements between the raters.

Authors’ response: We did not formally measure the level of agreement between the raters primarily because the raters met and discussed their ratings after rating each block of 5 reviews (as described in the paper). As the process had an iterative component to it, an overall measure of agreement would not be particularly useful.

Reviewer 2
Comment: Thank you for the opportunity to review this manuscript, which I read with great interest. This manuscript provides an interesting overview of the completeness of reporting of non-pharmacological stroke interventions in a sample of systematic reviews. It highlights the intervention components that are frequently not specified adequately, thereby illuminating areas where reporting practice could be improved in the future. The manuscript rightly notes that improvements require the action and collaboration of all those involved in the publishing process, including researchers, peer reviewers, and journal editors.

Authors’ response: Thank you for your comments about the value of this paper and for your detailed and constructive review comments.

1. The authors note in the discussion section on page 19 that ‘This study highlights an even greater problem in systematic reviews, where the poor reporting of interventions in trials is compounded. This reinforces the importance of trialists providing sufficient intervention details.’ However, this study focussed on reporting of interventions in systematic reviews, and did not consider reporting in the original trials themselves. Because reporting in the original trials was not reviewed, it is unclear the extent to which the reporting issues highlighted can be attributed to the reviews or to the original trials. Indeed, in paragraph 2 on page 16 of the discussion, the authors discuss previous work that found issues with intervention reporting in subsequent reviews even when details were present in the original trials. Taken together, these results would seem to highlight that improvements in intervention reporting are the responsibility of both trialists and reviewers. The manuscript would benefit from clarification of this, and it would also be useful to include as a study limitation the fact that reporting in the original trials was not assessed.

Authors’ response: A number of studies have analysed the reporting of interventions in trials and concluded that they are incompletely reported and this problem is no longer in doubt. Whereas, other than the current study, only 1 study (in the specific area of acupuncture) has examined intervention reporting in systematic reviews. Hence, the primary aim of this study was to examine the completeness of intervention reporting in systematic reviews, rather than to analyse the causes of this problem. Although undoubtedly this problem is contributed to by the incomplete intervention reporting in trials (as is mentioned in the Discussion), we did not seek to analyse this particular contributor,
there are likely many other contributors (as our author survey suggests). Exploration of the multiple
reasons for the poor reporting of interventions in systematic reviews is needed in future research. We
have clarified in the Discussion (last para) that both trialists and systematic reviewers share a
responsibility for ensuring that interventions are adequately reported in reviews.

2. On page 5 at the beginning of the methods section, the manuscript outlines that the study included
30 Cochrane systematic reviews and 30 non-Cochrane systematic reviews. The manuscript would
benefit from a rationale for arriving at this specific number of reviews. For example, was it because
the goal was to highlight key reporting issues, rather than to be exhaustive, in line with previous
research into reporting of non-pharmacological interventions (Boutron I, Tubach F, Giraudiere B,
Ravaud P: Methodological differences in clinical trials evaluating nonpharmacological and
pharmacological treatments of hip and knee osteoarthritis. JAMA 2003, 290:1062–1070)?

Authors’ response: Yes, the goal of this study was to highlight key reporting issues regarding the
reporting of interventions in systematic reviews as there has been almost no previous research
conducted in this area. Our intention was not to conduct an exhaustive study such as the study
conducted by Boutron et al. Though we did not do a formal sample size calculation, we were aware
that 60 systematic reviews should have at least 500 studies (Cochrane reviews have a median of 8
trials; non-Cochrane a median of 23*). Hence estimates of proportions would have 95% confidence
intervals of better than +/-7% which we felt sufficient for the descriptive purposes of this study. We
have added this into the Methods section in the ‘search strategy and selection of reports of trials’
subsection.

* Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics

3. Page 7 outlines how the completeness of intervention reporting was assessed. This section would
benefit from further detail to facilitate study evaluation and replication. For example, it is not currently
clear how interventions were scored according to the TIDieR criteria: I assume after reading the data
analysis section that each intervention was scored for each criterion as either adequately reported,
inadequately reported, or not at all reported? It may also be informative to provide details of the
coding disagreements that did occur: for example, were there many, did they relate to any of the
TIDieR items in particular?

Authors’ response: We have added further detail to this section of the Methods (‘rating of intervention
reporting in each systematic review’) and a table that describes each of the items (Table 1). See
response to reviewer’s 1 comment 3 about coding disagreements.

4. Linked to point 3, this section would also benefit from a clear summary of the TIDieR criteria. On
page 7 line 16, readers are referred to Figure 1, which appears later in the paper and includes only
the labels of the criteria, not their descriptions. The inclusion of a box summarising the criteria labels
and descriptions may assist readers with interpreting the meaning of the criteria. This may be
especially important given that the authors presumably want to encourage uptake of TIDieR for
investigating reporting practices and specifying interventions in both trials and reviews.

Authors’ response: We have added a Table (Table 1) with the descriptions of the TIDieR items.

5. On page 6 in the methods section it is stated that ‘The full text of the selected reviews was
examined to identify reviews that met the inclusion criteria’. Can the authors give more details about
this process? Was this also performed by the two independent coders? If not, the authors may want
to highlight this as a study limitation.
Authors’ response: Further detail about this process has been added to this section.

6. In the ‘search strategy and selection of reports of trials’ section in the methods, it may be beneficial to move the final paragraph, which clarifies the types of reviews that were included, to the beginning of this section. This may help the reader follow the search strategy development and the search process, since the reasons for inclusion of certain search terms are clarified beforehand.

Authors’ response: We have made this change.

7. There does not seem to have been any quality assessment of the systematic reviews included in this study, and so the overall quality of the reviews included remains unclear. This is an important limitation that should be discussed, since it is not clear whether or not high quality reviews in this area are lacking in intervention detail.

Authors’ response: The aim of this study was to assess the completeness of intervention reporting in systematic reviews, not to explore whether the completeness of intervention reporting was related to variables such as the overall quality (ie risk of bias) of the review and hence this is not a limitation of this study. As analysis of intervention reporting in systematic reviews has had almost no attention, this study aimed to generally explore this issue as a starting point. Further analysis such as that suggested by the reviewer would be appropriate to conduct in future research.

8. On page 8, the first paragraph in the results section states that interventions were coded into broad categories according to a relevant guideline: this would be better placed in the methods section, along with details of the coding process.

Authors’ response: We have made this change as suggested.

9. The authors also note in this section that two of the 30 Cochrane reviews were excluded, as no interventions reviewed were eligible for inclusion. Can the authors provide more detail with regard to why this occurred, and why they did not ensure at the review sampling stage that all reviews included eligible interventions? Considering that some of the reviews included many interventions, this may have resulted in data relating to a considerable number of interventions being lost, which is an important limitation to highlight.

Author’s response: Screening of reviews as to whether they met the inclusion criteria occurred at the level of the review question/criteria, not at the level of the individual trials that were included in the reviews. This is why two reviews met the inclusion criteria for the study, but upon subsequent analysis, did not contain trials/intervention which could be included in this study. Even though no intervention reporting data was included from these two reviews, our sample size remained more than adequate (as per response to Comment 2).

10. In the ‘completeness of intervention descriptions in systematic reviews’ section in the results, the authors note that intervention fidelity, materials, procedures, and tailoring were the least adequately reported TIDieR items. However, upon reviewing Figure 1, it appears that modifications were the least well reported item. This should be included amongst these key results, and also in the abstract and discussion.

Authors’ response: Thank you for noting this. We have made this change in the body of the text and the abstract.

11. In the final paragraph on page 16 of the discussion section, the authors note that ‘Elements that are easier and more succinct to report, such as the ‘when and how much’ of the intervention and its mode of delivery, were better reported.’ The authors could elaborate on this by discussing the correspondence between these results and the wider literature on behavioural interventions in particular, which notes that there is often a greater focus on reporting intervention objectives, settings, methods of administration, and providers (Davidson KW, Goldstein M, Kaplan RM, Kaufmann PG,

Authors’ response: Thank you for suggesting the inclusion of these articles, which we are familiar with. However, after careful reconsideration we have not included them in the discussion for the following primary reasons. The paper by Davidson et al does not include an analysis of the reporting of interventions and their elements (but advocates for why better reporting of interventions and other aspects of trials are important for behavioural interventions). The paper by Abraham et al is focused on the rater agreement of a taxonomy of behavior change interventions and their sample was gathered from studies of interventions about HIV/AIDS risk. Additionally, as not all of the interventions in the current study were behaviour change interventions, we do not feel that drawing comparisons with these behaviour change papers (whose primary goal was not to analyse intervention reporting quality) and this study is appropriate. However, in the Discussion we have compared the findings of this study with a study of intervention elements in trials of non-pharmacological interventions, in terms of the best and worst reported items.

12. As mentioned in various previous comments, the manuscript would benefit from elaboration of the study limitations in the discussion section, including that reporting in the original trials was not assessed, that two Cochrane reviews had to be excluded, and that included reviews were not quality-assessed. This section should also highlight that it is unclear the extent to which the random sample reflects systematic reviews in this particular area.

Authors’ response: As explained in responses to above comments (1, 5, 7), we do not agree that these three suggestions are limitations of the study. We have though added to the limitation section a comment about not knowing the extent to which this random sample reflects systematic reviews in stroke (Discussion, para 2).

13. The design section of the abstract on page 2 should specify that the study also involved an online survey of review authors.

Authors’ response: This change has been made.


Authors’ response: Thank you for noting this - this change has been made.

15. Page 3, line 39 would benefit from re-wording, such as ‘our sample contained only reviews of non-pharmacological and non-surgical stroke interventions, limiting generalisability of the results’.

Authors’ response: Thank you for noting this - this change has been made.

16. On page 3, there are full stops for some bullet points, but not others.

Authors’ response: We have added full stops for each bullet point.

17. Page 5, line 53: I assume that the two independent screeners were SE and ET, who are specified later in the methods section as conducting key tasks, but please specify here.

Authors’ response: This has been specified here.

18. Page 8 line 45: ‘non-Cochrane reviews’ should be singular.

Authors’ response: Thank you for noting this - this change has been made.

19. Page 11, Figure 2 title: please replace ‘and’ with ‘or’.

Authors’ response: This change has been made.
Reviewer 3
1. In my institution I’d consult with my ethics committee chair before approaching a set of review authors with these kinds of questions. I’d expect the chair would confirm no formal ethics approval would be required, but I would ask. No interaction with an ethics committee is reported here. Was action along these lines considered? (Editor: I’ve checked “no” above, if I had the option I’d have checked “don't know” - I don’t think it’s strictly necessary, and would come down to local institutional policies, and BMJ Open Policies)
Authors’ response: We have previously conducted research that involves contacting corresponding authors of published papers (whose contact details are provided for the purpose of answering questions about their paper) and our local ethics committee does not require that ethics approval is obtained in order to approach corresponding authors.

2. Why was the sample size of 60 selected? And within that 60, why 30 Cochrane and 30 non-Cochrane? (I’m not querying why Cochrane reviews were examined as a special case, this is set out - rather why these particular counts?)
Authors’ response: please see response to Reviewer 2 (comment 2).

3. I didn't find a justification for collapsing the ‘adequate for all’ and ‘adequate for some’ categories when comparing Cochrane and non-Cochrane reviews. This should be provided.
Authors’ response: Presenting separate data for both ‘adequate for all’ and ‘adequate for some’ for both Cochrane and non-Cochrane reviews for all 11 items produced a figure that was complicated and hard to interpret, hence the collapse of these two categories. However, we are happy to present the two lots of data as separate Figures (e.g. Fig 2a and Fig 2b) if the editors would prefer this.

4. In figure 2, why use a line chart? The items on the x axis are not directly related to each other. The ordering is arbitrarily determined by the TIDieR check list. I'd prefer a format which didn't imply any particular relationship, such as some form of column chart.
Authors’ response: We have changed the chart to a column chart as suggested by the reviewer.

**GENERAL COMMENTS**

1. In relation to the authors’ response to my original comment 3 and reviewer 1’s original comment 3 regarding disagreements between raters: I understand and agree that a formal, quantitative measure of agreement would not be possible or useful in this context. However, it may still be useful to include a narrative description of any disagreements, for example if there were any items for which disagreements were common, and how this was resolved. If there were no problematic disagreements, the authors could simply state this.

2. In response to my original comment 7 (quality assessment of included reviews) and reviewer 1’s original comment 2 (rating control interventions), the authors have noted useful avenues for further research. The authors could consider adding these to the manuscript.

**VERSION 1 – AUTHOR RESPONSE**

**VERSION 2 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Nicola McCleary</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>University of Aberdeen</td>
</tr>
<tr>
<td>Scotland</td>
<td>Scotland</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>10-Oct-2015</td>
</tr>
</tbody>
</table>
VERSİON 2 – AUTHOR RESPONSE

Reviewer’s Comments and Authors’ Responses

1. In relation to the authors’ response to my original comment 3 and reviewer 1’s original comment 3 regarding disagreements between raters: I understand and agree that a formal, quantitative measure of agreement would not be possible or useful in this context. However, it may still be useful to include a narrative description of any disagreements, for example if there were any items for which disagreements were common, and how this was resolved. If there were no problematic disagreements, the authors could simply state this.
Response: We have added information about this to the second paragraph of the Discussion.

2. In response to my original comment 7 (quality assessment of included reviews) and reviewer 1’s original comment 2 (rating control interventions), the authors have noted useful avenues for further research. The authors could consider adding these to the manuscript.
Response: We have suggested these future research avenues in the second last paragraph of the Discussion.

3. Thank you for noting in the abstract and the results that ‘modifications’ were the least well reported item. Could the authors also add this to the relevant paragraph in the discussion?
Response: We have added this to the third paragraph of the Discussion.