

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)	Peer Review Comments on Drug Trials Submitted to Medical Journals Differ Depending on Sponsorship, Results, and Acceptance: A Retrospective Cohort Study
AUTHORS	van Lent, Marlies; IntHout, Joanna; Out, Henk Jan

VERSION 1 - REVIEW

REVIEWER	Wendy Lipworth University of Sydney Australia
REVIEW RETURNED	02-Mar-2015

GENERAL COMMENTS	<p>The article could be set up a bit more strongly: rather than simply stating which parameters have and have not been studied, it would be helpful to tell the reader why you thought it was important/interesting to study the effects on review of study outcomes, sponsorship and decision about acceptance (as opposed to focusing on other things).</p> <p>The discussion and conclusion could also be enriched by considering systemic changes that might be needed, rather than simply saying that authors need to be more careful.</p>
-------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

REVIEWER	Roberta W. Scherer Johns Hopkins Bloomberg School of Public Health
REVIEW RETURNED	24-Mar-2015

GENERAL COMMENTS	<p>This manuscript describes a study examining and classifying reviewer comments for manuscripts describing drug randomized clinical trials. The methodology is sound and the findings add to our understanding of the peer review system. I enjoyed reading it. The main concerns I have are related to the description of the methods used for classifying the reviewer comments and in presentation of the data (interestingly, items 6 and 8 in their classification scheme).</p> <p>1. The description of the way in which the raters classified the comments is a bit opaque. It is clear that 2 raters independently classified comments for 50 manuscripts. It appears that at least one rater classified comments for the remaining manuscripts. What is not clear is whether a second person reviewed the remaining manuscripts for items with less than 85% agreement. Lines 50-57 seems to say that these items were graded by a second rater, but it is not clear. This text should be reworded to make clear how the remaining manuscripts were classified.</p> <p>2. Please define what is meant by "mean percentage of comments"</p>
-------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>in the methods. I assume it is the percentage of manuscripts with a specific comment – but nowhere is this explicitly defined. It makes a difference in the interpretation is this is not the definition.</p> <p>3. There should be a table with the total number of manuscripts and the “mean percentage of comments” for each item. The current manuscript includes this information for all the subgroups, but not for all manuscripts in total. In addition, the abstract lists the ranges of the “mean percentage of comments” across the subgroups, but could include the overall value for all manuscripts.</p> <p>4. For Tables 3, 4, and 5, it would be helpful to have the total number of manuscripts by subgroup within the table (I understand this information is present in Table 2, but the succeeding tables should be self-explanatory).</p> <p>5. A limitation not noted is that this is a sample of manuscripts describing drug RCTs – may not be generalizable to other study design or other types of RCTs (although I suspect it might be).</p> <p>6. In the discussion, page 14, paragraph in lines 42-58 – could not another explanation be that the type of particular comments found more often in rejected manuscripts (irrelevant research question, lack of novelty, and poor experimental design) be the reason for rejection rather than the overall number of comments.</p> <p>7. Minor comment – should not the abstract use the term “industry” instead of “pharmaceutical” in line 7-8, page 5.</p> <p>8. Can you comment on the one unexpected finding – that accepted manuscript more often had comments about the background literature?</p>
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

REVIEWER	Victoria Neale Wayne State University Detroit, Michigan USA
REVIEW RETURNED	15-Apr-2015

GENERAL COMMENTS	<p>This is an excellent study and well written paper on an important topic. The methods and results are clearly presented. I have only a few suggestions for improvement:</p> <p>1) I did not find a statement about reviewer awareness of industry support or sponsorship? (did I miss it?). If they were aware of industry involvement, this may have influenced their critique (possibly making it more negative would be my hypothesis).</p> <p>2) The paper could be shortened a bit, as there seems to be some redundancy in the Results and Discussion.</p> <p>3) Table 1 could be omitted or combined with Table 3 or the Supplemental Table 1.</p>
-------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Wendy Lipworth

The article could be set up a bit more strongly: rather than simply stating which parameters have and have not been studied, it would be helpful to tell the reader why you thought it was important/interesting to study the effects on review of study outcomes, sponsorship and decision about acceptance (as opposed to focusing on other things).

In the last paragraph of the Introduction (p.5), we have more explicitly described why we chose to investigate the parameters that are reported in this study:

'Analysis of reviewer comments provides more insight into the shortcomings of drug trials that are submitted to medical journals, both from the perspective of the design and conduct of trials and the reporting quality in articles. It would be interesting to determine whether the occurrence of specific shortcomings in manuscripts is affected by sponsorship and results being either positive or negative. In the current study, we performed a descriptive content analysis of peer review comments made on manuscripts on drug trials submitted to eight medical journals to investigate the relationship between the content of comments and sponsorship, direction of results, and decision about acceptance, using a previously reported cohort.(13)'

The discussion and conclusion could also be enriched by considering systemic changes that might be needed, rather than simply saying that authors need to be more careful.

For this comment, we like to refer to our answer to point 3 of the editorial comments; we adjusted the last paragraph of the Discussion to respond to this comment.

Reviewer 2: Roberta W. Scherer

This manuscript describes a study examining and classifying reviewer comments for manuscripts describing drug randomized clinical trials. The methodology is sound and the findings add to our understanding of the peer review system. I enjoyed reading it. The main concerns I have are related to the description of the methods used for classifying the reviewer comments and in presentation of the data (interestingly, items 6 and 8 in their classification scheme).

1. The description of the way in which the raters classified the comments is a bit opaque. It is clear that 2 raters independently classified comments for 50 manuscripts. It appears that at least one rater classified comments for the remaining manuscripts. What is not clear is whether a second person reviewed the remaining manuscripts for items with less than 85% agreement. Lines 50-57 seems to say that these items were graded by a second rater, but it is not clear. This text should be reworded to make clear how the remaining manuscripts were classified.

We agree with the reviewer that this information should be more clearly reported. We have adjusted this in the Methods (p.7):

'The final version of the checklist (Supplemental Table 1) was then tested on reviewer comments for a random sample of 30 manuscripts, that were independently classified by the two raters. Assuming an inter-rater agreement of at least 80% for each type of comment with the final checklist, 30 manuscripts were sufficient to estimate the agreement with a precision (standard error) of at most 7%. If the inter-rater agreement during this test was considered sufficiently high, a single reviewer (MvL) could continue with the rating process. After the classification of reviewer comments for these 30 manuscripts, we calculated the percentage of agreement between raters for each type of comment in the checklist. Kappa statistics were not considered suitable as some types of comments were rarely scored and resulting kappa values would be inaccurate.(21) For these 30 manuscripts, classification discrepancies were resolved by consensus between raters if the agreement for a comment was <85%. For the other types of comments, the score assigned by the rater who subsequently classified all comments for the other manuscripts (MvL) was decisive. Overall, reviewer comments for 50 manuscripts were scored by two raters in this study, which was equivalent to 20% (50 of 246) of the total number of included manuscripts.'

2. Please define what is meant by "mean percentage of comments" in the methods. I assume it is the percentage of manuscripts with a specific comment – but nowhere is this explicitly defined. It makes a

difference in the interpretation is this is not the definition.

'Mean percentage of comments' as reported in the results tables refers to the percentage of reviewers that will comment on a manuscript, depending on sponsor type, results, or decision about acceptance, estimated by the statistical model described in the Methods. This was explicitly stated in the Statistical analysis section (p.8):

'The model estimates the percentage of reviewers that will comment on a manuscript, depending on sponsor type, results, or decision about acceptance ("mean percentage of comments on a manuscript").'

3. There should be a table with the total number of manuscripts and the "mean percentage of comments" for each item. The current manuscript includes this information for all the subgroups, but not for all manuscripts in total. In addition, the abstract lists the ranges of the "mean percentage of comments" across the subgroups, but could include the overall value for all manuscripts.

As described above, the percentages reported for the subgroups represent the model estimates of the percentage of reviewers that will comment on a manuscript, depending on sponsor type, results, or decision about acceptance. By using this approach, individual reviewer reports could be included in the analysis (N=575) and the identification number that manuscripts received from a journal was used as cluster variable. If we would have calculated the percentage of manuscripts with a specific comment and would have compared this by subgroup, we would have lost a lot of data as we would look at manuscript level instead of at reviewer report level, and consequently our N would be decreased to N=246 manuscripts. For that reason, no table with the total number of manuscripts and the percentage of manuscripts with a specific comment was not reported in this article, and instead a range of point estimators is provided in the Results section.

4. For Tables 3, 4, and 5, it would be helpful to have the total number of manuscripts by subgroup within the table (I understand this information is present in Table 2, but the succeeding tables should be self-explanatory).

These numbers were added to the tables as suggested by the reviewer.

5. A limitation not noted is that this is a sample of manuscripts describing drug RCTs – may not be generalizable to other study design or other types of RCTs (although I suspect it might be).

This limitation was added to the Discussion (p.15):

'In addition, we included a sample of manuscripts describing drug RCTs and our results may therefore not be generalizable to other study designs or RCTs with other interventions.'

6. In the discussion, page 14, paragraph in lines 42-58 – could not another explanation be that the type of particular comments found more often in rejected manuscripts (irrelevant research question, lack of novelty, and poor experimental design) be the reason for rejection rather than the overall number of comments.

We agree with the reviewer that these specific comments may be the reason for rejection, but we cannot conclude this based on findings of the current study. There may be many reasons why papers get rejected beyond what is in the peer reviewers' report of the initial version of a manuscript. This is addressed in the Discussion on p.14, last paragraph.

7. Minor comment – should not the abstract use the term "industry" instead of "pharmaceutical" in line

7-8, page 5.

We could not find the term “pharmaceutical” in the abstract.

8. Can you comment on the one unexpected finding – that accepted manuscript more often had comments about the background literature?

This was indeed an unexpected finding. One explanation may be that the quality of manuscripts that were eventually rejected was too low for reviewers to comment on interpretation of the meaning of study results and comparison of results to previous literature, as they focused in their review on major flaws such as poor experimental design and/or irrelevance of the research question. For accepted manuscripts, overall quality may generally be higher and reviewers may focus more on the discussion of results. Inadequate discussion of the meaning of results and insufficient reference to the literature represent fixable errors that may not be a reason to reject manuscripts.

Reviewer 3: Victoria Neale

This is an excellent study and well written paper on an important topic. The methods and results are clearly presented. I have only a few suggestions for improvement:

1) I did not find a statement about reviewer awareness of industry support or sponsorship? (did I miss it?). If they were aware of industry involvement, this may have influenced their critique (possibly making it more negative would be my hypothesis).

At the included journals, reviewers were aware of the sponsorship of manuscripts as this has to be transparently reported by authors upon manuscript submission, and reviewers were not blinded to this information. In addition, manuscripts need to include a trial registration number and sponsorship can therefore also be checked in the trial registry by reviewers. It is possible that disclosure of funding and conflicts of interest may have influenced their critique, but we cannot conclude this based on the current study. However, reviewers should have access to this information to be able to estimate the validity of trial results and the influence of any potential bias introduced by the funding source. We added a sentence to the Methods section stating that reviewers were aware of the sponsorship of a study (p.6, last paragraph):

‘Information on sponsorship and the direction of results was previously extracted from manuscripts and classified according to predefined criteria.(14) Reviewers were aware of the sponsorship of trials.’

2) The paper could be shortened a bit, as there seems to be some redundancy in the Results and Discussion.

We reported the findings of our study and the interpretation of the meaning of the results as concise as possible.

3) Table 1 could be omitted or combined with Table 3 or the Supplemental Table 1.

As suggested by the reviewer, Table 1 has been removed from the main text of the manuscript and is now provided as a supplemental table.

VERSION 2 – REVIEW

REVIEWER	Roberta W. Scherer Johns Hopkins Bloomberg School of Public Health
REVIEW RETURNED	20-May-2015

GENERAL COMMENTS	Very nice paper - all my concerns have been addressed.
-------------------------	--------------------------------------------------------