Author’s response to reviews

Title: The effects of industry funding and positive outcomes in the interpretation of clinical trial results: a randomized trial among Dutch psychiatrists

Authors:

Joeri Tijdink (j.tijdink@amsterdamumc.nl)
Yvo Smulders (y.smulders@vumc.nl)
Lex Bouter (lm.bouter@vu.nl)
Christiaan Vinkers (c.h.vinkers@umcutrecht.nl)

Version: 1 Date: 07 Sep 2018

Author’s response to reviews:

Dear Editor,

On behalf of all authors, we would like to thank you and the reviewers for constructive comments to our manuscript entitled ‘How does industry funding disclosure influence psychiatrists? A randomized trial among Dutch psychiatrists’. We think that this input has enabled us to considerably improve the quality and readability of the manuscript. We have addressed all comments and suggestions in a point-to-point fashion in this rebuttal letter and we have incorporated all suggestions in the revised manuscript. As instructed, we have created a revised manuscript using track changes as well as a clean revised manuscript file. Note that in the answers to the reviewers, we refer to page numbers in the manuscript with the tracked changes.

Thank you for considering our rebuttal and our submission for publication in BMC Medical Ethics. We eagerly anticipate your decision and hope for acceptance. If any questions arise or if anything is unclear, don’t hesitate to contact me. I will be more than happy to provide extra information.

On behalf of all authors,

Best wishes,
For a formatted version of this rebuttal letter, we would kindly refer to the attached Cover Letter and Rebuttal letter in our submission.

Editor Comments:

The methodological and criticisms of the analysis from referee 3 indicate that your study is potentially flawed. We are offering you a chance to revise and respond to the reviewers' comments, but please be advised that the revised submission will be sent back to the referees for comment.

Reviewer reports:

Gwen Adshead (Reviewer 1):

Thank you for letting me see this interesting paper. I found it to be well constructed and well written; the methodology, aims and outcomes were clear. In my view the authors made a persuasive argument for their findings based on their study. I am not competent to assess the statistics; so my only comment may be misplaced BUT I thought the completion rate of 45% seemed low to me; and I wondered what difference that might have made. I would have welcomes some comment from the authors on why so few responded and whether the responders were spread equally across all groups.

Answer:

We thank the reviewer for the insightful comments. This is indeed an important limitation, and we have now addressed this in more detail in the revised manuscript. Even though a completion rate of 45% can be considered relatively high compared to other (randomized) surveys, it may also cause response bias. To check whether our findings can be generalized towards the Dutch population of psychiatrists, we have compared our sample with other samples from other studies conducted among Dutch psychiatrists and find that the gender ratio in our sample is comparable with the gender ratio of another study among Dutch psychiatrists.

We address these issues in the results section on page xx and in the discussion section on page xx.

We now write (see discussion section, page 13):
We did not survey non-respondents for the reasons not to participate. Psychiatrists may lack time and energy to engage in online surveys, or it may be related to the subject of the survey. Possibly, some psychiatrists might have been reluctant to participate as they might feel that they are not competent enough to judge a scientific abstract on its quality. Such bias could have affected the primary outcomes.

Moreover, in light of the response rate of 45%, we checked whether the gender of our study sample was comparable to the total population of Dutch psychiatrists. With 54% of the survey sample being male compared to 59% in another recent study of Dutch psychiatrists, (1), there is at least no reason to assume a large bias in the gender of participating psychiatrists.

I was also not completely convinced that the responders did not recognise the influence of sponsorship; is it not possible that they believed that they could account for that potential bias. The complexity of this debate seems to rest on whether it is possible for clinicians to 'handle' bias even if it is not fully conscious; and that there may be compensatory belief systems that are acquired through professional life and practice that mitigate against the industry effect. Worth a comment?

Answer: Thank you for this suggestion. We have addressed this issue in the Discussion.

We now write in the discussion section (see page 14):

Finally, the question remains if the responders did recognize the presence of the disclosure of industry funding. Some respondents may not have noticed the disclosure of industry funding in the scientific abstract, even though all seven psychiatrists who pretested the fictitious abstract noticed it. Others may have believed that they were able to account for this potential bias and thus have responded to the research findings with this sponsorship bias taken into account.

Wendy Lipworth (Reviewer 2):

This study explores the effects of industry funding disclosures on psychiatrists' appraisals of research studies in terms of their perceived credibility, clinical relevance, quality and interest.

The authors found that perceptions were not influenced by industry funding disclosures, but perceived credibility was higher for studies reporting negative outcomes. They interpret this as further evidence that physicians often fail to recognise the impact of conflicts of interest, which include withholding of negative results and other forms of publication misconduct.
As a bioethicist/qualitative researcher I cannot comment on the methods. It was not clear to me, however, why the side effect profile was reported as different between the positive outcome abstracts (limited side effects) and the negative outcome abstract (important side effects). Could this not have been a confounding factor if psychiatrists (implicitly) assumed that a study must be credible if it clearly reported severe side effects in its abstract? This might have contributed to the higher credibility apparently assigned to "negative" studies.

Answer:

We thank the reviewer for the insightful remarks. We intentionally added more severe side effects in the negative outcome abstract to make the outcome of the study more negative with the aim to be perceived as more negative. Indeed, psychiatrists and patients are focused on side effects as these effects are important discussion points of psychiatric consultation and follow-up consults with patients. Therefore we have added a sentence in the limitations section in the discussion on page 13:

Fifth, the side effect profile was different in the positive outcome abstract than in the negative outcome abstract. This was intentional to assure that the positive outcome abstract would be perceived as a positive study. However, this may have had a collateral effect on the perceived credibility, and reporting severe side effects might improve the intuitive credibility of a trial in psychiatrists as they frequently discuss side effects with patients and are focused on side effects in their treatments.

Notwithstanding the above, it is an interesting finding that psychiatrists did not assign less credibility or relevance to industry-funded studies. The authors conclude from this that psychiatrists are not swayed in their assessments by knowledge of industry funding, but I think it is important to at least consider the possibility that participants (probably not unreasonably) assumed that a study of this kind would be funded by industry even without a disclosure. This potential explanation at least warrants consideration.

Answer:

This is a good point, which we added to the limitation sections and we now write in the discussion section (page 14):

Participants may also have assumed that the study on a novel antipsychotic would automatically be funded by industry even without an explicit disclosure statement.

Finally, as a bioethicist, I would have liked to see a more developed discussion of the strengths and limitations of disclosure as a means of managing conflict of interest, and the range of
possible practical implications of these findings. For example, while it may be concerning that disclosures do not impact upon assessments of abstracts, is this necessarily what disclosures are supposed to do? While I am sympathetic to concerns about industry influence and conflict of interest, one could argue that it would not necessarily be desirable for an abstract to be judged as a less worthy simply because of an industry disclosure.

Answer:

These are valid points. We have now expanded the implications section of the Discussion (see page 12).

‘It also has to be taken into account that disclosure of industry funding in itself does not make study results more or less valid (25)’. It may well be that respondents does account for the disclosures. For example by knowing the literature on CoI. They may feel that they will not let this knowledge influence the perception of research results. Future research may also address the necessity of funding disclosures on scientific abstracts. One might argue that reporting industry disclosures may also unjustifiable influence the perceived credibility of research results simply because it was funded by or carried out by the pharmaceutical industry.

Maia Lesosky (Reviewer 3):

The manuscript describes a survey of Dutch psychiatrists. The survey itself randomised emails to individuals to read one of 4 different abstracts with different outcomes / industry funding disclosure. The main survey outcome was evaluated on a 10-point Likert scale.

The major issue is around the study design, and subsequent conclusions.

1. The claim of a 45% response rate is somewhat misleading, given that 1566 were randomised and only 395 surveys analysed.

The response rate (regardless of the denominator) is low enough to impact significant bias. The descriptive summaries in Table 1 are appropriate, but do nothing to assure the reader that the sample analysed is at all representative of the general (Dutch psychiatrists) population.

Answer:

Thank you for this comment. This is an important limitation and we have addressed this limitation in more detail in the discussion section:

To calculate the response rate, we used the psychiatrists who opened the email as the denominator. A stricter calculation of the response rate would use all invitees (n=1566) - whether
they opened the email or not - in our response rate determination. This would have resulted in a response rate of 25% (395 out of 1566 invitations).

We did not survey non-respondents for the reasons not to participate. Psychiatrists may lack time and energy to engage in online surveys, or it may be related to the subject of the survey. Possibly, some psychiatrists might have been reluctant to participate as they might feel that they are not competent enough to judge a scientific abstract on its quality. Such bias could have affected the primary outcomes.

Moreover, in light of the response rate of 45%, we checked whether the gender of our study sample was comparable to the total population of Dutch psychiatrists. With 54% of the survey sample being male compared to 59% in another recent study of Dutch psychiatrists (1), there is at least no reason to assume a large bias in the gender of participating psychiatrists.

2. In a classical hypothesis testing framework, the absence of statistical significance, which is the primary finding of this analysis does not impart the conclusion of no effect. The result of 'no significant association' based on a hypothesis test only allows one to conclude there is insufficient evidence to reject the null hypothesis of that test. The language throughout the manuscript needs to be changed to reflect an appropriate interpretation of these hypothesis tests.

Answer: This is indeed true and should be made clearly throughout the manuscript. We have gone through the manuscript to downplay our language and make sure this important limitation is addressed in the discussion section. For example, we now write:

Third, we did not find a difference between funding disclosure and perceived credibility. Although it may be enticing to conclude that there is no relation between the two, we cannot conclude this from our data as no evidence of an association is not evidence of no association.

3. It is hard to evaluate if ANOVA can be considered appropriate without any reporting of the distribution of outcomes. A 10-point Likert scale can easily behave as a discrete variable, for which ANOVA would be inappropriate.

Answer:

We have added the distribution of the primary and secondary outcomes and did not find an abnormal distribution of the Likert scale items that warrants additional statistical tests for skewed distributions. We have added the table below in the supplementary material. The variables are not discretely distributed.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>+/- SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
</table>

Table S4 Distribution of primary and secondary outcomes

In the manuscript, we write in the Methods section (see page 6/7):

We checked for distribution patterns of the primary and secondary outcomes and concluded that the distribution allows the use of ANOVA (see supplementary table S4).

4. The discussion and conclusion overstate the findings and should emphasis the effect sizes more that the results of the significance tests.

Answer:

We agree with the reviewer. We have now downplayed the interpretation of the significance levels in the discussion and conclusion. Furthermore, we have described this limitation of our interpretation of the findings in the discussion section. We also have added a Cohen’s d (effect size). We write whether it can be debated that a 19% difference (Cohen’s d 0.43) in the perceived credibility scores and 13% difference (Cohen’s d 0.19) in interest in reading the full article between groups is large enough to draw firm conclusions.

We now explicitly write in the Discussion section on page 13:

Secondly, it may be debated whether a 19% difference (Cohen’s d 0.43) in perceived credibility score and the 13% difference (Cohen’s d 0.19) in interest in reading the full article score between the two groups is large enough to allow firm conclusions.