Responses to manuscript committee meeting and journal reviewer comments

Responses to comments from the meeting:

1. The topics of publication policy and publication bias are interesting to our readers, though for the BMJ the limitation to just one company does reduce the interest. There have been other studies of publications from other single companies. For example Mooney LA et al. Cross-sectional study of Pfizer-sponsored clinical trials: assessment of time to publication and publication history. BMJ Open. 2016;6:e012362 . Response: While the topic of industry bias in reporting results of negative studies has been widely debated (most recently in BMJ by Ahn et al – BMJ 2017:356:i6770), to our knowledge no other study has systematically addressed the issue of submission and publication bias based on study outcome. We believe that this is an important topic for readers and that this paper presents the opportunity to promote further discussion and data sharing related to challenges and barriers to full transparency. We recognize the important limitation highlighted but suggest that this is mitigated by the study's strengths, in particular the large sample size, inclusion of all studies (regardless of phase or outcome), rigorous methodology, and the importance of the research question coupled with the current lack of data to inform discussion. The Mooney et al study (referenced in our paper), while providing valuable information on publication rates does not address the important issue of study outcome as it relates to potential submission or publication bias, one of the key factors cited in the findings of Ahn et al.
2. Can you be more precise about what you counted as "publication"? Does it include conference proceedings, a poster abstract, posting of results in a repository (such as clinical trials.gov) as well as full journal articles? If you have included other categories of publication please make this clear. Response: The term "publication" has been clarified in the manuscript to refer only to full journal articles. Information on a smaller additional set of studies disclosed as conference abstracts is also presented but more clearly separated in the revised manuscript. Posting of results on a public register was not considered "publication" (also clarified in Methods) although we do confirm that 1041 of the 1064 studies in this cohort have also been disclosed in one or more public registers and that 100% of FDAAA reporting requirements were met.
3. We did not understand the rational for starting with "All studies completing during this period that had a primary manuscript submitted or were due for submission". Does this not leave hidden the studies that the company decided not to submit for publication? Starting with the sample as it appears you have means we cannot obtain insight into publication bias from failure to submit and are limited to exploring the editorial selection, peer review and author persistance process only. Have we misunderstood or is this indeed the case? Are you able to add in the manuscripts that were not submitted? Response: We apologize for the misunderstanding. The entire cohort of 1004 drug studies completing between January 1 2009 and June 30 2014 was included in the analysis, whether they had been submitted or not, to ensure a comprehensive picture. The January 1, 2009 start date was chosen, because since this date, GSK policy requires all human drug research studies to be submitted for publication within 18 months of study completion. The June 30, 2014 end date was chosen to ensure capture the cohort of studies expected by this policy to have been submitted for publication no later than December 31, 2015. An additional 60 studies completing after June 2014 but also submitted for publication by December 31, 2015 were also included, for a total of 1064 studies that were, or should have been submitted for journal publication between 2010-2015. We have revised the Methods section to make these criteria more clear. We have added Supplemental Table 1 in the form of CONSORT flowchart to describe the selection in more detail. During the revision process, we revisited our data set, and in the process of tightening our study inclusion requirements , have added 6 studies and deleted 44 others to change the overall sample size from 1102 to 1064. Of the 44 studies deleted, 25 were due to change in sponsorship, 17 were due to updated study submission and completion dates falling outside of cutoffs and 2 studies never administered study drug to patients. All database changes were applied without regard to study outcome or publication status. This has had no appreciable effect on the descriptive statistics but has served to sharpen the sample's date boundaries (our original approach deliberately erred on the side of inclusion, i.e., when there was doubt studies were included). We enclose a complete list of

additions and deletions (supplemental file #1). If the editors or reviewers prefer our original, slightly over-inclusive data set, we are happy to revert to the latter.

4. If you feel you can address the points above to our satisfaction, please also revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below. Response: The revised manuscript addressing editorial committee and reviewer comments has been resubmitted. We thank the editors for the opportunity to revise the manuscript and the reviewers for their thoughtful and insightful suggestions and comments which have served to improve clarity and understanding.

Comments from Reviewers

Reviewer: 1

Recommendation:

Comments:
This paper addresses an important question for publications professionals (including academic-, industry-, and publisher-based): whether study outcome of industry-sponsored research influences ultimate publication. Like the authors state, I am also unaware of any other research addressing this question.

Aspects of the methodology of the study appear strong, particularly the sample size of more than 1100 papers, the scope of 5 years, and the use of an independent body to classify research by outcome. I note that the differences and similarities in publication metrics are descriptive only and not statistically evaluated. If there is a proper statistical approach to quantifying the meaning of the findings, it should be taken. Response: We have included 2 x 2 contingency testing of submission and publication rates for positive vs. negative studies in the manuscript. Both rates were significantly lower for positive vs. negative studies. We suggest that these results should not be over interpreted as evidence of publication bias against positive studies, but rather as clear evidence of lack of bias against negative studies in this large cohort.

The paper would benefit from a more thorough examination of bias, by stakeholder and by motive. The first sentence of the abstract concludes that "there was no evidence of submission bias," (addressing potential sponsor bias), although the relationship of bias to study outcome, role (sponsor or publisher), or study outcome is not addressed in the abstract. If the authors hypothesized that journals may be biased against publishing negative studies, they should say so. The second sentence of the abstract implies that the study found no evidence of journal bias, but this is not clearly stated. Response: The abstract and introduction have been revised to state the original hypothesis and related results more clearly. The introduction also discusses potential sources of bias from two key stakeholders, namely sponsor and publisher and also references the recent BMJ paper by Ahn et al. regarding potential sources of bias.

The introduction makes a strong case for the importance of publication of all study results regardless of outcome. However, the first sentence of the second paragraph may overstate the case. Not all research reporting "informs treatment decisions," particularly those of negative results in early-stage development programs. This area could serve as a springboard for discussion of potential journal bias, or at the least, a discussion of the importance of proper journal selection in publication planning (i.e., not submitting early-stage investigational work to a clinical treatment-oriented journal).
Response: The first sentence has been revised to avoid overstatement. The authors do believe that all research studies should be designed to either directly or indirectly inform research or treatment decisions (as negative studies often do). In the interest of brevity, we have not discussed potential journal bias (other than mentioning the possibility that it exists, given that there are so little data on

the subject) but do mention the importance of planning and journal selection, especially for “negative” studies in the discussion.

The research hypothesis should be more clearly described. Beyond study outcome having an effect on submission and other publication metrics, the authors should specifically state what they hypothesized the effect to be. This can be done diplomatically without alleging systematic journal bias against negative results; and, in fact, the study results found this not to be the case. Response: The research hypothesis has been more directly stated.

I applaud the use of an external party to classify the studies as to outcome. More information on the role of the 3rd party and their preexisting relationship and base level of knowledge about the sponsor would be valuable. Likewise, agreement metrics (concordance statistics) between the 3rd party and the verifying author (GE) would assure readers that the classification was rigorous. The verifying author may have been blinded to publication status; however, it is likely he knew the status of at least some of the studies given his role in the company, and therefore the extent of any disagreement in verification should be documented. If this is low or nil, it strengthens the rigor. Response: Concordance figures have been calculated and included and additional information on the third party is now included within disclosures. Every effort was taken to blind the verifying author to publication status when adjudicating study outcomes, although we cannot dismiss the possibility of inadvertent knowledge. We are providing a full listing of studies and their outcomes and invite readers and or reviewers to judge the study outcomes for themselves.

Figure 1 is an important piece of data, but it's not optimally presented. What it appears to conclude is (a) positive and negative studies appear to have similar rejection rates when stratified by the number of previous submissions; and (b) rejection rates decrease with increasing previous submission, perhaps demonstrating better journal selection, improvements on rewrite, or both. This would be clearer if the graph presented acceptance rates (1 minus the rejection rate) for each n of previous rejections. Finally, this figure would be significantly strengthened with the inclusion of data for mixed and noncomparative studies. Response: We have provided alternative versions to Figure 1 as supplemental materials and are prepared to discuss the optimal means of presentation with the editors. We continue to believe that our original figure quickly conveys the overall data pattern for acceptance for positive vs. negative studies and the overall progression of multiple submissions, which continually pares down the remaining number of non-accepted submissions.

The first paragraph of the discussion merely restates results and could be omitted. This space could be used for a more thorough, frank discussion of bias and its relationship to the study design and findings. The finding that negative studies took 98 additional days to publish is interesting. Potential reasons for this include competing priorities at the sponsor and at the journal. If possible, it could be useful to evaluate the journals for negative versus positive studies. Impact factor is acknowledged to be a flawed measure, but it could be hypothesized that negative studies are ultimately published in less prestigious journals. Response: Our understanding of the journal style guide is that the first discussion paragraph should be a statement of the main results. However we are happy to précis this further and allow more space for discussion if that is the editors’ preference. We have added some discussion on timeframes to the current text on journal selection. As a post hoc response to the reviewer query, we examined Impact Factor and did not detect any correlation between it and study outcome or submission attempts. Therefore this information was not included in the manuscript although we have provided a supplemental data tool containing journal titles and IF where rejected manuscripts were submitted.

In conclusion, this paper addresses an important question in scientific publishing and should be strongly considered.

Reviewer: 2

Recommendation:

Comments:
General Comments:
This article focuses on an important question relative to data transparency. It approaches the research question using different methods than the literature I have seen. Submission/publication bias is often difficult to assess systematically, but access to a comprehensive list of conducted studies enables the authors perform a systematic assessment.
Data transparency has the potential to have a large impact on clinical practice and the public perception of the pharmaceutical industry. It is important to encourage sponsors to be open about their policies and practice.

Minor Revisions:
- The phrase "study outcome" was confusing to me. Readers may incorrectly assume that this refers to which outcomes/endpoints were measured in each study, rather than whether the findings were positive or negative. Response: We have tried to make it clearer in our revisions that "study outcome" refers to whether the primary outcome measure achieved statistical significance. We have also changed "publication outcome" to "publication status" which we hope will also reduce the confusion for readers.
- Abstracts are short and contain much less information than journal articles. It may more appropriate to report all "publication" rates separately for abstracts and journal articles. "Publication" typically refers to journal articles alone. Some sections of the article report abstracts and journal articles separately, while other sections combine them; it's not always immediately clear which. Consistently reporting them separately would help to clarify. Response: The manuscript has been revised throughout to ensure that journal articles and abstracts are always reported separately and that "publication" is consistently defined as full journal article publication. We continue to include separate description of disclosure rates via congress abstracts because, although more limited in content, they do contribute to public disclosure of new data.
- Some sections of the article use the phrase "publication outcome" (e.g., page 4, line 53) while others use "publication status" (e.g., page 6, line 37). Publication status is more appropriate. Response: We have revised the manuscript to use the term "publication status".
- How specifically were the primary endpoints pre-specified? If they were not clearly defined (for example, using all five elements of an outcome: Zarin, D.A., et al., The ClinicalTrials.gov results database--update and key issues. N Engl J Med, 2011. 364(9): p. 852-60.), this ambiguity could lead to questions about the results of your analysis. You may want to address this in your paper. Response: We did not apply the five element scheme advocated by Zarin, but rather classified studies according to the primary endpoint pre-specified in the study summaries posted to www.clinicaltrials.gov or www.gsk-clinicalstudyregister.com. The need for clarity in specifying endpoints has been added to the discussion.
- A limitation of your research is that you did not assess selective outcome reporting. There has been some evidence showing that primary outcomes in protocols differed from primary outcomes in publications (Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. New England Journal of Medicine. 2009;361(20):1963-71.). Response: We agree that selective outcome reporting is an important issue and has been added to the discussion of study limitations. GSK's policy is to ensure concordance between the primary outcome measures specified in the protocol and results summaries posted, as well as in the final publication. However, we feel that validating this for such a large sample was beyond the scope of our analysis and addresses a different research question. We are providing a complete list of studies used here to allow others to evaluate this important question as a separate research project.

Reviewer: 3

Recommendation:

Comments:
Thank you for the opportunity to review this manuscript. This appears to be a well-conceived study with a straightforward and easily understandable methodology and objectives, and comprehensive supporting data in a supplemental file. The research question -- whether GSK's study outcomes influenced their submission and acceptance rates for publication – is answered in the negative.

This is interesting information, especially in light of the ongoing debate over publication of clinical trial results both in the medical literature and in public databases like clinicaltrials.gov. While the methodology of this study seems well-designed and well-executed, what is lacking in my opinion is a more thorough introduction and discussion that places the findings within this greater context. A review of the history leading to the goal of 100% submission of research studies would greatly increase the value of this study, as would a discussion of why the study's findings appear to be at odds with perceptions elsewhere in the academic community. A discussion of the state of transparency from the manufacturers' point of view and how GSK compares with other researchers, to the extent that information is available, would provide useful background. In that context, it would also be helpful to know how publication of these results in the medical literature compares to their publication on clinicaltrials.gov. In short, this study could be much more valuable if it were expanded to include information about the progress of disclosure at GSK, and the context within which that progress has occurred. Response: The introduction has been expanded to describe more fully the evolution and details of GSK's disclosure policies. These were not included in the initial submission because we were unsure of the level of interest to the reader. We have avoided further detailed discussion of our company's transparency policies (e.g., GSK's commitment to patient level data sharing for all studies) due to space limitations but would be happy to do so if the editors think it would be useful and appropriate.

On a more specific level, the references to the 1102 studies included in the present study are somewhat confusing. In the Results section (P 5, L 21-22) these studies are described as "GSK-sponsored human drug studies that completed between 2010 and 2015 and had a primary manuscript submitted/due to be submitted, prior to data cutoff." Yet 105 of them are noted as not submitted. Please clarify whether 1102 is the total number of relevant studies done at GSK between 2010 and 2015, or is the number of relevant studies submitted for publication. If this is not the total number of studies, it would strengthen the report to have that number. Response: The number of studies and the criteria used for inclusion has been clarified (see also response to previous comment).

More details and explanation of the nature of the studies within each of the four categories devised by the authors would be helpful. This is particularly true of the largest of the four categories, noncomparative studies. These studies constitute over half the studies considered, yet are by definition excluded from the research question. Given that the rate of publication is similar for positive and negative studies, but that there are almost twice as many positive studies as negative, the reasoning behind the assignment of studies to each category is essential to explaining the current study's outcome. It would be helpful to have some kind of breakdown of the noncomparative and mixed studies and more detailed discussion of how they were classified. Similarly, in the unsubmitted studies, a brief elucidation of the nature of the studies terminated early and an overview of the reason why 48 bioequivalence studies were not submitted would be helpful. Response: Our cohort included all GSK Pharmaceutical studies completing during the study period, irrespective of study phase. It is particularly common for Phase I studies to include only descriptive analyses due to small sample sizes. Further clarifying information has been added to the manuscript. In addition, the attached supplemental table #2 may be the most appropriate way to address the question of why so many studies cannot, by the study methods used here, be categorized as having positive or negative

outcomes. This has been provided together with additional supplemental files(#2&3) providing further descriptive detail on studies that were submitted only as congress abstracts and those not submitted in any form.

Finally, a minor but puzzling point: Why is the period from Jan 1, 2010 to Dec 31, 2015 considered five rather than six years? Response: The reviewer is indeed correct – this has been corrected. Our initial cohort of studies includes completions over a 5.5 yr period but the submission period is indeed a six year window.

I would recommend acceptance with revisions.

Reviewer: 4

Recommendation:

Comments:
I would like to thank the editors for giving me the opportunity to review this interesting paper.

The described study evaluated publication rates of 1102 GSK sponsored studies from 2010-2015. It provides stratified analyses by study phase, and study outcome. Based on the data presented it concludes that there is no evidence for submission bias.

I do have some remarks which I would the authors ask to address.

Major:
- Selection of sample: I am somewhat unclear whether the sample of 1102 studies includes all studies sponsored by GSK that completed between 2010-2015. I am raising this since it says in the abstract, design description, that only studies were included that were submitted, or were due for submission (per the sponsor's policy) for publication during this period (page 2, line 11-12). Please clarify. Also, if some studies were excluded because they were not due for publication per the sponsor's policy please add more information on number of studies excluded based on this. Also, would be interesting for the reader to learn more about what type of studies are not due for submission per the sponsor's policy, etc.). Response: Additional information has been added to the revised manuscript on this point (see also response to previous comments)
- When presenting data I would generally suggest to focus on data of full journal publications, and then add what proportion on top of this has been disseminated as abstract only. E.g. on page 2, line 36, when speaking of 90% that have been submitted for publication, I would suggest to rephrase to 84%, and then add that on top of this, 7% were disseminated as abstracts. Given the very limited amount of information that can be put in an abstract only, this seems more appropriate. Response: Reporting of full publication vs. abstracts has now been more clearly separated and clarified (see also response to previous comments)
- Given that the authors also worked with CT.gov, I wonder whether they assessed as well for how many studies complete summary results are posted in CT.gov. If not, this would be in my vie a valuable addition. Response: Although not directly related to our research hypothesis, results disclosure via CT.gov and other public registers is an important issue. We have provided additional information on posting of this cohort to public registries in the manuscript (98% had results posted to ct.gov or gsk-clinicalstudyregister.com). It is also important to remember that only a subset of studies in this cohort are required to have results posted on CT.gov (100% of those required by FDAAA were posted there).

Minor:
Abstract
- Page 2, line 52: While I agree that it journal publications are important, I don't think that a journal publication can achieve "..full transparency of human study outcomes…". So, suggest rephrasing

slightly. Response: This has been rephrased.

Article Summary

- Page 3, line 13: I think generalizability is limited given that sample only from one single sponsor, so suggest to delete “potentially”. Response: “Potentially” has been deleted.

Introduction

- Page 3, line 44-52: Wondering whether this para would better fit the methods section.

Methods Response: We have revised the paragraph in the Introduction to explain more clearly why given date frames were chosen and have included the descriptive detail in Methods.

- Page 4, line 16ff: Why were studies on vaccines, etc. excluded. Consider providing flow chart that indicates how many studies were screened, and how many were excluded for which reasons. Also, how was significant patient enrolment defined? Response: As previously noted, we have included a supplemental CONSORT diagram to clarify the sample selection process. Vaccine studies were not included for a number of reasons: at the time of the cutoff, the timelines for reporting vaccine studies differed from drug studies (e.g., measurement of antibody titers may lag significantly behind other outcomes), they are typically published in a different cohort of journals, and both the studies and publications are managed within a separate division of the company. Similarly other non-drug studies were excluded in the interest of focusing the research topic on publication/non-publication of negative drug studies.
- Page 4, line 20: I wonder how complete the list of outcomes posted on CT.gov or the GSK internal register is, and whether this would allow to identify re-classification of outcomes e.g. from secondary to primary, or the other way around. Given that these are all GSK sponsored trials, why did you not use the actual trial protocol (+amendments), which might provide the most complete description of outcomes. Response: The authors agreed on prospective criteria for assigning positive vs. negative studies that were predicated solely on outcome of the primary endpoint. GSK uses a unified process for reporting primary endpoints in study summaries and study reports to ensure consistent reporting of outcomes. Given that study outcome classification was based solely on primary endpoints, we did not deem it necessary to consult final study reports for the same information that was already available on registries such as clinicaltrials.gov and on GSK’s public clinical trial register www.gsk-clinicalstudyregister.com (see also response to comment from Reviewer 2).
- Page 5, line 5: Personally, I would think it is of more interest to present data separately for studies rejected and not submitted. While it makes no difference to the potential user/reader, separate presentation will shed some light on the mechanism/reason underlying non-publication. Response: We are not sure we understand the reviewer’s request. All tables report separately those studies that were submitted but rejected vs. those that were never submitted. Aggregated submission rates include rejected manuscripts but aggregated publication rates do not.

Results

- Page 5, line 22: I think it would be helpful to provide a clear definition/criteria how GSK defines studies that are “due for submission”. Is any study in which GSK is involved as a sponsor due for submission? Response: Further clarifying information has been added to the revised manuscript (see also responses to previous comments).