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)	Measuring clinical trial transparency: An empirical analysis of newly approved drugs and large pharmaceutical companies
AUTHORS	Miller, Jennifer; Wilenzick, Marc; Ritcey, Nolan; Ross, Joseph; Mello, Michelle

VERSION 1 – REVIEW

REVIEWER	Joel Lexchin York University Canada In 2015-2016 Joel Lexchin received payment from two non-profit organizations for being a consultant on a project looking at indication based prescribing and a second looking at which drugs should be distributed free of charge by general practitioners. In 2015 he received payment from a for-profit organization for being on a panel that discussed expanding drug insurance in Canada. He is on the Foundation Board of Health Action International.
REVIEW RETURNED	12-Jun-2017

GENERAL COMMENTS	This study looks at the transparency of clinical trial reporting by large drug companies conducting studies required for marketing in the United States. 1. While the CSR synopsis can give important information, I feel that the entire CSR should be available and that the authors should look at this metric. 2. The authors are equating the amount of data reported in registries, CSR synopses and journal articles. Is this justified, i.e., is the level of data the same in all three? 3. Page 9, line 25: Why did the authors exclude trials involving patients with renal impairment? 4. Page 20, line 46: The authors need to explain what the Final Rule is. 5. Page 21, line 3: The authors should note other limitations: their results only apply to large companies having drugs approved by the FDA in 2014 and only to drugs that are governed by the FDA rules. 6. The authors should outline measures that might improve transparency in reporting.

REVIEWER	Matthew Herder Dalhousie University, Faculties of Medicine and Law, Canada
REVIEW RETURNED	19-Jun-2017

GENERAL COMMENTS	The authors provide a useful follow on study to their 2015 cross-sectional analysis of drugs approved by the FDA in 2012. I believe the paper should be published, following certain minor revisions. My comments are as follows: 1. The relationship between the present paper and the 2015 paper should be explicitly clarified. I understand that the current paper analysed a much higher number of data sources. This should be explicitly contrasted with the 2012 paper, and perhaps used to deepen the discussion around one of the main findings; namely, that nearly all of the trials were registered in Clinicaltrials.gov. There has been worry expressed in the literature that the proliferation of trial registries would result in fragmentation of evidence. But the authors' finding about the wide reliance on Clinicaltrials.gov undercuts that concern. It is less clear whether the outcome measures analysed in the present paper are exactly the same as, or depart from, the measures used for the 2015 paper. This should be clarified. 2. On a related note, I found the terminology a little confusing. It appears that 'outcome measures', 'standards', 'metrics', and 'benchmarks' are used interchangeably throughout the paper. Each time the term shifted, I was unsure whether a new issue was being raised. As a reader, I would encourage the authors' to simplify and use 1-2 terms throughout. 3. Regarding the authors' sampling strategy, I think the authors' should provide more justification/explanation for two aspects. First, I was puzzled as to why only focus on the largest biotechnology / pharma companies. Perhaps this is another instance where the authors could highlight the fact that they were building on the 2015 study, which similarly examined larger companies. If, however, large companies are improving their compliance, then perhaps the authors' should note that greater attention to other, less well resourced players' compliance with transparency is needed going forward. Second, and more importantly, can the authors' provide more explanation as to why they excluded observational studies and whether that represents a serious gap? It may not be the case for the 19 drugs under scrutiny, however, I believe the literature suggests that observational studies are being relied upon (in concert with trials) more often for regulatory approval. If that is accurate, the decision to omit observational studies from the analysis - at least as against the ethical standard/metric - may be called into question. 4. On a fairly trivial note, I believe Table 2 is organized in a confusing way. Specifically, in the text (pp. 11-12), the results for the patient trials sub-sample as depicted in Table 2 are described before the results for the trials that include healthy volunteers. But the table, reading left to right, is in the opposite order: all trials first, then the patients only trials. I suggest switching their order in the table unless I've misunderstood the text. 5. If I understand it correctly, the ranking process described on p. 10 appears very favorable to industry. That is, the authors scored companies based on transparency of patient trials only (whereas the results for trial registration of all trials

	are much lower), and no matter which interpretation of the reporting requirement the company in question followed. Are the rankings different if 'all trials' are used instead? And/Or if the more stringent 'trial completion date' interpretation of the FDAAA is relied upon? Even if the rankings are more or less the same, is the 'transparency the score' the same? The reason I raise this issue is that some may charge that an inflated transparency score can be used to justify industry practices, which remain in need of improvement.

REVIEWER	Kay Dickersin Johns Hopkins Bloomberg School of Public Health
REVIEW RETURNED	07-Jul-2017

GENERAL COMMENTS	Objective: is this study part of a series or a semi-annual update? The objectives seem different from the, 2015 BMJOpen paper, yet there are many similarities, It would be helpful to have this clarified. Abstract: Some small wording changes in the Results and clonclusions would help with clarifications on (1) approvals are for drug-indication pairs, not the drug overall; (2) reporting results may not be equivalent to sharing CSR summaries; (3) the definition of FDAAA compliant (eg, "all trials" are confusing throughout since the way I read the law and final rule, phase 1 studies are not included); (4) defintion of publicly disclosed results; (5) the meaning of "standards" (eg, do you mean for FDA reviews?) Replicable methods: How did you find those trials subject to results reporting under FDAAA? My understanding is that the law pertains to off label trials for any approval, not just the indication of the approval. Not sure what it means (P. 8 lines 8-13) that you incorporated feedback from companies if it could be validated thorugh the public sources. See my comments otherwise Research ethics: Not sure where IRB information is located. References: you mention the Dickersin and Rennie 2003 JAMA article on trial registration, and there is a 2012 update too (not referenced). A lot happened in that time period, including FDAAA, so I would tend to refer to both. But I am biased. PRISMA: Why did you use PRISMA, given that as you say this is a cross sectional analysis? I would think STROBE makes more sense. I have addressed many of your questions in my comments so don't want to do all that again. Hope that is ok. Comments overall: There were a number of things I did not understand in this study and it would be good to explain them for people like me. First, why did you include phase 1 studies? As I understand it, the law does not apply to them, even under the Final rule, and so it seems odd to hold the companies to a standard that does not apply. If I am wrong, please be clear about this change and educate people like me! Second, please explain in the paper that FDA approval is for a drug-indication combination and does not imply that the drug is effective or safe for other uses. Third, could

	showing your findings beforehand to the companies seem to represent a conflict of interest to some people? I see why it was a "good idea" for your study, but it could also lead to a perception of working together rather than this being a true "scorecard". Fourth and finally, could you explain (assuming that this is what you mean) how you determined that CSR synopses are as good as CSRs. In our experience they really are not, but maybe this is a bigger sample. I am confused about timing of the Final Rule and your cohort of trials. Can you explain this for us? Abstract: Is this article one of a series with a semi-annual update? Please explain p.2 Line 12 Please explain that FDA approval is for a drug-indication combination and does not imply that the drug is effective or safe for other uses (ie, off label or other on label uses). p. 2 line 38: not all patient trials, just those in FDA reviews p. 2 line 39: What standards are you referring to? p. 3 Strengths and limitations: CSRs are not public documents in the US; you are looking at on label compliance p. 3 line 11: You certainly searched a lot, but I would be careful not to claim you searched every public registry (for example, you did not search 2 in China that I know of). Introduction P. 4 line 11: What's "it"? p. 4 line 25: What do you mean "under one study"? p. 5 line 11: what are the standards? I am confused by p. 5 line 25: It seems to imply that phase 1 studies in healthy volunteers are now included under expanded standards. Is this so? I don't know about this. Even if phase 1 studies are now included, don't the expanded rules apply starting in 2017? p. 5 line 48: I am a little worried that you are referring to "validation" when you mean checking with pharma to make sure that what you are saying is correct. Perhaps you should report what they asked you to change! You say later on (page 8 lines 8-13) that the feedback from companies was generally incorporated "if it could be validated through our public sources" which sounds both complex and is a little vague! Methods Page 6: State that you only included trials for "on label" indications. I wonder if looking at a sample of trials, one that includes drug companies that include the largest, medium sized and smallest, would give you the same answers? Page 7 line 4: what is a primary completion date? CT.gov says it is defined as: Definition: The date that the final participant was examined or received an intervention for the purposes of final collection of data for the primary outcome, whether the clinical study concluded according to the pre-specified protocol or was terminated. In the case of clinical studies with more than one primary outcome measure with different completion dates, this term refers to the date on which data collection is completed for all of the primary outcomes. Once the clinical study has reached the primary completion date, the responsible party must update the Primary Completion Date to reflect the actual primary completion date. Page 7 line 13: My understanding is that results reporting applies to trials done for approved drugs but off-label indications. Is this true? How did you find these? I would expect this would be a lot harder to do and "compliance" would be a lot less. Did you find this? Page 8 lines 18-30: Not sure about some of your language and

	assumptions here. (1) what is a "trial" (line 18)? Please make it clear how you know that CSR synopses tell you something similar to a CSR. I think you are using the terms interchangeably later on (eg, line 44). In my experience, CSR synopses are pretty uninformative, but maybe this is only true sometimes. Does "available" mean useful? For example, missing outcomes may be such that although the results are "available" they are not terribly useful. Page 8 line 51: You use the terms "metric" and "measure" here and elsewhere to refer to items (I think), a term which might be easier to define without bumping up against other meanings. Page 8 line 55: here and elsewhere why is "all trials" hyphenated? Page 9: the page is a single paragraph. Perhaps it can be divided by topic? Or use bullets that correspond to Table 1? Page 9 line 15: You refer several times to a meeting your group convened in August 2016. Is this reported somewhere? Who came, what happened? This reference is tantalizing but not terribly informative as is. Here are some word questions: Line 15: disclosing results "only for trials enrolling patients"? When are you looking at the others? Line 20: What are "patient trials"? Line 22: How can you rank or analyze "transparency"? Transparency must be defined for each situation where it is used. Line 25: sorry, but I am not sure why "renal impairment" is here? Line 32: Are you talking about registration or completion, in terms of confusion about FDAAA requirements? I thought the final rule clarifies what is required. If this is not the case because the final rule was not in effect for these trials, remind us! There should not be confusion any more. Table 1: I wonder of the table could be improved by putting in a column before the transparency "measure"? The column would refer the issues that the transparency "item" addresses. The column titled "Trials subject to FDAA" could be defined in a footnote Page 11 line 10: Excel was used. Maybe it would be good to move to a database (as opposed to a spreadsheet). Page 11 lines 14-22: see earlier comments on validation and meeting. Results: Page 11 lines 39-49: I am not sure how a median of 7/22 trials works. Is it a median of 7 and a median of 22? I can't quite picture the proportion…. Page 11 line 53: Your standard that results of trials in patients should be publicly available is not in Table 1 as stated. Maybe it is there but wording is different. Page 12 lines 13-18: Confession: I really am not that interested in phase 1 trials, since they are not required to be disclosed, are they? Some are in patients (eg, cancer trials) and some not. So I am still unsure of the point here. I feel embarrassed to ask, but what is your point? Page 12 Lines 25-35: I agree about a standard that all trials in a successful NDA should be publicly available (I thought they had to be whether the NDA was successful or not). These numbers are very low! Page 12 lines 37-44: I am confused. Why should it matter that at least one was publicly available, if all should be disclosed? Table 2: if drug names are not the brand name, please don't use capitalization. Are you saying you have 2 ethics standards or five? Page 14 Lines 11-15: I am confused by the trial that was registered "a few days late." Is this worth mentioning? Maybe put it into one

	bucket of the other, and say in the discussion (if you want) that sometimes being late was a matter of a day or two. Table 3: You mention the interpretation differences for trial completion, but ClinicalTrials.gov must have an opinion. Use that one. You have to make a decision. I don't know enough to do this and am relying on you. I have to admit I am not so impressed by these transparency scores, given previous comments. It might be interesting to see on Table 5 how these scores correlate with the number of drug-indication approvals and also how much they are making per drug! I like Appendix 5 and wonder if it could substitute for one of you other tables, assuming you decide to delete phase 1 studies or reduce the number of sensitivity analyses you show. Discussion Do your results mean that everything is fine? My comments in this section mainly relate to things I have already said. Either you will explain your decisions or make the changes so I won't say much more, except that you will need to check for statements in the Discussion after you decide what to do about the points I made earlier. P 20 line 20: you say that NIH policy requires disclosing information about all trials. But this paper is about drug companies so why are you looking at all trials (including phase 1)? P 20 line 30: You say that disclosing phase 1 trials may help speed innovation and save money. Could your paper examine this question? Again, I get confused when you use terms such as "metrics" which has a meaning to me related to outcomes assessed in trials. Any way you can use a different word? P 22 line 54-56: Can you be more specific about where your data will be available? Seems odd to write a paper about transparency and to be vague about where your data will be publicly stored.

REVIEWER	Capuano Annalisa Regional Center of Pharmacovigilance-University of Campania "Luigi Vanvitelli"
REVIEW RETURNED	19-Jul-2017

GENERAL COMMENTS	This study was designed to investigate a series of clinical trial transparency metrics, on ethics and legal levels, and applied them to large pharmaceutical companies and their 2014 FDA approved drugs. To sum up, I endorse this attempt with great delight. It carries an important message. I would advise towards some minor revisions in order to deliver an optimal result. The scientific article is to some extent innovative, given the fact that few studies have investigated this topic. The scientific accuracy is adequate. The title is adequate. The introduction clearly sums up the available evidence. There exist a lack of details in regards of the operational framework where the study was conducted (does a data extraction form exists? more information on the collaborators that evaluated trials should be provided). In the abstract: the objective should contain a more straightforward research question. In particular, the statement "we defined…" should be removed. In the study design section, it should be emphasized the descriptive nature of the analyses. It could be hard to understand the meaning of outcome measures in the abstract before reading the full-text. I will suggest a more elegant way of presenting them in

	order to improve the reading flow. The conclusion in the abstract is different from those written in the conclusion section of the full-text and is to some extent speculative. I will suggest a “milder” statement because transparency was high only for clinical trials involving patients, and only for some specific companies. References are adequate and resulted qualified and updated with the latest data. Tables sum up the study clearly and concisely. Figures highlight the key points. Statistical analysis is mainly descriptive. Strengths and limitations of this study: “we added a company ranking…” this is not a strength of this study and should be removed. I strongly discourage the authors to include this table in the final version of this manuscript because it is highly speculative. SPECIFIC COMMENTS My major concern is that due to study design, the generalizability of results to a non-US audience could be hard. This aspect should be emphasized throughout the manuscript. • Data sources – page. 6 line 23-27 Authors should submit for peer-review the information obtained through personal communications with drug manufacturers. • Sampled Clinical Trials – page. 6 line 51-52 Regarding the information extracted for each clinical trials, do authors provided a data extraction form to their collaborators? If yes, the data extraction form should be provided in order to improve reproducibility. In fact, in appendix 2, there was no formal definition for each of the information that was required to be extracted for each clinical trial. • Sampled Clinical Trials – page. 7 line 34-37 “trials with unknown…” this statement should be moved to the results section. • Outcome measures – page. 10 line 18 “we excluded trials (N=12)…” More details should be provided for these trials and should be provided in supplementary material. • Transparency score based on patients trials – page. 11 line 51 Results based on all trials should be provided prior to those involving only patients in the text. Moreover, it should be emphasized that clinical trials involving only patients represent only a small fraction of the overall trials. In fact, the transparency scores for the “all trials” are dramatically lower than trials involving only patients. This aspect is not emphasized appropriately throughout the manuscript. Therefore, the manuscript should be modified in order to provide this message to the reader. Firstly, I expect in the text at least some examples of the proportion of trials involving patients with the overall amount of trials for each specific drug. In example: Harvoni® have in total 60 trials analyzed and only 31 (51.7%) involved patients. The transparency parameters for all trials are 53%, 32%, 57% and 40% for %registered, %reported, %published and %publicly available respectively. For trials involving only patients instead, these tend to be higher (92%, 61%, 68% and 74% for %registered, %reported, %published and %publicly available respectively). This aspect should be emphasized also in: 1. Transparency Scores Based on All-Trials in an NDA, Including Healthy Volunteers section; 2. Compliance with FDAAA Requirements section; 3. Timing of Results Reporting section; • Company Ranking – page. 18 line 20 I have some concern. This table could provide a misleading message. For example, for FDAAA trials metrics, Sanofi has only 3 trials while other companies have more trials (therefore, different

	probabilities of having transparency-related problems). This aspect should be emphasized.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 (Lexchin)

1. While the CSR synopsis can give important information, I feel that the entire CSR should be available and that the authors should look at this metric.

Response: We debated this issue carefully in the course of developing the measures for our scorecard. We agree with the reviewer that CSRs, not just synopses, should be publicly available. We just disagree about when to incorporate this standard into the Good Pharma Scorecard. The next iteration of the Scorecard (which evaluates drugs approved by the FDA in 2015) incorporates data-sharing standards, including whether full CSRs are publicly available.

We decided to wait until the next iteration of the Scorecard to assess patient-level data sharing, including CSR sharing, because the present version of the Scorecard (assessing drugs approved by the FDA in 2014, involves trials completed before December 2012. Data sharing and public posting of CSRs was not an issue that received significant attention at that time. GSK, one of the first corporations to publicly commit to sharing CSRs, made this commitment in 2014. PhRMA, the US pharmaceutical trade association, began requiring its members to publicly share patient-level trial data in 2015. The Gates Foundation and Wellcome Trust instituted similar requirements around 2015. Therefore, we decided to wait until the next Scorecard (for 2015 drugs) to add this element.

In the interim, we concluded that the CSR synopsis would provide the key piece of information we sought to measure for the present Scorecard: whether summary results of a trial are publicly available. What is included in CSR synopses is quite similar to what is presented in summary results posted on clinicaltrials.gov. (For example, see NCT01043926, which has CSR synopses link in clinicaltrials.gov and has summary results available.) For that reason, it seemed reasonable to include CSR synopses as roughly equivalent to results posted on registries. Other parts of the CSR are orthogonal to our interest in results reporting for this Scorecard.

2. The authors are equating the amount of data reported in registries, CSR synopses and journal articles. Is this justified, i.e., is the level of data the same in all three?

Response: The three types of results disclosure do not provide the same amount of information – journal articles, in particular, provide types of information not available in registries. For that reason, we separately measured whether the trial was reported in each of these venues. Because our primary interest was in whether the key trial results were publicly available in some venue, we also constructed a composite measure coded “yes” if results were reported in one or more of these venues, as stated in the manuscript. Because they are not equivalent, we report each of these variables separately in the manuscript; readers can therefore hone in on the measure they care most about.

3. Page 9, line 25: Why did the authors exclude trials involving patients with renal impairment?

Response: Renal patients were not excluded from all analyses. They were included in the “all-trials” analysis. They were excluded from the analysis of “trials in patients” because that analysis looked at patients in the intent-to-treat (ITT) population. Participants with renal or hepatic impairment are important for determining ADME, PK, and PD, but the interventions are not attempting to treat these impairments, so they are not patients in the same sense as those in the ITT population. We clarified that our “trials in patients” analysis was limited to the ITT population (p.9 bottom).

4. Page 20, line 46: The authors need to explain what the Final Rule is.

Response: We have added a clarification and reference (p.10 middle and 22 top).

5. Page 21, line 3: The authors should note other limitations: their results only apply to large companies having drugs approved by the FDA in 2014 and only to drugs that are governed by the FDA rules.

Response: We have added this to the list of study limitations (p.22 middle). Please note that previous and subsequent publications include more drugs and more types of sponsors.

6. The authors should outline measures that might improve transparency in reporting.

Response: We appreciate the Reviewer’s comment and have designed the Good Pharma Scorecard as a key measure that may motivate improved reporting of clinical trial results by publicizing companies’ level of compliance with existing and aspirational standards. This effort complements other efforts, including FDAAA requirements, ICMJE policies, and recently issued NIH requirements that mirror those of FDAAA. We have revised the final paragraph of the manuscript to enrich our discussion of this topic (p.23) and the “Strengths and Limitations of this study” page.

Reviewer: 2 (Herder)

7. The relationship between the present paper and the 2015 paper should be explicitly clarified. I understand that the current paper analysed a much higher number of data sources. This should be explicitly contrasted with the 2012 paper, and perhaps used to deepen the discussion around one of the main findings; namely, that nearly all of the trials were registered in Clinicaltrials.gov. There has been worry expressed in the literature that the proliferation of trial registries would result in fragmentation of evidence. But the authors' finding about the wide reliance on Clinicaltrials.gov undercuts that concern. It is less clear whether the outcome measures analysed in the present paper are exactly the same as, or depart from, the measures used for the 2015 paper. This should be clarified.

Response: The reviewer is requesting clarification of how the present paper adds to a 2015 article we published reporting an earlier, pilot version of our scorecard. We have provided a summary of the innovations in the present manuscript on pages 3 and 5. These are as follows:
i. We added a measure of public availability of the results of trials in patients (intent-to-treat population) to accompany previous measures looking at the public availability of results for (1) all trials in the successful NDAs and (2) FDAAA-applicable trials.
ii. We added the CSR synopsis as sufficient for results reporting, responding to feedback about the ability of ClinicalTrials.gov to accommodate phase 1 reporting, and as a component of public availability.
iii. We added a measure of FDAAA legal compliance capturing the current interpretation of FDAAA requirements: that trial results for initial approvals should be publicly reported by 30 days following FDA approval of the indication.

a. At the same time, we removed the legal compliance measure that requires a trial to be controlled, in the “placebo-controlled” sense of the term, since the consensus among companies and regulators is that interventional trials are FDAAA applicable.
iv. We added a measure and ranking of company transparency, which aggregates the transparency of individual drugs for companies with multiple drugs approved.
v. We applied the measures to a more recently approved sample of drugs (drugs approved in 2014, vs. 2012).
vi. We added review of 39 new public trial registries, including international, corporate, and patient registries, rather than just ClinicalTrials.gov.

We have revised the “Strengths and limitations of this study” page to more clearly enumerate the most important of these innovations.

8. On a related note, I found the terminology a little confusing. It appears that 'outcome measures', 'standards', 'metrics', and 'benchmarks' are used interchangeably throughout the paper. Each time the term shifted, I was unsure whether a new issue was being raised. As a reader, I would encourage the authors' to simplify and use 1-2 terms throughout.

Response: In response to this helpful comment, we have edited the manuscript throughout to use the term “standard” to the broad standards on which we scored drugs (for example, our “ethics standard 1”) and “measures” to refer to individual components of those standards (for example, percentage of trials that were published).

9. Regarding the authors' sampling strategy, I think the authors' should provide more justification/explanation for two aspects. First, I was puzzled as to why only focus on the largest biotechnology / pharma companies. Perhaps this is another instance where the authors could highlight the fact that they were building on the 2015 study, which similarly examined larger companies. If, however, large companies are improving their compliance, then perhaps the authors' should note that greater attention to other, less well resourced players' compliance with transparency is needed going forward.

Response: The Good Pharma Scorecard is strategically created to expand in scope annually. Drugs sponsored by small and mid-sized companies will be reviewed in subsequent scorecards (within the next three years). We began with large companies because we believed they would be the most likely agents of industrywide change if they led by example, and because they have the most resources available to deploy toward satisfying transparency standards that are more comprehensive than those currently imposed by law. We have clarified our reasons for starting with large companies in the revised manuscript (p.6 middle).

10. Second, and more importantly, can the authors provide more explanation as to why they excluded observational studies and whether that represents a serious gap? It may not be the case for the 19 drugs under scrutiny, however, I believe the literature suggests that observational studies are being relied upon (in concert with trials) more often for regulatory approval. If that is accurate, the decision to omit observational studies from the analysis - at least as against the ethical standard/metric - may be called into question.

Response: Generally, observational studies were excluded from our sample of trials because they were ongoing at the time of our study. However, it is also worth noting that these studies are not covered under FDAAA legal requirement to report results.

The impact of our decision to exclude these trials is minimal because observational studies were rare for the drugs we examined. In the NDA records we reviewed, they appeared only 5 times out of the 553 studies we reviewed.

We added an explanation to the revised manuscript (p.7 top).

11. On a fairly trivial note, I believe Table 2 is organized in a confusing way. Specifically, in the text (pp. 11-12), the results for the patient trials sub-sample as depicted in Table 2 are described before the results for the trials that include healthy volunteers. But the table, reading left to right, is in the opposite order: all trials first, then the patients only trials. I suggest switching their order in the table unless I've misunderstood the text.

Response: We thank the reviewer for pointing out this problem. We have fixed it by reordering the columns in Table 2 as suggested.

12. If I understand it correctly, the ranking process described on p. 10 appears very favorable to industry. That is, the authors scored companies based on transparency of patient trials only (whereas the results for trial registration of all trials are much lower), and no matter which interpretation of the reporting requirement the company in question followed. Are the rankings different if 'all trials' are used instead? And/Or if the more stringent 'trial completion date' interpretation of the FDAAA is relied upon? Even if the rankings are more or less the same, is the 'transparency the score' the same? The reason I raise this issue is that some may charge that an inflated transparency score can be used to justify industry practices, which remain in need of improvement.

Response: In calculating our initial set of rankings, our team discussed whether to feature rankings based on all trials or only trials in patients. Ultimately, we decided to feature the results based on trials in patients for 4 reasons. First, we did not incorporate both the “all trials” and “patient trials” into the composite score, because this would double count the patient trials (they are in both samples). Second, patient trials constitute the most important trials supporting FDA approvals. Third, there is disagreement about the value of disclosing information from trials in healthy volunteers, as we discuss on pp. 19-20. Third, companies’ existing commitments, as enshrined in industry guidelines, center on trials in patients. Over time, we anticipate that these commitments may expand to trials in healthy volunteers, but as a first step we felt it was appropriate to measure the extent to which companies were honoring their own expressed commitments. Lastly, FDAAA only applies to trials in patients, and measuring compliance with FDAAA was a major component of our analysis.

We believe the discussion on pp.20 -21 helps readers understand why our final company rankings are calculated as they are. However, it is important to bear in mind that throughout the manuscript, we also provide calculations for each measure in the all-trials sample, and we comment on the differences between the trials-in-patients and all-trials samples. Thus, readers who are more interested in the all-trials sample are able to see what the scores for these drugs looked like.

It is worth noting that although the all-trials sample looks like a much larger sample when one looks at the number of trials, it isn’t much bigger in terms of the number of enrolled people. Only 7% of all research participants in the all-trials sample were healthy volunteers. Thus, our company rankings include the majority of enrolled research participants (93%). We have noted this on p.12 (middle) of the revised manuscript.

Reviewer: 3 (Dickersin)

13. Objective: is this study part of a series or a semi-annual update? The objectives seem different from the, 2015 BMJ Open paper, yet there are many similarities. It would be helpful to have this clarified.

Response: This manuscript is part on an ongoing project in which the Good Pharma Scorecard is being gradually expanded and applied to additional samples of companies and approved drugs. We intend to produce a report approximately once a year, analyzing drugs approved by the FDA each year. As with this manuscript, each successive report will provide a longitudinal perspective on whether companies' adherence to our original set of measures has improved, as well as new information on how they scored on new dimensions of the scorecard. We will seek to publish future results in scholarly journals if there continue to be significant innovations in our methods with each new study, but results will, at a minimum, be posted on Bioethics International's website.

We have not added this information to the manuscript, as it seems to fall in the category of nonessential background information, but would be happy to do so at the Editor's request.

14. Abstract: Some small wording changes in the Results and conclusions would help with clarifications on (1) approvals are for drug-indication pairs, not the drug overall; (2) reporting results may not be equivalent to sharing CSR summaries; (3) the definition of FDAAA compliant (eg, "all trials" are confusing throughout since the way I read the law and final rule, phase 1 studies are not included); (4) definition of publicly disclosed results; (5) the meaning of "standards" (eg, do you mean for FDA reviews?)

Response:
Regarding (1), although it is true that the originally approved indication could be seen as a drug pair, we are targeting new molecular entitles – what the FDA calls "novel new drugs" – which indexes the drug approval to the year. Thus, we have retained the simpler terminology "drug".

Regarding (2), the revised abstract now says "reported results or shared a CSR synopsis," which clarifies that these were two distinct measures. For a broader discussion of the differences between CSR synopses and other kinds of results reporting, please see our response to comment #1 above.

Regarding (3), the abstract reports that "Per drug, a median of 100% (IQR 75-100%) of FDAAA-applicable trials were compliant." We do not use the term "all trials" here because indeed, not all trials are subject to FDAAA's requirements.

Regarding (4), we define the specific types of reporting that satisfied our measure for "publicly reported results" early in the manuscript. We are not able to provide that wordy explanation within the confines of the abstract.

Regarding (5), we have clarified in the revised abstract that we mean "transparency standards".

15. Replicable methods: How did you find those trials subject to results reporting under FDAAA? My understanding is that the law pertains to off label trials for any approval, not just the indication of the approval. Not sure what it means (P. 8 lines 8-13) that you incorporated feedback from companies if it could be validated through the public sources.

Response: The reviewer is correct that trials studying indications other than those the drug was approved for are covered by FDAAA's reporting requirements if they are later than phase 1 and meet other FDAAA coverage requirements, such as start date. Their results are due following the approval of the new indication, 210 days following the withdrawal of such approval, or 2 years after the submission of an extension certificate to delay results beyond the standard 1-year deadline. However,

the law does not require that a certificate of delay be submitted at a particular time. Therefore, without knowing when the approval will be for the new indication, or whether and when a certificate will be submitted, there is no way to measure compliance. The Final Rule addresses this problem by requiring the certificate to be submitted within the first year following the trials PCD. To replicate our set of FDAAA-applicable trials, a researcher can determine which trials meet the conditions of FDAAA and study the originally approved indication. Notwithstanding, most of the trials for new indications fall outside our study timeframe, because the trials are either ongoing or too recently completed.

In the revised manuscript, we have clarified that our sample only included trials for the originally approved indication (p.7 middle).

The reviewer also requests clarification about the type of feedback from companies that we incorporated. The scenarios below typify the feedback we received:
1) A company indicates that results were reported for a trial on ClinicalTrials.gov, but we marked it as not having results there. We recheck ClinicalTrials.gov, confirm that the results are there, and revise our data. The records on ClinicalTrials.gov are constantly being updated with new information and there are occasional delays between when a company reports results and ClinicalTrials.gov updates its records, so it is not uncommon that companies have submitted information that does not yet appear. For this reason, it was useful to be pointed to an updated version of a trial's ClinicalTrials.gov page.

2) The company clarifies whether FDAAA applies to a particular trial. One requirement for FDAAA coverage turns on manufacturing data, which can be difficult to determine through public sources. FDAAA says that a trial must be conducted in the US, or involve a drug, biologic, or device that is manufactured in the US or its territories and is exported for research, to be covered. In a few instances, companies informed us that this requirement was not in fact satisfied, which caused us to correct an initial judgment that a trial was FDAAA-applicable.

3) The company provides a web link to a publication, where we had found no publication. As you might imagine, matching trial characteristics to publications is a difficult task, especially for phase 1 trials with no registration record, so we miss publications from time to time.

Response: We have provided further information about the validation process and our interactions with companies in the revised Appendix. We call out to the Appendix where we describe the validation method in the manuscript (p.8 top) and provide an example of a correction made in the validation process there.

As a side note, we recently received a late response to our request for feedback from one of the companies and as a result, have made a few small corrections to the data reported in our initial manuscript submission.

16. Research ethics: Not sure where IRB information is located.

Response: This study did not involve human subjects research. We have so indicated in the revised manuscript (p.8 middle)

17. References: you mention the Dickersin and Rennie 2003 JAMA article on trial registration, and there is a 2012 update too (not referenced). A lot happened in that time period, including FDAAA, so I would tend to refer to both. But I am biased.

Response: We thank the reviewer for the reference, which we have added (p.4 top).

18. PRISMA: Why did you use PRISMA, given that as you say this is a cross sectional analysis? I would think STROBE makes more sense.

Response: An editor at the journal recommended that we include the PRISMA checklist with our submission. It may have been based on a sense that our project was of a meta-research nature. We share the reviewer's sense that it does not entirely fit, but would need guidance from the Editor to move away from the first editor's recommendation.

19. There were a number of things I did not understand in this study and it would be good to explain them for people like me. First, why did you include phase 1 studies? As I understand it, the law does not apply to them, even under the Final rule, and so it seems odd to hold the companies to a standard that does not apply. If I am wrong, please be clear about this change and educate people like me!

Response: The reviewer is correct that FDAAA reporting requirements do not apply to Phase 1 studies. She suggests that trials in healthy volunteers should not be included in our analysis at all, which conflicts with Reviewer 2's suggestion that our rankings focus on the "all trials" rather than the "trials-in-patients" sample. We have preserved the focus of the original manuscript, but have underscored in the revised manuscript that our analysis looks at both compliance with FDAAA and meeting a more aspirational standard and that the all-trials analysis is relevant to the latter (p.7 middle).

20. Second, please explain in the paper that FDA approval is for a drug-indication combination and does not imply that the drug is effective or safe for other uses.

Response: Please see our response to comment #14 above regarding the "drug-indication combination" point. Respectfully, we do not think it is necessary or relevant to point out in this paper that FDA approval does not imply that the drug is safe or effective for unapproved indications.

21. Third, could showing your findings beforehand to the companies seem to represent a conflict of interest to some people? I see why it was a "good idea" for your study, but it could also lead to a perception of working together rather than this being a true "scorecard".

Response: We are certainly sensitive to the need to maintain the independence of our scoring, and aware that our practice of allowing companies to review preliminary data for accuracy for their drugs might raise eyebrows in that regard. For our Scorecard to have credibility, however, being accurate is just as important as being independent. For that reason, it is critical to allow companies to highlight potential errors in our data. All company comments are discussed by the research team. We declined to revise our results based on many of the comments; however, the few that raised important data issues to our attention, and that could be verified by our public data sources, are addressed. In response to comment #15 above, we provided three examples that capture much of the range of feedback we received.

Because more than one reviewer raised this issue, we have added to our Appendix additional details about the process of eliciting and responding to companies' feedback.

22. Fourth and finally, could you explain (assuming that this is what you mean) how you determined that CSR synopses are as good as CSRs. In our experience they really are not, but maybe this is a bigger sample.

Response: Please see our response to comment #1 above. It may be helpful to recall that we are not using CSR synopses as a replacement for CSRs, but rather as an alternative for reporting results in a registry.

23. I am confused about timing of the Final Rule and your cohort of trials. Can you explain this for us?

Response: The Final Rule does not apply to trials with primary completion dates before January 17, 2017. We excluded trials that were ongoing at the time of our study and trials that had not been completed for at least 1 year by our cutoff date, February 1, 2016.

24. Abstract: Is this article one of a series with a semi-annual update? Please explain.

Response: Please see our response to comment #12 above. Due to space limitations, we can't incorporate an explanation into the abstract.

25. p.2 Line 12 Please explain that FDA approval is for a drug-indication combination and does not imply that the drug is effective or safe for other uses (ie, off label or other on label uses).

Response: This comment repeats comments #14 and 20 above. Kindly refer to our earlier responses.

26. p. 2 line 38: not all patient trials, just those in FDA reviews

Response: The reviewer is referring to the sentence in the abstract reading, "…2 of 11 companies disclosed all patient trials and complied with legal disclosure requirements." We have clarified this line so that it now reads, "…all patient trials in our sample…".

27. p. 2 line 39: What standards are you referring to?

Response: As explained above (comment #4), we were referring to transparency standards and have made this clearer.

28. p. 3 Strengths and limitations: CSRs are not public documents in the US; you are looking at on label compliance

Response: We were unsure what the reviewer is driving at with the comment about CSRs. It is true that CSRs are not required to be public documents in the US. But we looked at CSR synopses, which some companies have made available on the internet. They aren't legally required to do so, but that isn't really germane to our analysis. Publishing results (in the CSR synopsis or an alternative venue) was part of our aspirational standard, not our measure of legal compliance.

To respond to the reviewer's comment about on label compliance, in the revised manuscript we have clarified that our trial sample was limited to those trials that supported the indications for which the drug was approved, that is all trials in approved NDAs (please see our response to comment #15 above). We have also added this to the "Strengths and limitations of the study" page and the limitations section of the manuscript (p.22 middle).

29. p. 3 line 11: You certainly searched a lot, but I would be careful not to claim you searched every public registry (for example, you did not search 2 in China that I know of).

Response: We take the point that we can't rule out having missed a registry. We have revised the "Strengths and limitations of the study" page and p.5 of the manuscript to state that we searched 39 registries rather than that we searched "every public trial registry". Notwithstanding, we likely captured the relevant registries for our sample because companies had the opportunity to inform of us if they registered trials in a different trial database.

30. P. 4 line 11: What's "it"?

Response: We have clarified this sentence to read, "transparency around clinical trial results does not always occur."

31. p. 4 line 25: What do you mean "under one study"?

Response: The lines to which the reviewer refers are as follows: "Moreover, studies that measure the transparency of trials, drugs, drug manufactures, and research sponsors often use markedly different transparency measures and standards, yielding different findings and progress reports. A research sponsor, trial, or drug may look transparent under one study and opaque in another."

We think it is clear that "one study" refers to the "studies" mentioned in the immediately preceding sentence. However, we have changed the wording to "in one study."

32. p. 5 line 11: what are the standards? I am confused by p. 5 line 25: It seems to imply that phase 1 studies in healthy volunteers are now included under expanded standards. Is this so? I don't know about this. Even if phase 1 studies are now included, don't the expanded rules apply starting in 2017?

Our language on line 25 was confusing, and we have revised it (p.5 middle and p. 7) to more clearly refer to the aspirational transparency standard that we defined.

The "line 11" the reviewer refers to is as follows: "These efforts have helped foster a culture of transparency in research, but they have also introduced ambiguous, and at times conflicting, standards." It follows a paragraph that lists a number of different sources of transparency standards, including FDAAA and NIH requirements. Therefore, it seems clear that the word "standards" refers to those sources. We are not sure whether the reviewer is asking us to describe the content of each of those standards, but that would significantly expand the introduction, does not seem germane to our analysis (we do describe the standards that are applicable in our analysis, i.e., FDAAA), and would disrupt the flow of this brief introduction to the paper.

33. p. 5 line 48: I am a little worried that you are referring to "validation" when you mean checking with pharma to make sure that what you are saying is correct. Perhaps you should report what they asked you to change! You say later on (page 8 lines 8-13) that the feedback from companies was generally incorporated "if it could be validated through our public sources" which sounds both complex and is a little vague!

Please see our response to comments #15 and 21 above and the new section we have added to the Appendix providing more detail about our interactions with companies. We believe the example we have added on p.8 (middle) helps clarify what we mean by validating through public sources. It reads, "For example, companies in some cases provided a web link to a publication missed in our matching process."

34. Page 6: State that you only included trials for "on label" indications. I wonder if looking at a sample of trials, one that includes drug companies that include the largest, medium sized and smallest, would give you the same answers?

Response: We have stated this in the revised manuscript, as described above in response to comment #28. We are not able to speculate about whether the scores would change had we included a different mix of companies. Future scorecards include reviews of medium and small sized companies (within the next three years).

35. Page 7 line 4: what is a primary completion date? CT.gov says it is defined as: Definition: The date that the final participant was examined or received an intervention for the purposes of final collection of data for the primary outcome, whether the clinical study concluded according to the pre-specified protocol or was terminated. In the case of clinical studies with more than one primary outcome measure with different completion dates, this term refers to the date on which data collection is completed for all of the primary outcomes. Once the clinical study has reached the primary completion date, the responsible party must update the Primary Completion Date to reflect the actual primary completion date.

Response: The term “primary completion date” is sui generis to the US context, and specifically to ClinicalTrials.gov, so the definition the reviewer provided is correct, and the one we used. In the revised manuscript (p.7 top), we have clarified that the primary completion date was defined as the date listed in that field in ClinicalTrials.gov.

36. Page 7 line 13: My understanding is that results reporting applies to trials done for approved drugs but off-label indications. Is this true? How did you find these? I would expect this would be a lot harder to do and “compliance” would be a lot less. Did you find this?

Response: Results reporting applies to all trials of FDA-regulated products, regardless of whether they relate to on-label or off-label uses. However, our Scorecard is focused solely on the trials supporting initial FDA approval, as we have clarified in the revised manuscript. Trials for different, as yet unapproved, indications are used in the initial NDA only for safety purposes. We assess those trials (for different indications) under the “all trials” and “patient trials” analyses. They are not included in our FDAAA analysis because we are focused on the trials supporting evidence of efficacy for the approved indications and because the current law (FDAAA) lacks the specificity to require compliance for unapproved indications, so these cannot be assessed ex ante. At any rate, the small number of trials to which this question pertains were generally excluded from our analysis, in practice, because they were ongoing or not completed by our study cutoff date.

37. Page 8 lines 18-30: Not sure about some of your language and assumptions here. (1) what is a “trial” (line 18)? Please make it clear how you know that CSR synopses tell you something similar to a CSR. I think you are using the terms interchangeably later on (eg, line 44). In my experience, CSR synopses are pretty uninformative, but maybe this is only true sometimes. Does “available” mean useful? For example, missing outcomes may be such that although the results are “available” they are not terribly useful.

Response: We have revised the sentence to which the reviewer refers (“line 18”) to clarify that we are referring to the trials in our samples (p.8 bottom).

With regard to the CSR vs. CSR synopsis issue, please see our response to comment #1 above.

38. Page 8 line 51: You use the terms “metric” and “measure” here and elsewhere to refer to items (I think), a term which might be easier to define without bumping up against other meanings.

Response: As discussed above in response to comment #8, we have changed our language throughout the paper to be more consistent.

39. Page 8 line 55: here and elsewhere why is “all trials” hyphenated?

We intended to use hyphenation only where “all-trials” modified a noun, but the reviewer helpfully points out that we were not consistent about it. We have corrected this in the revised manuscript.

40. Page 9: the page is a single paragraph. Perhaps it can be divided by topic? Or use bullets that correspond to Table 1?

Response: We agree that this paragraph was too long, and have broken in into two paragraphs in the revision.

41. Page 9 line 15: You refer several times to a meeting your group convened in August 2016. Is this reported somewhere? Who came, what happened? This reference is tantalizing but not terribly informative as is.

Response: We have implemented this useful suggestion by adding a new section to the Appendix that describes the Roundtable meeting. We call out to the Appendix in a new parenthetical on p.11 (bottom).

42. Here are some word questions: Line 15: disclosing results “only for trials enrolling patients”? When are you looking at the others?

Response: The revised methods section makes is clearer exactly how we analyzed the 2 separate samples, trials in patients and all trials (including healthy volunteers). The line to which the reviewer refers isn’t describing our analysis, it’s simply stating that companies have only committed to sharing results for trials in patients.

43. Line 20: What are “patient trials”?

Response: We share the sense that this term is ambiguous, and have used the term “trials in patients” consistently in the revised manuscript.

44. Line 22: How can you rank or analyze “transparency”? Transparency must be defined for each situation where it is used.

Response: We have clarified that we are talking about applying the transparency standards we have defined (p.9 bottom).

45. Line 25: sorry, but I am not sure why “renal impairment” is here?

Response: Please see our response to comment #3 above.

46. Line 32: Are you talking about registration or completion, in terms of confusion about FDAAA requirements? I thought the final rule clarifies what is required. If this is not the case because the final rule was not in effect for these trials, remind us! There should not be confusion any more.

Response: The Final Rule does clarify what is required, but it did not apply to our sample of trials. There was indeed substantial confusion about the requirement before the Final Rule was promulgated, even though reading FDAAA now, in light of the Final Rule, would lead one to conclude that certificates of delay were always required. For instance, there is an old NIH PowerPoint presentation on the internet indicating that certificates were never required. We have clarified that the

Final Rule didn't help resolve the ambiguity for our sample of trials in the revised manuscript (p.10 middle).

47. Table 1: I wonder if the table could be improved by putting in a column before the transparency "measure"? The column would refer the issues that the transparency "item" addresses. The column titled "Trials subject to FDAAA" could be defined in a footnote.

Response: We have added the footnote as requested and retitled Table 1 to refer to "measures" rather than "benchmarks".

Because each measure is defined clearly (for example, "published in a journal indexed in PubMed, Google Scholar, or EMBASE by 13 months post FDA approval") we have not added another column explaining what issues each of the measures address.

48. Page 11 line 10: Excel was used. Maybe it would be good to move to a database (as opposed to a spreadsheet).

Response: We thank the reviewer for this suggestion, which we will consider for future analyses. As our Scorecard becomes more extensive in the measures it includes, we agree this would be helpful. An advantage of the Excel format we have used to date is that the calculations and data are very transparent at a glance to anyone who might care to look at them.

49. Page 11 lines 14-22: see earlier comments on validation and meeting.

Response: Please see our responses to comments #15, 21, and 33, above.

50. Page 11 lines 39-49: I am not sure how a median of 7/22 trials works. Is it a median of 7 and a median of 22? I can't quite picture the proportion….

Response: The statement to which the reviewer refers is "A median of 7 out of 22 trials, per drug, were conducted in patients." This means that across all drugs, a median of 22 trials were conducted and a median of 7 of those were conducted in patients. We agree that it was in artfully worded, and have reworded it as follows: "A median of 22 trials were conducted per drug. A median of 7 trials per drug were conducted in patients." (p.12 middle).

51. Page 11 line 53: Your standard that results of trials in patients should be publicly available is not in Table 1 as stated. Maybe it is there but wording is different.

Response: The wording was indeed confusing. We have clarified it (p.12 middle) as follows: "We first report results for the sample of trials in patients."

52. Page 12 lines 13-18: Confession: I really am not that interested in phase 1 trials, since they are not required to be disclosed, are they? Some are in patients (eg, cancer trials) and some not. So I am still unsure of the point here. I feel embarrassed to ask, but what is your point?

Response: This sentence alluded to the fact that about half the undisclosed trials were phase 1 trials, which many people (like the reviewer) may not care about too much. But this point was not important, and we have removed it.

53. Page 12 Lines 25-35: I agree about a standard that all trials in a successful NDA should be publicly available (I thought they had to be whether the NDA was successful or not). These numbers are very low!

Response: We do not perceive that this comment calls for a response, though we have addressed the issue above in response to comment #12.

54. Page 12 lines 37-44: I am confused. Why should it matter that at least one was publicly available, if all should be disclosed?

Response: The reviewer may have misread the sentence, which reads, “All 19 drugs had at least one publicly unavailable trial conducted in patients or healthy volunteers.”

55. Table 2: if drug names are not the brand name, please don’t use capitalization. Are you saying you have 2 ethics standards or five?

Response: We believe the reviewer’s first sentence here is referring to column 1 of Table 2. This column does indeed present brand names, so no changes to capitalization are needed.

We have clarified in the revised manuscript that each “standard” we scored drugs on is composed of several individual “measures”. So, for example, “Ethics Standard B” has 4 component measures. Please see our response to comment #8 for additional information.

56. Page 14 Lines 11-15: I am confused by the trial that was registered “a few days late.” Is this worth mentioning? Maybe put it into one bucket of the other, and say in the discussion (if you want) that sometimes being late was a matter of a day or two.

Response: The reviewer is referring to 2 sentences describing why we counted one trial as compliant although it was registered a few days past the deadline. We agree that this text was unnecessary because we had already stated that we counted trials as meeting a deadline if they were less than 7 days late, “to account for delayed postings, weekends, holidays, time zones.” We have removed these 2 sentences (p.15, top).

57. Table 3: You mention the interpretation differences for trial completion, but ClinicalTrials.gov must have an opinion. Use that one. You have to make a decision. I don’t know enough to do this and am relying on you.

Response: The issue to which the reviewer refers is that there are 2 different ways to interpret FDAAA’s requirements, both of which are reasonable. The Final Rule should clarify the confusion for FINAL RULE applicable trials. However, as stated, our analyzed trials pre-date the Final Rule.

We considered the two interpretations carefully when calculating our company rankings, which did require us to make a decision on which interpretation to include. In the revised manuscript (p.10, middle), we have clarified which we used in calculating rankings—the one based on approval date.

We choose this interpretation (the approval date one) because it is now the general consensus interpretation for FDAAA since the passing of the Final Rule, confirmed by several law firms (Like Goodwin Procter) and NIH Staff at conferences. Unfortunately, our sense is that ClinicalTrials.gov’s interpretation of FDAAA on its website does not clarify the ambiguity in FDAAA. Current wording on its website still appears to hew closer to the other (first) interpretation, because it mentions the certificate of delay. However, NIH staff presentations at conferences contradict this.

58. I have to admit I am not so impressed by these transparency scores, given previous comments. It might be interesting to see on Table 5 how these scores correlate with the number of drug-indication approvals and also how much they are making per drug! I like Appendix 5 and wonder if it could

substitute for one of you other tables, assuming you decide to delete phase 1 studies or reduce the number of sensitivity analyses you show.

Response: Particularly because the reviewers do not agree about which samples and analyses are most important or useful to feature, we have not deleted phase 1 studies or reduced the number of samples for which we describe results. Given that, we are not able to add tables to the manuscript, although we are pleased the reviewer finds the Appendix helpful.

We are not able to provide correlations between the drugs' scores and how much money the companies make on each drug, as this is beyond the scope of the Scorecard. Although we could potentially add information about the number of indications for which each drug was approved, we feel this would add further complexity to an already complex presentation and is not absolutely necessary in order to convey the key findings of our analysis.

59. Discussion: Do your results mean that everything is fine? My comments in this section mainly relate to things I have already said. Either you will explain your decisions or make the changes so I won't say much more, except that you will need to check for statements in the Discussion after you decide what to do about the points I made earlier.

Response: We appreciate the reviewer's comment and in many ways have deliberately avoided editorializing about our results. We have clarified in the discussion our overall conclusion: the results are encouraging if you look at the trials-in-patients sample, but there are much lower rates of transparency on our measures when you look at all trials, including those in healthy volunteers. In the discussion, we summarize both sides of the debate about whether transparency for trials in healthy volunteers is really important. We believe reasonable readers could disagree about that issue, and thus, about whether current levels of transparency are adequate. Our own conclusion is that "Among large pharmaceutical companies, clinical trial transparency is high based on many measures, although opportunities for improvement remain."

60. P 20 line 20: you say that NIH policy requires disclosing information about all trials. But this paper is about drug companies so why are you looking at all trials (including phase 1)?

Response: We have addressed this issue above (e.g., in response to comment #19). Although our project is ultimately about ranking drug companies, those rankings derive from an analysis of individual drugs and the individual trials that supported approval of those drugs. Thus, we had to make decisions about which trials to include. We've done the best we can to explain those decisions and show how things look differently if we focus on a different sample of trials.

61. P 20 line 30: You say that disclosing phase 1 trials may help speed innovation and save money. Could your paper examine this question?

Response: Our empirical analysis can't address the question of whether and how disclosure of phase 1 trial results could help speed innovation, but we've explained why it may help speed innovation on p.21 (bottom).

62. Again, I get confused when you use terms such as "metrics" which has a meaning to me related to outcomes assessed in trials. Any way you can use a different word?

Response: Yes. Please see our response to comment #8 above.

63. P 22 line 54-56: Can you be more specific about where your data will be available? Seems odd to write a paper about transparency and to be vague about where your data will be publicly stored.

Response: We have revised the data sharing statement to be more specific about where our data can be found.

Reviewer: 4 (Capuano)

64. There exist a lack of details in regards of the operational framework where the study was conducted (does a data extraction form exists? more information on the collaborators that evaluated trials should be provided).

Response: Research assistants are listed by name in the Acknowledgements section of the paper (p.23 bottom) and all material researchers are listed as authors on this paper. We added affiliations next to the names of the research assistants listed in the acknowledgement section (p.23 bottom). We also included a revised Appendix containing a substantial amount of additional detail about our data-collection and analysis methods, including the qualifications of the data collectors. We also noted in the methods section that the researchers used a standard form for data-collection and listed extracted variables (ie trial characteristics) in Appendix 2.

65. In the abstract: the objective should contain a more straightforward research question. In particular, the statement “we defined…” should be removed.

Response: We have revised the Objective statement in the abstract so that it simply reads, “To define a series of clinical trial transparency measures and apply them to large pharmaceutical companies and their 2014 FDA approved drugs.”

66. In the study design section, it should be emphasized the descriptive nature of the analyses. It could be hard to understand the meaning of outcome measures in the abstract before reading the full-text. I will suggest a more elegant way of presenting them in order to improve the reading flow. The conclusion in the abstract is different from those written in the conclusion section of the full-text and is to some extent speculative. I will suggest a “milder” statement because transparency was high only for clinical trials involving patients, and only for some specific companies.

Response: We have made the following revisions to the abstract to respond to this group of comments:
• Added the word “descriptive” to the Study Design section.
• Improved the Conclusions section so that it better tracks our results and the presentation of conclusions in the main text of the manuscript.

We agree that readers cannot get a full understanding of how our outcome measures were defined by reading only the abstract. This is not uncommon, and not fixable within the word limit of the abstract.

67. Strengths and limitations of this study: “we added a company ranking…” this is not a strength of this study and should be removed. I strongly discourage the authors to include this table in the final version of this manuscript because it is highly speculative.

Response: The creation of company rankings is at the very heart of the Good Pharma Scorecard project. We believe strongly, based on 9 years of engagement with nearly 25 large pharmaceutical companies, and many years of researching the promises and pitfalls of these types of quality improvement initiatives, that it is widely publicized company rankings that will move companies to change their behavior. Our earlier paper in BMJ Open reporting our first set of company rankings attracted broad media attention and greatly sharpened companies’ engagement in the Good Pharma

Scorecard project. Far from being “speculative,” our company rankings are based on actual data and measures that we have described in this paper and our earlier article. Respectfully, we do not agree with the reviewer’s suggestion and have preserved this Table in the manuscript.

68. My major concern is that due to study design, the generalizability of results to a non-US audience could be hard. This aspect should be emphasized throughout the manuscript.

Response: We are unsure as to whether the reviewer is saying our methodology could not be applied to drugs approved by non-US regulators, or that non-American readers would for some reason have trouble understanding or replicating our analyses. We do not believe either is the case (indeed, in ongoing work we are incorporating approvals by the European Medicines Agency, without difficulty), and would need additional information to be able to respond to the reviewer’s comment. Further, because the majority of new drugs are first approved by the FDA and all information used for the Scorecard is available to the public or the profession (via biomedical journals), we expect our findings have implications for the field globally. Additionally, the trials in our sample are conducted in many different countries.

69. Data sources – page. 6 line 23-27: Authors should submit for peer-review the information obtained through personal communications with drug manufacturers.

Response: In general, the correspondence consisted of us sending the companies their raw data in an excel spreadsheet and them correcting any flawed data in the spreadsheet and returning the spreadsheet. We then used a rigorous standard process to adjudicate on whether to incorporate any proposed changes. The process included needing to be able to verify suggested data changes with our other public sources. We made both the pre and post validation scores public in the paper. The pre data sets are those we compiled for the study and the post data sets include any changes made through the company validation process. Very few changes were made to data.

70. Sampled Clinical Trials – page. 6 line 51-52: Regarding the information extracted for each clinical trials, do authors provided a data extraction form to their collaborators? If yes, the data extraction form should be provided in order to improve reproducibility. In fact, in appendix 2, there was no formal definition for each of the information that was required to be extracted for each clinical trial.

Response: As explained above in response to comment #64, we have included in the revised Appendix a list of all the data fields that were included in our data-collection spreadsheet. As we explain in the Appendix, many of the variables came directly from ClinicalTrials.gov and we used the NIH data dictionary. In the manuscript and Appendix, we also refer interested readers to the additional information about study methodology contained in our earlier BMJ Open article.

71. Sampled Clinical Trials – page. 7 line 34-37: “trials with unknown…” this statement should be moved to the results section.

Response: We have moved this sentence to early in the Results section (p.12 middle).

72. Outcome measures – page. 10 line 18: “we excluded trials (N=12)…” More details should be provided for these trials and should be provided in supplementary material.

Response: These 12 trials were cases in which the NDA holder (the unit of analysis for our company rankings) was not the company that conducted the trial, but later acquired the company that conducted it. In such a situation, the NDA holder is not the one responsible for satisfying the reporting requirements. We give several examples of these situations in footnotes to Table 2 – for example, “Novartis states that the unregistered trial ‘was conducted before the Novartis position to register all

phase I trials in patients became applicable to Alcon. Alcon was purchased by Novartis in April 2011." We doubt that readers will be interested in going beyond this, but would be happy to provide a comprehensive list of the circumstances surrounding all 12 trials if the Editors disagree.

73. Transparency score based on patients trials – page. 11 line 51: Results based on all trials should be provided prior to those involving only patients in the text. Moreover, it should be emphasized that clinical trials involving only patients represent only a small fraction of the overall trials. In fact, the transparency scores for the "all trials" are dramatically lower than trials involving only patients. This aspect is not emphasized appropriately throughout the manuscript. Therefore, the manuscript should be modified in order to provide this message to the reader.
Firstly, I expect in the text at least some examples of the proportion of trials involving patients with the overall amount of trials for each specific drug.
In example: Harvoni® have in total 60 trials analyzed and only 31 (51.7%) involved patients. The transparency parameters for all trials are 53%, 32%, 57% and 40% for %registered, %reported, %published and %publicly available respectively. For trials involving only patients instead, these tend to be higher (92%, 61%, 68% and 74% for %registered, %reported, %published and %publicly available respectively).
This aspect should be emphasized also in:
1. Transparency Scores Based on All-Trials in an NDA, Including Healthy Volunteers section;
2. Compliance with FDAAA Requirements section;
3. Timing of Results Reporting section;

Response: It is not the case that trials in patients comprise "only a small fraction of the overall trials". They account for 233 of the 505 trials. They also account for 93% of trial participant data, as trials in healthy volunteers tend to have few enrollees. We have added this information to the revised manuscript (p.12 middle).

To add to our discussion above of why we have chosen to emphasize the trials in patients rather than the all trials sample, we would note that the difference in results across the two samples is very prominently discussed in the Discussion section. We devote more than a page of text to it, beginning with, "The gap between transparency of results from the all-trials and trials-in-patients samples is striking. The median proportion of trials available per drug was markedly lower in the all-trials sample …." However, we have made some edits to the discussion to make doubly sure that readers do not miss it—including moving this paragraph up to be the second paragraph in the Discussion (p.20 middle). As the reviewer suggests, we have included an example of the difference in scores for one drug (p.20 middle).

The reviewer also asks that we add commentary on the difference in results between the two samples earlier in the paper, when we present the descriptive results for the different analyses. We would like to minimize duplication between the Results and Discussion sections, but have added brief comparative statements to 2 of the subsections within Results as requested (p.12 bottom, p.18 top). We did not add them to the subsection on FDAAA compliance because the critique does not actually apply to that sample (the FDAAA compliance sample differs from both the all-trials and trials-in-patients samples).

74. Company Ranking – page. 18 line 20: I have some concern. This table could provide a misleading message. For example, for FDAAA trials metrics, Sanofi has only 3 trials while other companies have more trials (therefore, different probabilities of having transparency-related problems). This aspect should be emphasized.

Response: We discussed at various points in the project the point the reviewer raises here: it's easier to be compliant when you've conducted only a handful of trials than when you've conducted a large

number of them. We ultimately decided not to weight the rankings by volume of trials. There is no right or wrong way to do it, but the reviewer's comment reminds us that the issue is worth mentioning in the paper. We have added it to the revised manuscript on p.23 (bottom).

VERSION 2 – REVIEW

REVIEWER	Joel Lexchin York University Canada In 2015-2017 Joel Lexchin received payment from two non-profit organizations for being a consultant on a project looking at indication based prescribing and a second looking at which drugs should be distributed free of charge by general practitioners. In 2015 he received payment from a for-profit organization for being on a panel that discussed expanding drug insurance in Canada. He is on the Foundation Board of Health Action International.
REVIEW RETURNED	23-Sep-2017

GENERAL COMMENTS	The authors have addressed my initial concerns. There is one remaining relatively minor point that needs to be corrected. In the Highlights the authors say "we obtained NDA holders' feedback on their results and convened 18 companies to discuss methods and root causes of transparency performance." This statement makes it sound as if the feedback and meeting was part of this study but in the Methodology the authors only vaguely refer to an August 2016 meeting and don't mention the number of companies attending. Furthermore, there is nothing about this meeting in the Results.

REVIEWER	Matthew Herder Dalhousie University, Faculties of Medicine and Law, Canada
REVIEW RETURNED	04-Oct-2017

GENERAL COMMENTS	My comments to the authors are very minor in nature: 1. Page 5, line 27 - the phrasing "This paper and efforts are a continuation of efforts..." reads awkwardly. Suggest removing 'and efforts'. 2. Page 6, line 39 - It would be helpful to have a little information in the main text about how company subsidiaries were identified. 3. Page 6, line 48 - The word 'will' appears to be missing. "our scorecard [will] expand..." 4. Page 10, lines 27-32 - it would be helpful to know what the Final Rule actually stated/clarified. (There is also a vague statement about the 'Final Rule being helpful' in the Discussion section (p. 22, lines6-8), which should be explained more clearly. What position did the Final Rule take?) Finally, I would like to encourage the authors to consider undertaking in future studies an evaluation of compliance with these transparency standards for drugs that do not obtain FDA approval. Arguably, the need for, and potential benefits of transparency, are

	even greater for unapproved drug indications. The authors discuss several other lines of inquiry for future research (e.g. data transparency, focusing on small and medium sized companies). Given, however, that companies in the present study demonstrate stronger compliance with the 'at the time of approval' interpretation of the FDAAA, I think it would be very telling to examine compliance for unapproved drugs in a given year that have, for instance, reached the stage of a phase 3 clinical trial. This comment is meant simply as a suggestion. The current study nicely builds upon previous research by these authors and documents the slow progress toward meaningful transparency standards in pharmaceutical research.

REVIEWER	Matthew Herder Dalhousie University, Faculties of Medicine and Law, Canada
REVIEW RETURNED	04-Oct-2017

GENERAL COMMENTS	My comments to the authors are very minor in nature: 1. Page 5, line 27 - the phrasing "This paper and efforts are a continuation of efforts..." reads awkwardly. Suggest removing 'and efforts'. 2. Page 6, line 39 - It would be helpful to have a little information in the main text about how company subsidiaries were identified. 3. Page 6, line 48 - The word 'will' appears to be missing. "our scorecard [will] expand..." 4. Page 10, lines 27-32 - it would be helpful to know what the Final Rule actually stated/clarified. (There is also a vague statement about the 'Final Rule being helpful' in the Discussion section (p. 22, lines6-8), which should be explained more clearly. What position did the Final Rule take?) Finally, I would like to encourage the authors to consider undertaking in future studies an evaluation of compliance with these transparency standards for drugs that do not obtain FDA approval. Arguably, the need for, and potential benefits of transparency, are even greater for unapproved drug indications. The authors discuss several other lines of inquiry for future research (e.g. data transparency, focusing on small and medium sized companies). Given, however, that companies in the present study demonstrate stronger compliance with the 'at the time of approval' interpretation of the FDAAA, I think it would be very telling to examine compliance for unapproved drugs in a given year that have, for instance, reached the stage of a phase 3 clinical trial. This comment is meant simply as a suggestion. The current study nicely builds upon previous research by these authors and documents the slow progress toward meaningful transparency standards in pharmaceutical research.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1 Comments:
The authors have addressed my initial concerns. There is one remaining relatively minor point that needs to be corrected. In the Highlights the authors say “we obtained NDA holders’ feedback on their results and convened 18 companies to discuss methods and root causes of transparency performance.” This statement makes it sound as if the feedback and meeting was part of this study but in the Methodology the authors only vaguely refer to an August 2016 meeting and don’t mention the number of companies attending. Furthermore, there is nothing about this meeting in the Results.

Response to Reviewer 1:
Thank you for this thoughtful comment. As the meeting did not yield findings reported in the results section of this paper, we removed references to the meeting in the body of the manuscript. We left one reference and explanation about the meeting in the appendix per the request of previous reviewers. This can also be removed, per editorial request.

Reviewer 4 Comments:
Dear Authors, This is an excellent study on data transparency in clinical trials.
The subject matter is of relevant importance and the paper is methodologically well structured.
Tables are very helpful and you provided a useful insight of strength and limitation. I would definitely encourage the continuation of such studies also for a longest time period.

Response to Reviewer 4:
Thank you.

Reviewer 2 Comments:
My comments to the authors are very minor in nature:
1. Page 5, line 27 - the phrasing "This paper and efforts are a continuation of efforts..." reads awkwardly. Suggest removing 'and efforts'.
2. Page 6, line 39 - It would be helpful to have a little information in the main text about how company subsidiaries were identified.
3. Page 6, line 48 - The word 'will' appears to be missing. "our scorecard [will] expand..."
4. Page 10, lines 27-32 - it would be helpful to know what the Final Rule actually stated/clarified. (There is also a vague statement about the 'Final Rule being helpful' in the Discussion section (p. 22, lines6-8), which should be explained more clearly. What position did the Final Rule take?)
Finally, I would like to encourage the authors to consider undertaking in future studies an evaluation of compliance with these transparency standards for drugs that do not obtain FDA approval. Arguably, the need for, and potential benefits of transparency, are even greater for unapproved drug indications. The authors discuss several other lines of inquiry for future research (e.g. data transparency, focusing on small and medium sized companies). Given, however, that companies in the present study demonstrate stronger compliance with the 'at the time of approval' interpretation of the FDAAA, I think it would be very telling to examine compliance for unapproved drugs in a given year that have, for instance, reached the stage of a phase 3 clinical trial.
This comment is meant simply as a suggestion. The current study nicely builds upon previous research by these authors and documents the slow progress toward meaningful transparency standards in pharmaceutical research.

Response to Reviewer 2:

1. We removed “and efforts” from, Page 5, line 27.
2. We added information in the main text about how company subsidiaries were identified. Generally, we used information from company websites, press releases, SEC filings and communications w companies. See Page 6, line 39.
3. We added the word 'will' to Page 6, line 48.
4. We added a statement about what the Final Rule clarified in regards to FDAAA. See Page 10, line 27 and Page 22, lines 6-8. Essentially, the Final rule newly requires that trial results be reported for both approved and unapproved indications.

Thank you for the suggestion to examine compliance for unapproved drugs in a given year that have, for instance, reached the stage of a phase 3 clinical trial. It is a good suggestion that we will consider adding.