Author’s response to reviews

Title: Reporting quality of randomized controlled trial abstracts in the seven high-ranking anesthesiology journals

Authors:
Katja Janackovic (katja.janackovic@gmail.com)
Livia Puljak (livia@mefst.hr)

Version: 1 Date: 29 Jun 2018

Author’s response to reviews:
Ana Kowark, M.D.
Trials

June 27, 2018

Re: Revision of the manuscript TRLS-D-18-00102, title: Reporting quality of randomized controlled trial abstracts in the seven high-ranking anesthesiology journals: a meta-epidemiological study

Dear Dr. Kowark,

hello, we are very grateful for the detailed comments we received from you and anonymous reviewers. We appreciate the time that was taken to assess our manuscript. We did our best to address the comments and revise the manuscript accordingly. Hereby you will find our point-by-point response to the comments. We highlighted changes in the manuscript using the ‘track changes’ feature in Word.

Editors comments:

Beside the other comments of the reviewers, the following two issues posed by the reviewers are more critical and require careful evaluation:
1. Please consider extending your analysis by looking through the journals themselves rather than use of PubMed! Why did you choose 7 journals and analysis only over 3 years?

Response: As suggested, we searched each journal’s web site to make sure that our search of PubMed gave us accurate list of trials in analyzed anesthesiology journals. We have now expanded our methods’ section to describe this additional search and revised results’ section accordingly.

Our decision to analyze 7 journals and time span of 3 years was based on our hypothesis that leading journals in the chosen field have reporting deficiencies when it comes to reporting of randomized controlled trial (RCT) abstracts. We took time frame of 3 years because we wanted to analyze current, i.e. more recent period. This was now clarified in the manuscript.

2. Please consider using a real statistical software.

Response: As suggested, we have now used MedCalc software. To be perfectly honest, we did not know that Microsoft Excel can yield erroneous statistics. Change was now indicated in the Methods.

Reviewer #1:

1. I am not too familiar with 'meta-epidemiology', but it seems to be an alternative to systematic reviews/meta-analyses for combining results from separate RCTs. I do not think you have done any meta-analyses, only descriptive analyses with some subgroups (i.e. percentages per journal and per CONSORT item), so I think the description is incorrect. If you think this definitely is a meta-epidemiology study, I would justify the description in the text and also define the term 'meta-epidemiology', with appropriate references, as I think it is quite a recent term and may not be familiar to everyone. In the text of the manuscript we used the expression methodological study instead.

Response: As described by Hassan-Murad and Zhang (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5537553/), meta-epidemiology, or research on research is a fairly new field of study where unit of analysis is a study, and not a patient. The outcomes of such study are usually not clinical outcomes. Instead, meta-epidemiological studies, which are also sometimes called methodology research, provide data for methodological analysis. Since using the term meta-epidemiology in the title is not crucial for us, we dropped the word from the title, so that it would not confuse readers that are not familiar with this term.
2. Consider rewording some of the second and third paragraphs of the Background section - I can see why you might like to say the issue of poor abstract reporting is a specific problem for anaesthesia studies, but so far I'm not convinced - it seems like a general issue, affecting all areas of research. So the section on pain relief being a basic human right seems out of place to me. But perhaps it can be improved and clarified.

Response: As suggested by the reviewer, we removed the section about pain relief being a basic human right.

3. More clarity required in methods: you mention inclusion criteria, but no exclusion criteria, though these are mentioned in the Results.

Response: As suggested, we expanded our section on inclusion criteria in the methods to also highlight the exclusion criteria.

4. Methods: did you consider that, given you were looking only through seven journals over three years, it might have been easier to just look through the journals themselves rather than use PubMed? More seriously, you report in the Results that there were some false positives in your searching, but is there a risk of false negatives? Did you look through at least some of the journal issues to check you had everything? It would be useful to have at least a comment on this issue.

Response: Searching individual journal web sites includes two search options, where one option is hand-searching through all eligible issues, and the second option is using a search box available on their web sites. Journals do not offer indexing option, to search for a specific type of manuscript. Instead, one would need to use keywords to search for the RCTs. For example, journal Anesthesia allows truncation and if a user types in random*, with the intention to search for any word that begins with random, the output (on June 14) is 395 manuscripts (limited to our targeted time frame). Many of the search results are not eligible because the journal browser returns entries that have a word random anywhere in the text. Journal Pain gives 3685 results after searching for random*, but when trying to limit the search for the time frame we analyzed, there is no simple option as offered by the Anesthesia journal. Pain journal gives option to search individual volumes, or pre-defined time periods that include: All Dates // Current Issue // Last 12 Months // Last 3 Years // Last 5 Years // Last 8 Years. The example of these two journals only indicate complexity of searching web sites of the journals. For this reason, we resorted to searching PubMed, which offers indexing of different manuscript types. But, we do understand that our methodology may have cost us perhaps losing some studies which were RCTs, but not indexed as RCTs on PubMed. We have indicated this now in the Discussion as a limitation of our study.
5. Methods, outcome: description of the primary outcome needs clarifying. I think your primary outcome measure was actually proportion of adherence to the CONSORT-A checklist within each article, presented as a median of all articles.

Response: We revised the primary outcome as suggested.

6. Methods: secondary outcomes also not described totally clearly here - you've used the word 'subcategories' here, but in Results you say 'individual items'; subcategories could mean 'Background', 'Methods' etc. Also you've said 'adherence to the checklist across the individual journals' - I think this could be clarified that you were looking only at overall adherence here, not within individual items.

Response: We have clarified now in the secondary outcomes that we analyzed overall adherence across domains of CONSORT-A, so that we would not give erroneous impression about what we aimed to do.

7. You say you have marked all the abstracts' contact details as Yes only if an e-mail address was given - but CONSORT-A doesn't ask specifically for this. I don't think you can justify evaluating abstracts against something beyond what CONSORT-A asks for (and you also acknowledge that CONSORT-A says this is only for conference abstracts). Even if an author followed CONSORT-A perfectly, they could still fail this test. I think you should consider removing this criterion from your exercise, as I assume you had no conference abstracts in any case. In Results, you say that email address was only given in 16% of abstracts - this can't really be criticized using CONSORT-A as it doesn't ask for it. This also needs addressing in the Discussion (paragraph 9).

Response: We acknowledge that the Reviewer is correct about this, and we tried to emphasize and explain our approach. Our rationale was the following: the purpose of contact details is to contact somebody. And in the current internet era, it is difficult to imagine that someone would write an actual letter to a corresponding author. Instead, people rely on emails to contact colleagues. Therefore, we hypothesized that nowadays contact details are basically useless, unless there is e-mail address available. However, it is definitely correct that this was not specified in the CONSORT-A. Therefore, we have now presented an additional result indicating whether there were any contact details (affiliation of a corresponding author) in the abstract on PubMed.

8. Results, first paragraph: did some of the results really have no abstracts? I find this amazing given you were looking for papers in high impact factor journals. What's the explanation?
Response: Indeed, some of the RCTs are being published as a letter to the editor. We also did not expect this, and therefore originally we did not use it as an exclusion criteria. But then we found multiple RCTs published without an abstract and thus we had to exclude them.

An example of such RCTs is a study of Klein et al. from 2016: https://www.ncbi.nlm.nih.gov/pubmed/26086287. It is not even acknowledged as an RCT in the title, but the letter to the editor is very clear about its methodology: “After written informed consent on the day before neurosurgery, patients (n=24) scheduled for routine craniotomy were included in the randomised crossover study.”

We have now indicated in the Methods why some RCTs were excluded because they did not have an abstract – with an explanation that those were published as letters to the editor.

9. Some other results are quite surprising and are not explained. For example, did 25 abstracts really not give Conclusions? I note that very little space in Discussion is taken up by detailed discussion of your results. After a short paragraph, you go on to discuss others’ works.

Response: We did find that some rare abstracts (4%) that did not have conclusion statements; instead, these abstracts simply list their results without final conclusions. An example is abstract from Dhir et al., published in the journal Anaesthesia, link: https://www.ncbi.nlm.nih.gov/pubmed/26566960. That particular abstract has 116 words, so we can speculate that, considering the word limit for Summary of that journal, that the authors did not have space to also provide conclusion statement. Based on the comment of a reviewer we have now analyzed those 25 abstracts in more detail and found that 18 of them were published in the journal Anaesthesia, which had requirements for summary of 150 words. The remaining seven were published in journals Pain (N=5), Anesthesiology (N=1) and European Journal of Anesthesiology (N=1). Beyond word limit, one potential other reason for the lack of conclusion statement is journal requirement for authors to write a non-structured abstract. If the authors are not required to use formal sections such as Aims, Methods, Results, Conclusions, then they may be more prone to organizing the abstract the way they want. One additional reason for the lack of conclusion statement could be insufficient academic writing skills. We added these consideration to the Discussion now.

10. Results suddenly present some statistics that were not mentioned in the Methods, such as medians. As mentioned elsewhere, I think it might be useful to present an interquartile range (probably instead of confidence intervals, which I don't think are too informative). Then there is a statistical test - Komogorov-Smirnov - it is not previously mentioned, it is not clear what this is a test of. How should we interpret the p-value given? The test is also not mentioned again in the paper. This needs some work.
Response: We removed mentioning of the Kolmogorov-Smirnov test, as suggested by the second reviewer, as we used count data, and therefore there should be underlying assumption that these are non-continuous data. As suggested, we presented interquartile range instead of confidence interval.

11. Is it worth statistically testing the difference between the different journal and reflecting more on the difference/lack of difference?
Response: We are not sure whether this additional analysis would add much to the results.

Minor comments:

1. Abstract - Results - second line reports an 'overall adherence of 41%' but actually this is the per-article median, and you should say so more clearly. Otherwise the reader might assume this is 41% across all articles, with denominator being number of articles x 17 (number of CONSORT items per article).
Response: We corrected this segment of Abstract as suggested.

2. Abstract and main paper Results - might it be informative to add the interquartile range with the median?
Response: As suggested, we added interquartile range to the abstract.

3. Background and Discussion, you refer to readers not having access to full text articles, or these 'not being publicly available'. Surely the issue is that they are available, just at a cost, which people may not be able/willing to pay. I think you could clarify this.
Response: This is a very valid point to emphasize; thank you. We have indicated this now in both Background and Discussion.

4. Background, line 5 should read 'Reporting of research manuscripts…'
Response: This was revised as suggested.
5. In most cases, 'CONSORT' should not have 'the' before it. A general rule to apply would be when it's followed by another noun, it should have 'the', but when it's on its own it should not. So 'the CONSORT extension' or 'the CONSORT statement', but 'CONSORT provided limited guidance for writing an RCT abstract.'

Response: This was revised as suggested, thank you.

6. Background, paragraph 2, final sentence should be 'For this purpose it is important to have clearly reported RCT abstracts…'

Response: This was revised as suggested.

7. Background, paragraph 3: use written 'seven', not number '7'

Response: We used the written word seven instead of number, as suggested.

8. Methods, Search, second paragraph, should say 'In case of doubt we downloaded full texts in electronic format to judge whether an article was indeed an RCT.'

Response: This was revised as suggested.

9. Methods, scoring the checklist, first line should say 'We designed a data scoring table in a Microsoft Excel spreadsheet…'. Add a full stop at the end of the sentence - currently missing.

Response: This was revised as suggested.

10. Methods, scoring the checklist, line three says 'One author…' - who did which role here?

Response: We have now indicated in brackets who did what. As a senior author and methodologist, LP verified extractions.

11. Methods, scoring the checklist, line four says a second author verified extractions - clarify, was this 100% of them?

Response: Yes, all of them. This is now clarified in the methods.
12. Methods, scoring the checklist, line six: '...to ensure a consistent interpretation of each checklist item' - do you find the checklist ambiguous? Is this something to comment on in Discussion and Limitations?

Response: Indeed, we found that some items could be ambiguous and this is why we clarified in our methods how we scored certain items – this section of Methods begins with “After calibration exercise and considering potential items …”. We found that items 2, 8, 9 and 11 could lead to different interpretations and different scoring. For this reason, we clarified how we will proceed with scoring these four items and we described this in the Methods. We have now also added a consideration about this in the Discussion.

13. Methods, scoring the checklist, paragraph 2, first line should say 'After the calibration exercise…'

Response: This was revised as suggested.

14. Methods, scoring the checklist, paragraph 2, third line: use 'second' not '2nd'

Response: This was revised as suggested.

15. Methods, scoring the checklist, paragraph 3, line six should say '...explicitly stated that the study was completed…'

Response: This was revised as suggested.

16. Results, secondary outcome: adherence with CONSORT-A - title needs amending to make clear this is by journal.

Response: This was revised as suggested.

17. Results, secondary outcome: adherence with CONSORT-A - the first sentence is not very useful without an actual number. Although I appreciate it's in a table, we might as well just look at the table instead of having this sentence. Might it be better to say something that adds to the table, like range or median or mean?
Response: As suggested, we expanded this first sentence about secondary outcome and added a text that describes the table in more detail. Our idea was not to repeat the table too much, and this is why we indicated in the second sentence that for all those domains adherence was under ten percent. We hope that the newly revised sentence will be adequate.

18. Discussion, line three - perhaps CONSORT-A 'recommends' rather than 'suggests' about RCT abstract reporting?
Response: This was revised as suggested.

19. Discussion, paragraph 2, line 7 should say '2.7% indicates that the 2008-2009…'
Response: This was revised as suggested.

20. Discussion, paragraph 3, line 2: I think it should say 'Cui et al showed that significant improvement was observed only in a few items…'
Response: This was revised as suggested.

21. Same line: you say Cui et al found some improvement in a few items - which? And did you also note that these same items had improved?
Response: We have now clarified which three items were improved and where exactly were these manuscripts published.

22. Discussion, paragraph 4, first line - it has been suggested...by whom?
Response: We have now specified that this was suggested by Hays et al, and also we explained for the next sentence who made the mentioned suggestion.

23. Discussion, paragraph 5, first line - 'In this study we found that the median total adherence score…' - make clear this is per article
Response: This sentence was revised as suggested.
24. Discussion, paragraph 5 - you say six out of seven journals in this study endorse CONSORT - do you mean CONSORT or do you mean CONSORT-A? Which one does not endorse CONSORT and did you notice this one was worse? Or were they similar? This could be a useful finding.

Response: The Reviewer is indeed right that the endorsement refers to CONSORT, and not CONSORT-A. Although not all the analyzed journals are listed on the CONSORT web site as endorsers, all of the seven analyzed journals mention explicitly CONSORT in their instructions for authors. But the CONSORT checklist says this in the domain about writing an abstract: (for specific guidance see CONSORT for abstracts). We have revised this fragment of Discussion accordingly.

25. Discussion, paragraph 6: 'funding may be associated with higher frequency of positive results' - I think I know what you're getting at, but all studies have funding, so you need to clarify this.

Response: The reviewer is right; this sentence was not clear enough. We clarified now that we refer to commercial funding.

26. Discussion, paragraph 6, line 5: amend to 'skews the picture', and remove the first of the two 'available'.

Response: These two changes were made, as suggested.

27. Discussion, paragraph 6, line 7, should say '...problems with selective reporting, there is now an imperative...'

Response: This was revised as suggested.

28. Discussion, paragraph 7, line 2 - should say '1.6% of the analyzed RCT abstracts'

Response: This was revised as suggested.

29. Discussion, paragraph 7, line 5 - you say 'reporting of this item did not improve', but the numbers indicate that it did, from 1.6% to 4%. You might clarify to 'did not improve substantially' or similar.
Response: The reviewer is right; it is an improvement, but not substantial. So we added a word substantially.

30. Discussion, paragraph 8, line 5 should say 'For example, the journal Anesthesia…'
Response: This was revised as suggested. We also changed verbs into past tense, as this was the past situation; journals may change their word limit requirements.

31. Discussion, paragraph 8, line 6 should say 'the journal Anesthesia and Analgesia…'
Response: This was revised as suggested.

32. Discussion, paragraph 9, line 1 should say 'second' not '2nd'
Response: This was revised as suggested.

33. Discussion, paragraph 10, line 1 should be: 'limitations of this study are the limited number…'
Response: This was revised as suggested.

34. Discussion, paragraph 11, line 4 should be 'In February 2017, the journal…'
Response: This was revised as suggested.

35. Discussion - why so much about Penelope? It might be worth a mention, but not several sentences, I don't think.
Response: We thought Penelope could be one of the solutions for the problem we described. But we definitely agree that the same message can be conveyed in less words, so we shortened the section in Discussion about Penelope.

36. Discussion, final paragraph, line 5, duplication in sentence, suggest removing 'Although poor reporting is not equal to poor methodology'.

Response: This sentence was shortened as suggested.

37. Discussion, final line should say 'deprives a reader of important information…'

Response: This was revised as suggested.

Reviewer #2

1. The methods are described well and the paper seems straightforward. However the ms is too long for a descriptive study and should be shortened quite a bit.

Response: Based on the suggestions of the first Reviewer, we added additional explanations that were suggested, and shortened text where it was suggested. We made some additional shortening. We would be happy to shorten the manuscript further, if we can kindly get specific suggestions for shortening the text.

2. were reviewers blinded to journal title?

Response: No, reviewers were not blinded to manuscript and journal title. We indicated this now in the methods.

3. did they do a sample size calculation?

Response: We did not do a formal sample size calculation for this study; we used a convenience sample of all abstracts that were published in all of the chosen journals in the analyzed years.

4. were the authors trained in critical appraisal? how were assessor assessments standardized?

Response: One of the authors, LP, is a trained methodologist. For ensuring consistency, we used the first cohort of scored abstracts to align our scores with the checklist. Whenever we saw that there were discrepancies stemming from different interpretations of CONSORT-A domains, we noted clarifications in the methods, and these clarifications were presented in the methods. Additionally, LP checked all extractions subsequently.
5. The investigators should use a REAL statistical package for descriptive statistics and tests they may wish to perform. Microsoft Excel is NOT an appropriate software tool for statistical analyses, no matter how basic. Excel’s statistical procedures have been shown to have serious errors; although many were identified back in 1994 it seems that they have not been fixed. Problems included serious rounding errors, missing values handled inconsistently & incorrectly; misleading estimates for variance, confidence intervals etc.; inconsistent in the type of p-values returned.

Response: We have now used MedCalc software, and indicated this in the manuscript.

6. There is no necessity to report percentage data to two decimal places; whole numbers will do.

Response: We reported two decimals only for our results with numbers under 10. We have now reduced this to one decimal. Otherwise, we would get misleading numbers if we round 0.16 to zero. When reporting others’ results we used exact percentages that they reported, but we have now rounded this to whole numbers.

7. It is not clear why they want to use a Kolmogorov-Smirnov test. This is not an hypothesis testing study! In any case, the K-S test is a nonparametric hypothesis test that estimates the probability that two sets of data are sampled from populations with identical distributions (the two-sample KS test) or a continuous model (the one-sample KS test). In both cases, the underlying population distribution is assumed to be continuous. This will not be the case here - these data are count data.

Response: We removed mentions of the Kolmogorov-Smirnov test, and indicated here that we assumed that the count data are not continuous.

8. The investigators might want to check out and reference other similar studies to provide a broader context e.g. Sriganesha, et al. Quality of abstracts of randomized control trials in five top pain journals: A systematic survey. Contemporary Clinical Trials Communications 7 (2017) 64-68

Response: We cited the reference that was suggested.
We hope that the revised manuscript and responses will be satisfactory.

Sincerely,

Livia Puljak and Katja Janačković