PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Anti–Vascular Endothelial Growth Factor Treatment for Retinal Conditions: A Systematic Review and Meta-analysis
AUTHORS	Tricco, Andrea; Pham, Ba; Thomas, Sonia; Lillie, Erin; Lee, Taehoon; Hamid, Jemila; Richter, Trevor; Janoudi, Ghayath; Agarwal, Arnav; Sharpe, Jane; Scott, Alistair; Warren, Rachel; Brahmbhatt, Ronak; Macdonald, Erin; Straus, Sharon

VERSION 1 - REVIEW

REVIEWER	Gianni Virgili Department of Translational Surgery and Medicine, University of Florence, Italy
REVIEW RETURNED	17-Feb-2018

GENERAL COMMENTS	This is a comprehensive and thorough piece of work. Given its strategic importance, this manuscript could be published as it is.

REVIEWER	Ravi Parikh MD MPH Harvard Medical School Department of Ophthalmology Massachusetts Eye and Ear Infirmary Boston USA
REVIEW RETURNED	19-Aug-2018

GENERAL COMMENTS	The authors should be commended on having a study with strong methods and a robust analysis/review of the literature. The major value of a study such as this would be to answer from pooled data how bevacizumab compares to aflibercept in AMD. The answer for DME and RVO is clearer as there have been head to head trials, but drawing a conclusion for AMD would greatly strengthen the paper. Essentially, everyone knows (anecdotally and from numerous head to head clinical trials) that bevacizumab and ranibizumab are clinically similar, but aflibercept is the point of debate as it has a slightly different mechanism as a VEGF trap (not a monoclonal antibody) and has been shown to decrease the treatment burden with longer interval dosing (see VIEW 1 and VIEW2 trials). Aflibercept is also much more expensive than bevacizumab so the controversy with anti-VEGF medications is should I use an expensive medication and if so what is the marginal utility of that. We know the answer between bevacizumab and ranibizumab. Along with shedding more light on the aflibercept vs. bevacizumab issue in AMD, addressing the below points would also strengthen a paper that is likely going to be an important reference for the field. Please note the pages and lines refer to the

	authors' word processing document not the combined PDF from BMJopen. 1. p4 line 62 what "few exceptions" are there where bevacizumab is not a reasonable alternative to ranibizumab and aflibercept? Also what are the authors' conclusions regarding the difference between ranibizumab and aflibercept and becavizumab and aflibercept? These should be stated. 2. p 9 line 169. The authors also excluded studies comparing anti-VEGF to photodynamic therapy (PDT) which is a mainstay adjuvant treatment for the polypoidal choroidal vasculopathy (PCV) subtype of nAMD which is very common globally in places such as in Asia, where the majority of AMD currently is. This should be mentioned. 3. p 10 line 186: Are the 22% that gained vision specifically referring to those who gained 3 or more lines of vision which was often used as the measurement in many of the studies such as the CATT trial? This should be clarified and how many gained 3 or more lines of vision should be explicitly stated as that is often the benchmark in many of these trials. 4. line 192 p 11: It is widely known that bevacizumab and ranibizumab have similar treatment effects but what has your study shown us about bevacizumab and aflibercept as that is often the clinically and public health wise important issue as the question comes down to using low cost bevacizumab vs. high cost aflibercept while use of ranibizumab is relatively dropping per the literature. 5. P 12 line 211: It is important to note that for mortality, PE, MI, etc none of the individual studies was powered to properly answer if one medication has a higher rate of any of the serious adverse events mentioned. 6. P 13 line 244: In the patients with DME section, the authors specifically mention information from DRCR Protocol T but the 2 other studies included (from Egypt and Turkey) are not mentioned. The data from each study and how that relates to the overall assessment of each medication for DME. If these studies do not have names the lead author or some type of explicit (ie first authoer et al) mention of the study should be done to make the readership aware of these smaller single center studies. 7. P.13 line 248: DRCR Protocol T specifically split the outcomes by initial visual acuity and found that at 2 years aflibercept had superior outcomes to bevacizumab. Ranibizumab has statistically similar outcomes to both aflibercept and bevacizumab at 2 years essentially splitting the difference. This is the key finding of the study and cannot be overlooked. Although mentioned in the discussion, it should also be mentioned clearly in the results. 8. P. 15 line 275: What study was this, where was it done, who is the author, was it single center etc. this should be mentioned more explicitly if there are only 2 studies being discussed in this section. SCORE2 is sufficiently well known but it should be mentioned how it is multi-center etc in the main text not just in a supplement or appendix that practically speaking very few people will read.

	9. P. 17 line 325: Is higher mortality in CATT and IVAN a conclusion from the metanalysis? If so please clarify. If not please show where this conclusion is in the primary trials. 10. P. 19 363: What exceptions specifically? I think this should be stated more explicitly. The only evidence I am aware of regarding visual outcomes where bevacizumab maybe inferior is in center involving DME in patients with 20/50 or worse vision. 11. It should be noted in limitations or elsewhere that the RCTs included have not looked at treat and extend treatment regimens, which is often considered the optimal standard to balance visual outcomes and treatment burden. This method is employed widely in the United States for example. 12. It should be noted that no RCTs are included or exist for ziv-aflibercept a low cost anti-VEGF alternative to aflibercept that is cheaper than bevacizumab. 13. Some discussion should be made regarding a multinational comparison of anti-VEGF medication use and the global public health implications of using bevacizumab as a viable alternative as many nations outside of the US use predominantly more expensive medications despite supposedly more cost conscious healthcare systems. 14. Although possibly outside the scope of the paper, the study would be more relevant if it also included the HAWK and HARRIER trials looking at brolucizumab (compared to aflibercept) which is poised to be the newest anti-VEGF on the market.

REVIEWER	Prof. Brian Faragher Liverpool School of Tropical Medicine U.K.
REVIEW RETURNED	10-Oct-2018

GENERAL COMMENTS	This is an excellent systematic review and meta-analysis. The paper is well written and very comprehensive. Background A well reasoned and persuasive rationale for this review is provided. Methods The methods used are described briefly in the main text of the paper, but in sufficient detail to allow readers to repeat the review if they should so wish; a fully comprehensive description of the methods are provided inn the Supplementary Material and Appendices. This reviewer is concerned that the search for relevant RCTs for inclusion in this review was confined to studies published in English, solely on the grounds of "time and resource constraints". The possible implications of this decision are discussed briefly in the Discussion as a study limitation. Given the wide range of their

	collective expertise in this clinical area, it is likely that the authors would have been confident that there are no non-English study reports available that would have been relevant to this review - is such a reassurance possible? The very comprehensive search for relevant studies within the grey literature is applauded. Inevitably, there was heterogeneity across the studies in terms of the primary outcome measure and the length of follow-up of patients. The methods used to deal with these issues are fully described and, for this reviewer, appear to be fully appropriate for the review objectives; however, sufficient information is provided to allow readers who may disagree with the approaches taken by the authors to repeat the meta-analyses using their own preferred analytical methods if desired. This reviewer agrees with the authors' decision to conduct their meta-analyses using a random effects model given the heterogeneity between studies. Results The review findings are reported in excellent detail. A comprehensive PRISMA statement is provided, along with a CONSORT-type diagram detailing the study flow. The statistical methods used to combine the results extracted from the studies identified in the search are fully appropriate for the study objectives and for the chosen primary outcome measure(s). Detailing the findings of such a comprehensive review in the text is always difficult, and nearly always results in a long list of summary statistics with their confidence intervals. These can be difficult to read and absorb. Excellent summary Tables are provided to help with this - but this reviewer is a great advocate of diagrams rather than lists of numbers so would urge that consideration be given to moving the Forrest plots currently in Appendix 7 to the main text as they summarise the study findings simply and comprehensively in an easily digestible form (it is appreciated that this may have to be an Editorial decision ultimately). Two small typos: the confidence interval in line 279 (page 16) needs a minus sign in front of the lower CI limit (-8.0); the confidence interval on line 212 (page 12) would be better reported as -3.3 to -0.5. Discussion A thoughtful discussion of the study findings is provided, fully supported by the results of the statistical analyses reported. I do not sense any issues in the context of this study, but as a statistician I would like to see a little more consideration of the clinical significance (or non-significance) of statistically significant (or non-significant) findings to confirm the extent to which these two concepts are in agreement.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Comment 1: This is a comprehensive and thorough piece of work. Given its strategic importance, this manuscript could be published as it is.

Response:

We thank you reviewer 1 for taking the time to review and provide feedback on our manuscript.

Reviewer 2

Comment 1: The authors should be commended on having a study with strong methods and a robust analysis/review of the literature. The major value of a study such as this would be to answer from pooled data how bevacizumab compares to aflibercept in AMD. The answer for DME and RVO is clearer as there have been head to head trials, but drawing a conclusion for AMD would greatly strengthen the paper. Essentially, everyone knows (anecdotally and from numerous head to head clinical trials) that bevacizumab and ranibizumab are clinically similar, but aflibercept is the point of debate as it has a slightly different mechanism as a VEGF trap (not a monoclonal antibody) and has been shown to decrease the treatment burden with longer interval dosing (see VIEW 1 and VIEW 2 trials). Aflibercept is also much more expensive than bevacizumab so the controversy with anti-VEGF medications is should I use an expensive medication and if so what is the marginal utility of that. We know the answer between bevacizumab and ranibizumab. Along with shedding more light on the aflibercept vs. bevacizumab issue in AMD, addressing the below points would also strengthen a paper that is likely going to be an important reference for the field. Please note the pages and lines refer to the authors' word processing document not the combined PDF from BMJopen.

Response:

Thank you for the comment. We have added the results regarding the comparative effectiveness of bevacizumab and aflibercept from the indirect comparisons of bevacizumab versus ranibizumab and aflibercept verus ranibizumab to our manuscript as follows:

Lines 221-230 –

"Comparative effectiveness of bevacizumab and aflibercept:

There were no RCTs that directly compared bevacizumab and aflibercept (Table 2, and Additional file 2: Appendix 6). Regarding BCVA change, the mean difference between bevacizumab and ranibizumab was -0.03 (95% CI: -1.08, 1.02) whereas the mean difference between aflibercept and ranibizumab was -0.05 (95% CI: -2.5, 2.4), suggesting a mean difference between bevacizumab and aflibercept of 0.02 (95% CI: -2.60, 2.64)50. For vision gain, the corresponding risk ratio estimate was 0.95 (95% CI: 0.84, 1.07) for bevacizumab versus ranibizumab, and 0.99 (95% CI: 0.81, 1.22) for aflibercept versus ranibizumab, suggesting a risk ratio estimate of 0.96 (95% CI: 0.75, 1.22) between bevacizumab and aflibercept."

In addition, we have placed the results of the "comparative effectiveness of aflibercept and ranibizumab" (lines 211-220) alongside those for the "comparative effectiveness of bevacizumab and ranibizumab" to facilitate the indirect comparison of bevacizumab and aflibercept.

Comment 2: p4 line 62 what "few exceptions" are there where bevacizumab is not a reasonable alternative to ranibizumab and aflibercept? Also what are the authors' conclusions regarding the difference between ranibizumab and aflibercept and becavizumab and aflibercept? These should be stated.

Response:

We have revised our conclusion to clarify this as follows:

Lines 65-72 - “Intravitreal bevacizumab was a reasonable alternative to ranibizumab and aflibercept in patients with cn-AMD, DME, RVO-ME and m-CNV. The only exception was for patients with DME and low visual acuity (<69 ETDRS letters), where treatment with aflibercept was associated with significantly higher vision gain (≥15 ETDRS letters) than bevacizumab or ranibizumab at 12 months; but the significant effects were not maintained at 24 months. The choice of anti-VEGF drugs may depend on the specific retinal condition, baseline visual acuity, and treatment regimen.”

Comment 3: p 9 line 169. The authors also excluded studies comparing anti-VEGF to photodynamic therapy (PDT) which is a mainstay adjuvant treatment for the polypoidal choroidal vasculopathy (PCV) subtype of nAMD which is very common globally in places such as in Asia, where the majority of AMD currently is. This should be mentioned.

Response:

We excluded RCTs evaluating pegaptanib and brolucizumab (not licensed in Canada at onset of review) and RCTs comparing the anti-VEGF drugs with other comparators, such as photodynamic therapy, intravitreal corticosteroids, grid laser photocoagulation, and cataract removal surgery, among others. This has been described in our detailed methods in Appendix 1, and we have revised the manuscript as follows:

Lines 135-137 - “We excluded RCTs comparing anti-VEGF drugs with other comparators, such as photodynamic therapy, intravitreal corticosteroids, and grid laser photocoagulation (Appendix 1).”

Comment 4: p 10 line 186: Are the 22% that gained vision specifically referring to those who gained 3 or more lines of vision which was often used as the measurement in many of the studies such as the CATT trial? This should be clarified and how many gained 3 or more lines of vision should be explicitly stated as that is often the benchmark in many of these trials.

Response:

Yes, we defined vision gain as a gain in Best-Corrected Visual Acuity (BCVA) letter score of ≥ 15 (3 or more lines) on the Early Treatment Diabetic Retinopathy Study (ETDRS) chart. We defined this in our methods section (see lines 139-141), and have now clarified this in the results on line 199.

Comment 5: line 192 p 11: It is widely known that bevacizumab and ranibizumab have similar treatment effects but what has your study shown us about bevacizumab and aflibercept as that is often the clinically and public health wise important issue as the question comes down to using low cost bevacizumab vs. high cost aflibercept while use of ranibizumab is relatively dropping per the literature.

Response:

We have added the results of the comparison of bevacizumab and aflibercept in the Results section as follows:

Lines 221-230 –

“Comparative effectiveness of bevacizumab and aflibercept:

There were no RCTs that directly compared bevacizumab and aflibercept (Table 2, and Additional file 2: Appendix 6). Regarding BCVA change, the mean difference between bevacizumab and ranibizumab was -0.03 (95% CI: -1.08, 1.02) whereas the mean difference between aflibercept and

ranibizumab was -0.05 (95% CI: -2.5, 2.4), suggesting a mean difference between bevacizumab and aflibercept of 0.02 (95% CI: -2.60, 2.64)50. For vision gain, the corresponding risk ratio estimate was 0.95 (95% CI: 0.84, 1.07) for bevacizumab versus ranibizumab, and 0.99 (95% CI: 0.81, 1.22) for aflibercept versus ranibizumab, suggesting a risk ratio estimate of 0.96 (95% CI: 0.75, 1.22) between bevacizumab and aflibercept."

In addition we have revised the Conclusion section accordingly (lines 62-69).

Comment 6: P 12 line 211: It is important to note that for mortality, PE, MI, etc none of the individual studies was powered to properly answer if one medication has a higher rate of any of the serious adverse events mentioned.

Response:

We have added this to our discussion section as follows:

Lines 383-385 - "None of the included RCTs were designed with sufficient statistical power to detect significant differences between the treatments with respect to the incidence of harms."

Comment 7: P 13 line 244: In the patients with DME section, the authors specifically mention information from DRCR Protocol T but the 2 other studies included (from Egypt and Turkey) are not mentioned. The data from each study and how that relates to the overall assessment of each medication for DME. If these studies do not have names the lead author or some type of explicit (ie first authoer et al) mention of the study should be done to make the readership aware of these smaller single center studies.

Response:

We agree and have updated this section to specifically mention and cite the two small single-centered RCTs as follows:

Lines 289-291 – "Besides the DRCR.net RCT, two small single-centered RCTs reported BCVA data, one comparing aflibercept with ranibizumab14, and another comparing bevacizumab and ranibizumab31."

Comment 8: P.13 line 248: DRCR Protocol T specifically split the outcomes by initial visual acuity and found that at 2 years aflibercept had superior outcomes to bevacizumab. Ranibizumab has statistically similar outcomes to both aflibercept and bevacizumab at 2 years essentially splitting the difference. This is the key finding of the study and cannot be overlooked. Although mentioned in the discussion, it should also be mentioned clearly in the results.

Response:

We reported the stratified results from the DRCR.net trial in the Results section as follows:

Lines 295-310 – "The DRCR.net trial reported results stratified by baseline visual acuity at 12 and 24 months (Appendix 11). In patients with high baseline visual acuity (BCVA ≥ 69 letters), approximately 16% of patients treated with bevacizumab, 15% of patients treated with ranibizumab and 18% of patients treated with aflibercept attained vision gain at 12 months (RR of bevacizumab versus aflibercept: 0.91 [95% CI: 0.50, 1.65]; RR of aflibercept versus ranibizumab: 1.18 [95% CI: 0.64, 2.17]). Vision gain at 24 months was 17% with bevacizumab, 19% with ranibizumab and 20% with aflibercept (RR of bevacizumab versus aflibercept: 0.84 [95% CI: 0.47, 1.52]; RR of aflibercept versus ranibizumab: 1.10 [95% CI: 0.63, 1.92]). In patients with low baseline visual acuity (BCVA < 69 letters), approximately 41% of patients treated with bevacizumab, 50% of patients treated with ranibizumab and 67% of patients treated with aflibercept attained vision gain at 12 months (RR of bevacizumab versus aflibercept: 0.62 [95% CI: 0.47, 0.81]; RR of aflibercept versus ranibizumab: 1.35

[95% CI: 1.06, 1.72]). At 24 months, vision gain was 52% with bevacizumab, 55% with ranibizumab and 58% with aflibercept (RR of bevacizumab versus aflibercept: 0.90 [95% CI: 0.69, 1.16]; RR of aflibercept versus ranibizumab: 1.05 [95% CI: 0.82, 1.35]).”

Comment 9: P. 15 line 275: What study was this, where was it done, who is the author, was it single center etc. this should be mentioned more explicitly if there are only 2 studies being discussed in this section. SCORE2 is sufficiently well known but it should be mentioned how it is multi-center etc in the main text not just in a supplement or appendix that practically speaking very few people will read.

Response:

We included details of these trials and the patient populations in the Results section as follows:

Lines 332-333 – “Results from one randomized, double-blind, controlled and non-inferiority trial conducted in India (including 77 patients with ME due to branch RVO)…”

Lines 338-339 – “Results from the SCORE2 randomized non-inferiority trial conducted in 66 centers in the United States (348 patients with ME due to central RVO)…”

Comment 10: P. 17 line 325: Is higher mortality in CATT and IVAN a conclusion from the metanalysis? If so please clarify. If not please show where this conclusion is in the primary trials.

Response:

We have clarified that the mortality results are based on meta-analysis of 2 trials, and included the references to the IVAN and CATT trials as follows:

Lines 252-255 – “Compared to the monthly regimen, the as-needed regimen was associated with a significant increase in mortality of 1.8% (95% CI, 0.1% to 3.4%, meta-analysis of mortality data reported in 2 RCTs, 1795 patients) [RR, 2.0; 95% CI, 1.2 to 3.5]35 51”

Comment 11: P. 19 363: What exceptions specifically? I think this should be stated more explicitly. The only evidence I am aware of regarding visual outcomes where bevacizumab maybe inferior is in center involving DME in patients with 20/50 or worse vision.

Response:

We have clarified the exception related to baseline visual acuity in the conclusions section as follows:

Lines 431-436 – “Intravitreal bevacizumab was a reasonable alternative to ranibizumab and aflibercept in patients with cn-AMD, DME, RVO-ME and m-CNV. The only exception was for patients with DME and low visual acuity (<69 ETDRS letters, 20/50 or worse), where treatment with aflibercept was associated with significantly higher vision gain (≥15 ETDRS letters) than bevacizumab or ranibizumab at 12 months; but the significant effects were not maintained at 24 months.”

We have also revised the abstract accordingly:

Lines 65-70 - “Intravitreal bevacizumab was a reasonable alternative to ranibizumab and aflibercept in patients with cn-AMD, DME, RVO-ME and m-CNV. The only exception was for patients with DME and low visual acuity (<69 ETDRS letters), where treatment with aflibercept was associated with significantly higher vision gain (≥15 ETDRS letters) than bevacizumab or ranibizumab at 12 months; but the significant effects were not maintained at 24 months."

Comment 12: It should be noted in limitations or elsewhere that the RCTs included have not looked at treat and extend treatment regimens, which is often considered the optimal standard to balance visual outcomes and treatment burden. This method is employed widely in the United States for example.

Response:

We included studies that used the treat and extend regimen (for example, Berg 2015), and this is highlighted in Table 3 and Appendix 10.

Comment 13: It should be noted that no RCTs are included or exist for ziv-aflibercept a low cost anti-VEGF alternative to aflibercept that is cheaper than bevacizumab.

Response:

We have noted this in our limitations as follows:

Lines 422-424 – "We did not include ziv-aflibercept (a low-cost anti-VEGF alternative to aflibercept and bevacizumab58), the old anti-VEGF pegaptanib, or the newest anti-VEGF brolucizumab."

Comment 14: Some discussion should be made regarding a multinational comparison of anti-VEGF medication use and the global public health implications of using bevacizumab as a viable alternative as many nations outside of the US use predominantly more expensive medications despite supposedly more cost conscious healthcare systems.

Response:

We have refrained from discussing this issue since it is beyond the scope of this review.

Comment 15: Although possibly outside the scope of the paper, the study would be more relevant if it also included the HAWK and HARRIER trials looking at brolucizumab (compared to aflibercept) which is poised to be the newest anti-VEGF on the market.

Response:

We have included this as a limitation in our manuscript as follows:

Lines 422-424 – "We did not include ziv-aflibercept (a low-cost anti-VEGF alternative to aflibercept and bevacizumab58), the old anti-VEGF pegaptanib, or the newest anti-VEGF brolucizumab."

Reviewer 3

Comment 1: This is an excellent systematic review and meta-analysis. The paper is well written and very comprehensive.

Response:

We thank reviewer 3 for taking the time to review our manuscript and provide thoughtful feedback.

Comment 2: [Background] A well reasoned and persuasive rationale for this review is provided.

Response:

Thank you.

Comment 3: [Methods] The methods used are described briefly in the main text of the paper, but in sufficient detail to allow readers to repeat the review if they should so wish; a fully comprehensive description of the methods are provided in the Supplementary Material and Appendices.

Response:

Thank you.

Comment 4: [Methods] This reviewer is concerned that the search for relevant RCTs for inclusion in this review was confined to studies published in English, solely on the grounds of "time and resource constraints". The possible implications of this decision are discussed briefly in the Discussion as a study limitation. Given the wide range of their collective expertise in this clinical area, it is likely that the authors would have been confident that there are no non-English study reports available that would have been relevant to this review - is such a reassurance possible?

Response:

We have searched the references of studies included in previous systematic reviews of similar review questions (references 11, 53, 56). These systematic reviews were without language restrictions.

However, as we did not search for non-English studies, we are unable to say with confidence that we did not miss any relevant studies in other languages. However, we are confident in our inference that the impact of including additional studies reported in other languages, if any, would be insignificant (lines 414-418).

Comment 5: [Methods] The very comprehensive search for relevant studies within the grey literature is applauded.

Response:

Thank you.

Comment 6: [Methods] Inevitably, there was heterogeneity across the studies in terms of the primary outcome measure and the length of follow-up of patients. The methods used to deal with these issues are fully described and, for this reviewer, appear to be fully appropriate for the review objectives; however, sufficient information is provided to allow readers who may disagree with the approaches taken by the authors to repeat the meta-analyses using their own preferred analytical methods if desired.

Response:

We agree with the reviewer’s assessment.

Comment 7: [Methods] This reviewer agrees with the authors' decision to conduct their meta-analyses using a random effects model given the heterogeneity between studies.

Response:

Thank you.

Comment 8: [Results] The review findings are reported in excellent detail. A comprehensive PRISMA statement is provided, along with a CONSORT-type diagram detailing the study flow.

Response:

Thank you.

Comment 9: [Results] The statistical methods used to combine the results extracted from the studies identified in the search are fully appropriate for the study objectives and for the chosen primary outcome measure(s).

Response:

Thank you.

Comment 10: [Results] Detailing the findings of such a comprehensive review in the text is always difficult, and nearly always results in a long list of summary statistics with their confidence intervals. These can be difficult to read and absorb. Excellent summary Tables are provided to help with this - but this reviewer is a great advocate of diagrams rather than lists of numbers so would urge that consideration be given to moving the Forrest plots currently in Appendix 7 to the main text as they summarise the study findings simply and comprehensively in an easily digestible form (it is appreciated that this may have to be an Editorial decision ultimately).

Response:

This review evaluated the benefits and harms of three anti-VEGF drugs and 4 retinal conditions. We therefore elected to summarize the findings in the main text and summary tables. While we appreciate the reviewer's suggestion, we have included the forest plots in the Appendix given the journal's limit on tables and figures.

Comment 11: [Results] Two small typos: the confidence interval in line 279 (page 16) needs a minus sign in front of the lower CI limit (-8.0); the confidence interval on line 212 (page 12) would be better reported as -3.3 to -0.5.

Response:

We thank the reviewer for catching these typos, and have corrected them in the text.

Comment 12: [Discussion] A thoughtful discussion of the study findings is provided, fully supported by the results of the statistical analyses reported. I do not sense any issues in the context of this study, but as a statistician I would like to see a little more consideration of the clinical significance (or non-significance) of statistically significant (or non-significant) findings to confirm the extent to which these two concepts are in agreement.

Response:

We recognize the challenges in reporting statistically significant findings and interpreting the corresponding clinical importance. We therefore presented treatment effect estimates in absolute terms (e.g., proportion differences) as well as relative terms (e.g., relative risk). We believe the former would help the readers to interpret the results, taking into account their prior notion of how large is importantly large.