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)	Reporting Quality of Randomised Controlled Trial Abstracts on Age-Related Macular Degeneration Health Care – A Cross Sectional Quantification of the Adherence to CONSORT Abstract Reporting Recommendations
AUTHORS	Baulig, Christine; Krummenauer, Frank; Geis, Berit; Tulka, Sabrina; Knippschild, Stephanie

VERSION 1 – REVIEW

REVIEWER	Jamie Kirkham University of Liverpool, United Kingdom
REVIEW RETURNED	30-Jan-2018

GENERAL COMMENTS	This article addresses the issue of reporting standards of abstracts in AMD studies in accordance to CONSORT for Abstract reporting standards. While there are many such studies that have looked at reporting guideline adherence, this is possibly the first to look at the abstract extension to this particular clinical condition, although the results provide no particular surprise compared to all other studies. Here are my specific comments: 1) Pre-Post period: The authors specify the post period to be 5 years from the publication of the reporting standard such that this matches the length of the pre-period. This is a limitation - the uptake of a reporting guideline is not always immediate and therefore the results to the post-period maybe underestimated. There was no obvious reason why the length of the post-period needed to match the pre-period. 2) The hypothesis and parameters to justify the sample size calculation was unclear. The hypothesis was not tested in the main results. The study was largely descriptive and it made more sense to evaluate all relevant article identified in the search. 3) Pg 6 line 57 'calcululation' 4) It was unclear in this application why stepwise procedures were used in the regression. You identified risk factors associated with reporting standards and reported all the effects sizes of these irrespective of significance - was backward variable selection really used? 5) It was unclear from the methods how pre-post period will be analysed

	6) Pg 8 - Line 23 How can Q3 = 10 when Q2 = 46? 7) Word count was the only factor that appeared to make a difference on improved reporting. Surely this is more of a consequence of authors being able to write more rather than actually following the checklist. Perhaps this needs discussion. 8) From 7) - the improvement in the post-period was only minimal, on average one more item. Seems hard to imagine that the guidance was followed at all when only half the items are reported in the post-period - this slight improvement may just be a coincidence? 9) I read many of these reporting guideline uptake studies and all seem to make the same recommendation - including details in the authors instructions section doesn't seem to work. Can the authors suggest perhaps a more novel solution - there have been some intervention studies looking at reporting guideline uptake - are there any specific to abstracts?

REVIEWER	Erik Cobo Bercelona-Tech, Spain
REVIEW RETURNED	13-Feb-2018

GENERAL COMMENTS	I read the bmjopen-2018-021912 paper and I think it is a very important paper that has to be prioritized for publication in BMJ Open. I offer some recommendations to authors. Major recommendations 1) You do not provide a link to a previous protocol; and so your report might be affected by the selective outcome reporting risk of bias. Please, highlight this as a major limitation. 2) In the assessment criteria of table 2, to improve readability, please employ exactly the same wording as on the Consort for abstracts checklist (i.e, on harms, the meaning has been lost). To avoid confusion among readers, please consider to further highlight the denominator on the two “effect size and precision” entries —or just to delete this sub-classification not included in the original checklist. 3) You are in general cautious to not interpret absence of evidence as evidence of absence. But not always, P17L55, “…did not affect…”. Please, be aware that non-significant results on your multivariate model (P16L40) can be the consequence of multicollinearity and over-adjustment. Please, discuss the concordance between adjusted and non-adjusted coefficients. If relevant, please highlight which variables are responsible for the loss of significance of other coefficients.[Please, find useful information about the interpretation of those results on the STROBE explanation and elaboration document.] 4) Please, provide arguments for the way “independent” numerical variables are included in the model (i.e., justify any cut-point to dichotomize variables). For example, could another cut-point, or treating the variable as numerical, have changed the results about the number of words? 5) Please clarify if the application of the exclusion criteria on the flow chart are sequential [I was expecting some papers to be excluded for more than one criteria.] Other suggestions: 6) Please, consider to avoid presenting your results both in the table and in the text. For example, reduce to the minimum necessary the repetition on the text of results already presented on the tables

	(all text results after table 2) or on the flux-diagram (P8, last sentence). Please, consider to discussion altogether the concordance results and to concentrate their values on a table, including uncertainty measures, such as 95%CI. 7) Your results are very important, but I see your interpretation as too prudent. RGs aims to improve transparency and reproducibility, which is essential to science, since non-reproducible papers can be considered as no science at all. Can an abstract with just 7 out of 14 items be considered transparent and reproducible? If the corresponding full-texts report those missing items, then those abstracts alone are not enough to offer a full picture of what was done and observed; and the journals (and authors) are responsible for this poor (not just "sub-optimal") reporting. From my point of view, this is a scandal, contributing to the big waste of research resources (including volunteer efforts) —a waste that has been estimated to be around 85% of research articles (DOI: 10.1016/S0140-6736(09)60329-9). If the abstract does not report the effect size with its precision, can the reader had an idea of the intervention effect? If not, what can be more important to be reported in a 250-words abstract? 8) Please, note an inconsistency at the Q3 in "median 46, Q1 28; Q3 10" (P9L23). 9) Please, avoid undesired repetitions in methods (i.e., "two independent reviewers" is repeated in P6L49 and 4 lines later).

VERSION 1 – AUTHOR RESPONSE

Point to Point Reply to Editorial Requests and Reviewers' Comments

Manuscript ID bmjopen-2018-021912 entitled "Reporting Quality of Randomised Controlled Trial Abstracts on Age-Related Macular Degeneration Health Care – A Cross Sectional Quantification of the Adherence to CONSORT Abstract Reporting Recommendations" which you submitted to BMJ Open, has been reviewed. The comments of the reviewers are included at the bottom of this letter.

Editorial Requests:

- Can you please provide a rationale for the dates of coverage for the literature search (01/2004 to 12/2013)? Why was the time period not more recent than this?
Was justified in the manuscript (p 15)

- When you submit your revision please ensure that you submit any supplementary material as a supplementary file. We note that the full search strategy was submitted under the file type 'review history' rather than a supplementary file in the current submission.
Done.

Reviewer: 1
Reviewer Name: Jamie Kirkham

1) Pre-Post period: The authors specify the post period to be 5 years from the publication of the reporting standard such that this matches the length of the pre-period. This is a limitation - the uptake of a reporting guideline is not always immediate and therefore the results to the post-period maybe

underestimated. There was no obvious reason why the length of the post-period needed to match the pre-period.
This is correct. Appropriate changes were included in the manuscript and discussed. As the project already started in 2014 - according to the start of the funding period - the publication period was limited to 2013.

2) The hypothesis and parameters to justify the sample size calculation was unclear. The hypothesis was not tested in the main results. The study was largely descriptive and it made more sense to evaluate all relevant article identified in the search.
This remark is absolutely correct. Actually, we carried out a full census of all available RCT abstracts (and provided the sample size argument rather to demonstrate the accuracy as achievable by the actual sample size 136). Corresponding changes have been made to the section "Material and Method": the sample size and accuracy argument has been omitted, the binary surrogate endpoint introduced for the corresponding formal hypothesis and sample size declaration has also been omitted.

3) Pg 6 line 57 'calcululation'
Paragraph has been deleted (see topic 2 above).

4) It was unclear in this application why stepwise procedures were used in the regression. You identified risk factors associated with reporting standards and reported all the effects sizes of these irrespective of significance - was backward variable selection really used?
Thanks for your important suggestion. In terms of a sensitivity analysis we recalculated the full model using continuous variables whenever possible (only exception “publication year before vs after 2008”). The maximum fitted model turned out to present the same significant variables as did the model previously displayed in our publication using the binary variables. According to your recommendation and a corresponding recommendation of reviewer 2 we substituted Table 3 by only displaying locally significant explanatory variables; again model fitting was based on the AIC criteria and backward-selection. Interaction terms were introduced into model selection, but were not found relevant and were therefore obmitted in Table 3.
The abstract information on the Poisson regression results has been updated accordingly.

5) It was unclear from the methods how pre-post period will be analysed.
The pre-post period was analysed as part of the multivariate analysis, but was not the primary question here.

6) Pg 8 - Line 23 How can Q3 = 10 when Q2 = 46?
Sorry – a “0” was missing: Q3 = 100. It has been edited.

7) Word count was the only factor that appeared to make a difference on improved reporting. Surely this is more of a consequence of authors being able to write more rather than actually following the checklist. Perhaps this needs discussion.
From our point of view it should be possible to include the desired information in a total of 250 words according to the checklist. But one has to know how. Please refer to pages 14 and 17.

8) From 7) - the improvement in the post-period was only minimal, on average one more item. Seems hard to imagine that the guidance was followed at all when only half the items are reported in the post-period - this slight improvement may just be a coincidence?
Yes, that's to be expected. Especially since the difference was so small that one could not speak of a relevant improvement.

9) I read many of these reporting guideline uptake studies and all seem to make the same recommendation - including details in the authors instructions section doesn't seem to work. Can the authors suggest perhaps a more novel solution - there have been some intervention studies looking at reporting guideline uptake - are there any specific to abstracts?
We tried to do so with explicit recommendations on p. 16 (by proposing that journals should provide annotated best practice abstract examples held in the clinical terminology and perspective of the respective journal) - thereby less focusing on an additional guideline explanation, but rather on something more practical and author-related, as we have the impression that reading and following the CONSORT recommendations for abstracts is a difficulty or at least exhaustive for clinical investigators.

Reviewer: 2
Reviewer Name: Erik Cobo

I read the bmjopen-2018-021912 paper and I think it is a very important paper that has to be prioritized for publication in BMJ Open.

Thanks a lot for your encouraging estimation of our work !!!

1) You do not provide a link to a previous protocol; and so your report might be affected by the selective outcome reporting risk of bias. Please, highlight this as a major limitation.
This is true, done – p 16.

2) In the assessment criteria of table 2, to improve readability, please employ exactly the same wording as on the Consort for abstracts checklist (i.e, on harms, the meaning has been lost). To avoid confusion among readers, please consider to further highlight the denominator on the two “effect size and precision” entries —or just to delete this sub-classification not included in the original checklist.
Done. Note, however, that Table 2 Is now part of supplementary material according to your recommendation of avoiding redundant presentations in text and table format. As we found, that the text should explicitly report important outcome results, we decided in favor of the text format and propose Table 2 (results on CONSORT criteria in detail) for the supplementary material; the previous Table 2 is now referred to as ”Table S”.

3) You are in general cautious to not interpret absence of evidence as evidence of absence. But not always, P17L55, “…did not affect…”.
Thanks! Done.

Please, be aware that non-significant results on your multivariate model (P16L40) can be the consequence of multicollinearity and over-adjustment. Please, discuss the concordance between adjusted and non-adjusted coefficients. If relevant, please highlight which variables are responsible for the loss of significance of other coefficients.[Please, find useful information about the interpretation of those results on the STROBE explanation and elaboration document.]
We fully agree with your suggestion and recalculated our Poisson regression models by introducing continuous variables whenever possible as well as appropriate interaction terms. This recalculation (also serving as a sensitivity analysis for a model building) resulted in the same significance pattern as the model previously presented. We now substituted Table 3 by presentation of the best fitted model after correction for interactions.
The abstract information on the Poisson regression results has been updated accordingly.

4) Please, provide arguments for the way “independent” numerical variables are included in the model (i.e., justify any cut-point to dichotomize variables). For example, could another cut-point, or treating the variable as numerical, have changed the results about the number of words?
See 3 (above).

5) Please clarify if the application of the exclusion criteria on the flow chart are sequential [I was expecting some papers to be excluded for more than one criteria.]
Done. Flowchart was edited and the legend added.

6) Please, consider to avoid presenting your results both in the table and in the text. For example, reduce to the minimum necessary the repetition on the text of results already presented on the tables (all text results after table 2) or on the flux-diagram (P8, last sentence).
Done. We have included Table 2 in the supplements and adjusted the text to the flowchart.

Please, consider to discussion altogether the concordance results and to concentrate their values on a table, including uncertainty measures, such as 95%CI.
For this purpose a new table (now denoted Table 2) has been inserted, which contains all Kappa estimates with one-sided 95% confidence intervals; the text has been re-edited, accordingly. The item-wise kappa statements have been deleted in the results section, a global statement has been introduced summarizing the table’s contents.

7) Your results are very important, but I see your interpretation as too prudent. RGs aims to improve transparency and reproducibility, which is essential to science, since non-reproducible papers can be considered as no science at all. Can an abstract with just 7 out of 14 items be considered transparent and reproducible? If the corresponding full-texts report those missing items, then those abstracts alone are not enough to offer a full picture of what was done and observed; and the journals (and authors) are responsible for this poor (not just “sub-optimal”) reporting. From my point of view, this is a scandal, contributing to the big waste of research resources (including volunteer efforts) —a waste that has been estimated to be around 85% of research articles (DOI: 10.1016/S0140-6736(09)60329-9). If the abstract does not report the effect size with its precision, can the reader had an idea of the intervention effect? If not, what can be more important to be reported in a 250-words abstract?
This definitely meets our opinion. We have tried to clarify the final statement (p 17) in this respect.

8) Please, note an inconsistency at the Q3 in “median 46, Q1 28; Q3 10” (P9L23).
Done (it was a typo with one “0” missing in Q3 = 100).

9) Please, avoid undesired repetitions in methods (i.e., “two independent reviewers” is repeated in P6L49 and 4 lines later).
Done.

VERSION 2 – REVIEW

REVIEWER	Jamie Kirkham University of Liverpool, UK
REVIEW RETURNED	13-Mar-2018

GENERAL COMMENTS	I would like to thank the authors for addressing my initial comments. I have no specific additional comments to add. The discussion section I felt was overly long but I will leave this decision to the editors.

REVIEWER	Erik Cobo

	Barcelona-Tech, Spain
REVIEW RETURNED	13-Mar-2018

GENERAL COMMENTS	The authors have addressed or clearly discussed my previous concerns and recommendations.