BMJ Open

The development of ARCADIA: a tool for assessing the quality of peer review reports in biomedical research

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-035604
Article Type:	Original research
Date Submitted by the Author:	08-Nov-2019
Complete List of Authors:	Superchi, Cecilia; Barcelona-Tech, Statistics and Operations Research Department; Université de Paris 1 Hren, Darko; University of Split, School of Humanities and Social Sciences Blanco, David; Barcelona Tech, Statistics and Operations Research Department; Université de Paris 1 Rius, Roser; Barcelona-Tech, Statistics and Operations Research Department Recchioni, Alessandro; Springer Nature Boutron, Isabelle; Université de Paris 1; Hôpital Hôtel-Dieu de Paris, Centre d'épidémiologie Clinique González, José Antonio; Barcelona-Tech, Statistics and Operations Research Department
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, PUBLIC HEALTH, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The development of ARCADIA: a tool for assessing the quality of peer review reports in biomedical research

Authors: Cecilia Superchi[1,2], Darko Hren[3], David Blanco[1,2], Roser Rius[1], Alessandro Recchioni[4], Isabelle Boutron[2,5†], José Antonio González[1†]

[†]Equal contributors

Affiliations:

[1] Statistics and Operations Research Department, Barcelona-Tech, UPC, c/ Jordi Girona 1, C5- 213, 08034, Barcelona, Spain

[2] Université de Paris, CRESS, INSERM, INRA, F-75004 Paris, France

[3] Chair of Psychology, Faculty of Humanities and Social Sciences, University of Split, Split, Croatia

[4] Springer Nature, Heidelberger Platz 3, 14197 Berlin, Germany

[5] Centre d'épidémiologie Clinique, Hôpital Hôtel-Dieu, Assistance Publique Hôpitaux de Paris (APHP), F-75004 Paris, France

Corresponding author:

Cecilia Superchi, Statistics and Operations Research Department, Barcelona-Tech, UPC, c/ Jordi Girona 1, C5- 213, 08034, Barcelona, Spain. cecilia.superchi@upc.edu, +34 631654317

Abstract

Objective: To develop a tool to assess the quality of peer review reports in biomedical research.

Methods: We conducted an online survey intended for biomedical editors and authors. The survey aimed to 1) determine if participants endorse the proposed definition of peer review report quality; 2) identify the most important items to include in the final version of the tool; and 3) identify any missing items. Participants rated on a 5-point scale whether an item should be included in the tool and they were also invited to comment on the importance and wording of each item. Principal component analysis (PCA) was performed to examine items redundancy and a general inductive approach was used for qualitative data analysis.

Results: A total of 446 biomedical editors and authors participated in the survey. Participants were mainly male (65.9%), middle-aged (mean=50.3, SD=13) and with PhD degrees (56.4%). The majority of participants (84%) agreed on the definition of peer review report quality we proposed. The 20 initial items included in the survey questionnaire were generally highly rated with a mean score ranging from 3.38 (SD=1.13) to 4.60 (SD=0.69) (scale 1 to 5). Participants suggested 13 items that were not included in the initial list of items. A steering committee composed of five members with different expertise discussed the selection of items to include in the final version of the tool. The final checklist includes 14 items encompassed in five domains (Importance of the study, Robustness of the study methods, Interpretation and discussion of the study results, Reporting and transparency of the manuscript, Characteristics of peer reviewer´s comments).

Conclusion: ARCADIA tool could be used regularly by editors to evaluate the reviewers' work, and also as an outcome when evaluating interventions to improve the peer review process.

Words count: (abstract: 290, word limit: 300), (main text: 3999, word limit: 4000), 4 tables, 1 figure, 1 web application, 7 supplementary files

Keywords: Peer review, Report, Checklist, Quality control, Survey

Strengths and limitations of this study

- ARCADIA constitutes the first evidence-based tool that has been systematically developed to assess the quality of peer review reports.
- Its development is based on an exhaustive review of the literature and on empirical data from a large and heterogeneous sample of both biomedical editors and authors.
- ARCADIA has not yet been validated.

Background

Editorial peer review stands as the gateway to scientific publication. The process was established to ensure that research papers are vetted by independent experts before they are published, although it is recently being increasingly questioned due to beliefs that it is flawed [1,2] . Despite efforts over the last 30 years to "make peer review scientific", its impact is still considered suboptimal [3].

Peer reviewers, who are the pivotal actors in this process, are requested to write a review report evaluating the submitted manuscript. A peer review report helps authors improve the quality of their manuscripts, and it also helps editors make an informed decision about the outcome of the manuscript. However, evidence shows that these peer review reports are often of poor quality [4,5].

Tools for assessing the quality of peer review reports have been proposed, of which we have conducted a systematic review and identified 24 tools: 23 scales and 1 checklist [6]. However, none reported any definition of peer review report quality, only one described the scale development, and 10 provided measures of reliability and validity. Further, the development and validation process resulted from a small consensus of people, and the concepts evaluated by these tools were quite heterogeneous.

In 2016, Bruce et al. published a review evaluating the impact of interventions to improve the quality of the peer review process [5]. The authors showed that it is essential to clarify the outcomes (such as, for example, the quality of peer review reports), which should be used in randomized controlled trials to evaluate these interventions.

A validated tool is direly needed to clearly define the quality of a peer review report in biomedical research. This tool could be used regularly by editors to evaluate the reviewers' work, and also as an outcome when evaluating interventions to improve the peer review process. In the present study, we report on the development of a new tool to assess peer review reports in biomedical research.

Methods

The study was approved by the Research Committee of the Governing Council of the Universitat Politècnica de Catalunya, Barcelona Tech, Spain (Reference: EC 02, Date: 02/05/2018).

Steering committee

We formed a steering committee of five members (CS, DH, AR, IB and JAG), whose expertise include clinical epidemiology, biostatistics, social science and editorial peer review. The steering committee agreed on how to define peer review report quality; they agreed on the survey questionnaire based on the results of a previous systematic review [6]; they interpreted the results of the survey; and they agreed on the final version of the tool.

Defining the tool's objective

The tool aims to assess the quality of peer review reports in biomedical research. We defined the quality of a peer review report as "the extent to which a peer review report helps editors make a fair decision and authors improve the quality of the submitted manuscript".

Generating the items

A systematic review allowed identifying 24 tools aimed at assessing the quality of peer review reports [6]. We extracted 132 items from these tools. All redundant items were merged and we included all items that met our definition of peer review report quality. Overall, 20 items were identified for assessing peer review report quality (Table 1).

Survey

We conducted an online survey of editors and authors in order to: 1) determine if they endorse the proposed definition of peer review report quality; 2) identify the most important items to include in the final tool; and 3) identify any new items that should be included.

Survey questionnaire

The questionnaire was constructed using the online survey software SurveyMonkey [7]. It was structured into four main parts and included both open and multiple-choice questions. First, the participants were asked to agree ("yes/no/partially") on the definition we provided for peer review report quality. They were also invited to add any comments or ideas on how to improve the definition. Second, they were asked to rate the importance of the 20 items for assessing the quality of peer review reports we identified. Their responses were based on a 1–5 Likert scale (1 being not important and 5

very important). In particular, we asked the participants if the item should be included in a tool for assessing the quality of peer review reports. Moreover, they were invited to comment on the importance and wording of each item. In order to eliminate the question order effect, the items appeared in random order for each respondent. Third, the participants were invited to suggest any additional items missing that they considered important for assessing the quality of peer review reports. Finally, the questionnaire included nine demographic questions related to sex, age, education level, job title, referring institution and job experience as biomedical editor and/or author. We developed two versions of the questionnaire because biomedical editors and authors were recruited differently, despite the fact that some of them could play both roles (see Supplementary file 1). The two versions were structured in the same way; they only differed in some questions related to the demographic characteristics. The questionnaire was piloted among six experienced scientific editors and authors, followed by a subsequent revision based on their feedback.

Participants and recruitment strategy

We targeted biomedical editors and authors using a purposive sampling approach to recruit a heterogeneous sample of information-rich cases [8].

Biomedical editors

By means of standardized email, we invited two groups of editors to participate in the survey: 586 biomedical editors from 43 journals in the BMJ Publishing group; and 478 editors from 235 journals identified in a previous cross-sectional bibliometric study [9] (see Supplementary file 2). The survey was also distributed to 27 editors from 48 journals in BMC (part of Springer Nature), using internal email, and to members of the European Association of Science Editors (EASE) through their newsletter. In the invitation email and newsletter, the editors were encouraged to forward the survey to colleagues who might be interested in issues related to peer review. This recruitment strategy, known as snowballing, allowed us to identify "information-rich key informants" [8]. On the first page of the survey, participants were informed that the collected data would be anonymous, and they were further asked if they would agree to share their de-identified data in an open access repository. Two reminder emails were sent to non-respondents. Finally, the survey was promoted on Twitter and on the EASE blog [10] and Methods in Research on Research (MiRoR) [11] websites.

Authors

Searching the top 30-biomedical journals with the highest impact factors, we identified 4396 corresponding authors of articles that reported original research and which were published in Medline

between February 1 and October 31 2018 (see Supplementary file 3). We used the R package easyPubMed to extract the email contacts [12]. The corresponding authors received a standardized email that explained the purpose of the study and included a link to the survey (see Supplementary file 2). The first page of the survey informed participants that the data were collected anonymously and also asked if they would agree to share their de-identified data in an open access repository. Two reminder emails were sent to non-respondents.

Data analysis

We described the demographic data in terms of frequencies and percentages. The importance of the 20 items to assess peer review report quality is described in means and proportions of editors or authors who rated the importance of the items from 1 to 5. The items were also sorted according to the mean raking of all participants and either editors or authors. We also calculated Pearson correlations among items. The calculations and graphical representations were all obtained using the statistical software R 3.5.0 [13].

Principal component analysis of quantitative data

We conducted a principal component analysis (PCA) to examine item redundancy among the 20 items to assess peer review report included in the survey. PCA is a multivariate statistical technique used to reduce the number of variables in a dataset to a smaller number of dimensions [14]. The new dimensions (or *principal components*) are mutually independent and are determined by choosing the directions that explain the most variation in the data. The first principal component (PC1) accounts for the largest possible variance in the data, and each succeeding PC accounts for decreasing amounts of the remaining. This exploratory analysis helps reveal simple underlying structures in complex datasets. We performed PCA using the R package FactoMineR [15].

Inductive content analysis of qualitative data

We used a general inductive approach for qualitative data analysis. In particular, we followed the five steps of inductive analysis proposed by David R. Thomas: 1) Preparation of raw data files; 2) Close reading of text; 3) Creation of codes; 4) Overlapping coding and uncoded text; 5) Continuing revision and refinement of themes system [16]. In the third phase, two investigators (CS and DB) created independently the initial codes from the responses of the first 100 participants for each open-ended question. In order to ensure consistency and credibility, the initial codes were discussed with a third investigator (DH) and a codebook was developed and was used for analysing the remaining responses. In case new codes were successively created from the remaining responses, the emerging

codes were added to the codebook and applied to entire dataset. Two investigators (CS and DH) reviewed and refined the codebook and further clustered the codes into major themes. We used the software NVivo V.12 for data management and analysis [17].

Selecting items

The steering committee reviewed all items and, ultimately, drafted and refined the final version of the tool. Based on the participants' qualitative and quantitative answers, redundant items were combined, existing items were modified and/or expanded on, and new items proposed by survey participants were added.

Patient or public involvement

Patients and members of the public were not involved in the study.

Results

Participants

Between November 7 2018 and February 4 2019, 198 biomedical editors and 248 authors completed the survey. Participants were mainly male (263/399, 65.9%) with a PhD degree (225/399, 56.4%), and their ages were equally distributed across ranges (mean=50.3, SD=13). They were mainly located in Europe (219/389, 56.3%) and North America (118/389, 30.3%). More than half of the editors had work experience of more than 5 years (91/165, 55.2%), while over one-third of the authors had work experience of more than 20 years (84/224, 37.5%) (see Table 2 and Supplementary file 4).

Definition of peer review report quality

Overall 84% (362/431) participants, precisely 85% (160/188) editors and 83% (202/243) authors, agreed on the definition of peer review report quality that we provided in the survey. The definition was slightly modified to take into account participants comments (Supplementary file 5). The quality of a peer review report is now defined as "the extent to which a peer review report helps, first, editors make an informed and unbiased decision about the manuscripts' outcome and, second, authors improve the quality of the submitted manuscript".

Quantitative results

We created a web application that is publicly available at https://www-eio.upc.edu/redir/ReportQuality. Through the application, the readers can easily access and explore the quantitative results of the survey.

Rating the importance of items

The items were generally highly rated, with a mean score ranging from 3.38 (SD=1.13) to 4.60 (SD=0.69). All the items were scored 4 or 5 by >50% of the participants (see web application). The three items rated as the most important were: 1) *Knowledgeability*; 2) *Methodological quality*; and 3) *Fairness*. The three least important items were: 1) *Originality*, 2) *Presentation and organization*; and 3) *Adherence to RG*.

A peer review report aims to help authors improve their submitted manuscripts and assist editors in taking editorial decisions. Due to this dual objective, we compared editors' and authors' mean scores in order to investigate whether any difference is found in their perceptions regarding the importance of the 20 items that assess peer review report quality. We found little discrepancy in the mean scores between biomedical editors and authors, with only two items indicating any difference: 1) *Timeliness* and 2) *Detail/Thoroughness*. The *Timeliness* of the peer review report was considered more important to authors than to editors (respectively, in the 12th and 16th rank positions). Meanwhile, editors rated the *Detail/Thoroughness* of the reviewer's comments higher than did authors (respectively, in the 11th and 16th rank positions).

Correlations among items

Overall, we found relatively weak positive correlations among items. The largest positive correlations were found between *Relevance* and *Originality*, and between *Fairness* and *Objectivity* (r = 0.55 and 0.43, respectively).

Principal Component Analysis

The first principal component (PC1) accounted for 22.1% of data variability. The next two dimensions (PC2 and PC3) accounted for 38.5% of the cumulative variability and contributed gradually, that is, they increased at only small increments. PC1 was positively correlated to all items (or variables), and it showed correlations higher than 0.4 —which is the figure commonly used as a threshold reference for factor loadings — for 16 out of 20 items (see web application). These results illustrate that the data variance was not concentrated in a few components but distributed across all of them; hence, reducing the number of items is not recommended, since this would imply an important loss of data information.

Qualitative results

Comments on importance and/or wording of items

Out of 446 survey participants, 267 (59.9 %) made at least one comment on the importance and/or wording of the items. Based on the initial coding of the comments, we were able to identify eight general themes that they addressed: Peer reviewer; Wording; Importance; Dependency; Responsibility; Item; Structure and content; and Improvement. Table 3 reports the eight themes together with their definition and the most frequent codes (n>5), with example quotes. The entire codebook is found in Supplementary file 5.

New items

Participants suggested 13 items that were not included in the initial list of items. These items are listed in Supplementary file 6. The entire codebook is found in Supplementary file 5.

Steering committee meeting

The steering committee met on the 19/07/2019 to discuss the selection of items to include in the final version of the tool. Their decisions were based on the participants' quantitative and qualitative answers. The flow of the items is summarized in Figure 1.

The items *Relevance* and *Originality* were merged into a new item named *Contribution* (of the study). This decision was based on the high positive correlation found between the two items (0.55) and on the participants' opinions. Furthermore, participants suggested in their comments that the item *Relevance* was "*highly subjective*", because "*each reviewer's decision on relevance reflects what is relevant to them, which may not reflect relevance to the journal*". They also believed that the *Originality* of a study is not always an important aspect for comments in a peer review report, because some manuscripts "*are trying to duplicate findings from previous studies*". They therefore suggested reformulating the two items by asking the reviewer what the study "*adds to our knowledge*".

The steering committee decided to include the item *Interpretation of results* as a domain of the tool instead of a single item, changing the name into *Interpretation and discussion of the study results*. This decision resulted from the addition of two new items (*Conclusions* and *Limitations*), based on the suggestions of survey participants. The domain *Interpretation and discussion of the study results* now encompasses three items: 1) *Study conclusions*; 2) *Study limitations and* 3) *Applicability and generalizability.*

Overall, survey participants believed that the items *Strengths and weaknesses (general)* and *Strengths and weaknesses (methods)* were "*confusing to separate*". Additionally, the steering committee agreed that *Strengths and weaknesses (methods) and Methodological quality* were also redundant; thus, it was ultimately decided to merge the three items into a new item named *Study methods.*

The items *Objectivity* and *Fairness* were merged because of both the moderate correlation between them (0.43) and the participants' opinions. Participants suggested that the total objectivity of the reviewer's comments is not possible because "*all decisions contain some personal biases and subjectivity*" and they also believed that the term fairness was "*very subjective*" and difficult to define. Additionally, the steering committee agreed to also combine these two items into *Supported by evidence*. The committee finally decided to merge all three items into *Objectivity*, and this was defined as "comments provided in a peer review report should be as objective as possible and, if considered appropriate, include references to support the reviewer's statements".

The steering committee agreed to merge *Structure of reviewer's comments* and *Clarity*, because participants considered both important for making the peer review report easy "*to read for both editors and authors*". Moreover, participants suggested that the *Detail/Thoroughness* of a peer review report was mostly associated with the quality of a manuscript, because in certain occasions a study can be so poorly conducted that "*a reviewer can highlight one or two major methodological flaws*" without conducting a detailed review. They therefore believed that a detailed report is not "*always necessary*" and instead preferred a succinct report that "*cuts straight to the critical points*". Taking into account the participants' opinions, the steering committee finally decided to include a single item named *Clarity*, which is defined as "a peer review report should be clear, succinct and well organized in order to be understood correctly by editors and authors".

The items *Tone* and *Constructiveness* were merged into *Constructiveness*, which is defined as "a peer review report should contain constructive and polite comments that allow the authors to improve the quality of their work". This decision was based on the participants' opinions that "*the comments should be polite and constructive*".

The item *Adherence to RG* and the new item *Reproducibility* suggested by survey participants were merged into *Reporting* based on the steering committee decision. The item *Reporting* was defined as "the reviewer should comment if the reporting of the study is clear, complete and transparent enough

for facilitating its reproducibility by verifying the adherence of the manuscript to the corresponding reporting guideline."

The items *Timeliness* and *Knowledgeability* were not included in the final version of the tool. Survey participants suggested that *Timeliness* was not "*directly tied to review quality*" because "*some of the best reviews come in past the deadline*". Furthermore, the steering committee agreed that the item *Knowledgeability* was generally difficult to assess, because it implied that anyone using the tool would have enough competence to evaluate the reviewer's knowledge and expertise. Five new items suggested by survey participants (*Data availability and software*, *Study protocol*, *Study conclusions*, *Study limitations* and *Relevant literature*) were finally included in the tool.

The ARCADIA tool

The ARCADIA (**A**ssessment of **R**eview reports with a **C**hecklist **A**vailable to e**DI**tors and **A**uthors) tool was finally developed. The tool is a checklist that includes five domains and 14 items (Table 4). Brief explanations of the items included in the five domains are provided in Supplementary file 7.

Discussion

This study resulted in a checklist of items to assess the quality of peer review reports in biomedical research. The checklist constitutes the first evidence-based tool that has been systematically developed to assess the quality of peer review reports.

The checklist is simple, applicable to any biomedical field, and consists of five domains covering 14 items, each of which is phrased as a question. Each item should be ticked as yes, no or not applicable (NA). An item could be checked NA if it is not covered in the study (e.g., there are no data, software or other materials attached to the manuscript) and/or the peer reviewer is not qualified to comment on that specific aspect (e.g., statistical methods).The ARCADIA tool has several strengths. It is the first tool ever developed based on an exhaustive review of the literature [6] and on empirical data from a large sample of both biomedical editors and authors. Further, it is the only tool that clearly defines the quality of peer review reports, as its definition was based on the perspectives of 446 authors and editors.

To develop the tool, we recruited a large sample of biomedical editors and authors with varying experience and backgrounds. We found the percentage of female participants who took part in the survey to be quite low (129/399, 32.3%). This is in line with evidence showing that gender equity in

academic medicine careers remains far behind [18]. Moreover, we recruited corresponding authors (who are usually first authors) from the top 30 biomedical journals. Evidence also shows that women are underrepresented as first authors among biomedical journals with high impact factors [19].

The present study also has some limitations. The survey questionnaire included some open-ended questions, which allowed participants to voluntarily express their opinions. However, we were not able to inquire further to clarify and verify some information provided by the study's participants. Therefore, the interpretation of some information could be affected by the perception of the three investigators who conducted the qualitative analysis. Additionally, since participants could comment voluntarily on the importance and wording of each item, the number of comments among items differed greatly. Furthermore, the majority of editors (132/165, 80%) who took part in the survey were from Europe. This result may be due to the recruitment strategy we used to identify biomedical editors. Although we also utilized a snowballing strategy, we mainly contacted editors through European biomedical journals. Finally, the present study reports on the first version of the ARCADIA tool, which has not yet been validated.

Implications

The tool is a general checklist available to all biomedical editors and authors. It could be regularly used by editors to evaluate the reviewers' work, and it can also be used as an outcome when evaluating interventions in order to improve the peer review process.

Conclusions

ARCADIA is the first checklist that has been systematically developed to assess the quality of peer review reports. It is based on the perspectives of a large and heterogeneous sample of biomedical editors and authors. Our plans for future work are to validate the ARCADIA tool.

Abbreviations: ARCADIA: Assessment of Review reports with a Checklist Available to eDItors and Authors; EASE: European Association of Science Editors; EQUATOR: Enhancing the Quality and Transparency of Health Research; MiRoR: Methods in Research on Research; NA: Not Applicable; PC1: First principal component; PCA: Principal component analysis; RG: Reporting guidelines

Acknowledgements

We thank Patrick Bossuyt, David Moher, Elizabeth Moylan, Albert Selva O'Callaghan, Alessandro Recchioni and José Jiménez Villa for piloting the survey questionnaire. We thank all people who participated in the survey.

References

1. Stahel PF, Moore EE. Peer review for biomedical publications: we can improve the system. BMC Med. 2014 Dec;12(1):179.
2. Smith R. Peer review: a flawed process at the heart of science and journals. 2006;99:5.
3. Rennie, Drummond. Make peer review scientific. Nature. 2016;
4. Jefferson T, Rudin M, Brodney Folse S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. Cochrane Methodology Review Group, editor. Cochrane Database Syst Rev [Internet]. 2007 Apr 18 [cited 2019 Jun 20]; Available from: http://doi.wiley.com/10.1002/14651858.MR000016.pub3
5. Bruce R, Chauvin A, Trinquart L, Ravaud P, Boutron I. Impact of interventions to improve the quality of peer review of biomedical journals: a systematic review and meta-analysis. BMC Med. 2016;14(1):85.
6. Superchi C, González JA, Solà I, Cobo E, Hren D, Boutron I. Tools used to assess the quality of peer review reports: a methodological systematic review. BMC Med Res Methodol. 2019;19(1):48.
7. SurveyMonkey Inc. SurveyMonkey. In San Mateo, California, USA; Available from: www.surveymonkey.com
8. Putton, M. Purposeful sampling. In: Qualitative evaluation and research methods. Sage; p. 169–86.
9. Sharp MK, Tokalić R, Gómez G, Wager E, Altman DG, Hren D. A cross-sectional bibliometric study showed suboptimal journal endorsement rates of STROBE and its extensions. J Clin Epidemiol. 2019 Mar;107:42–50.
10. EASE Blog. Biomedical editors survey on peer review [Internet]. 2018. Available from: https://ese-bookshelf.blogspot.com/2018/11/biomedical-editors-survey-on-peer-review.html
11. MiRoR. Biomedical editors survey on peer review [Internet]. 2018. Available from: http://miror-ejd.eu/2018/11/12/biomedical-editors-survey-on-peer-review/
12. Fantini,Damiano. Search and Retrieve Scientific Publication Records from PubMed [Internet]. 2019. Available from: https://www.data-pulse.com/dev_site/easypubmed/
13. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; Available from: https://www.R-project.org/
14. Jolliffe, I.T. Principal Component Analysis. 2nd ed. Springer; 2002.
15. Husson F, Josse J, Le S, Mazet J. FactoMineR: A Package for Multivariate Analysis. J Stat Softw. 2008;25(1):1–18.
16. Thomas, David R. A general inductive approach for qualitative data analysis. AJE. 2013;27.
17. QSR International. NVivo [Internet]. Available from: https://www.qsrinternational.com/nvivo/home
18. Bates C, Gordon L, Travis E, Chatterjee A, Chaudron L, Fivush B, et al. Striving for Gender Equity in Academic Medicine Careers: A Call to Action. Acad Med. 2016 Aug;91(8):1050–2.
19. Filardo G, da Graca B, Sass DM, Pollock BD, Smith EB, Martinez MA-M. Trends and comparison of female first authorship in high impact medical journals: observational study (1994-2014). BMJ. 2016 Mar 2;i847.

Contributors: All authors provided intellectual contributions to the development of this manuscript. CS, DH, IB and JAG jointly contributed to the study conception, design and interpretation of data. CS conducted the survey. CS, RR and JAG conducted the quantitative analysis and created the web application. CS, DB and DH conducted the qualitative analysis. CS, DH, AR, IB and JAG formed the steering committee. CS led the writing of the manuscript. IB and JAG led the supervision of the manuscript preparation. All authors provided detailed comments on earlier drafts and approved the final manuscript.

Funding: This project was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement no 676207. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscripts.

Competing interests: The authors declared no competing interests.

Patient consent for publication: Not required.

Ethics approval: The study was approved by the Research Committee of the Governing Council of the Universitat Politècnica de Catalunya, Barcelona-Tech (Spain).

Provenance and peer review: Not commissioned; externally peer reviewed.

Data availability statement: The dataset supporting the conclusions of the research reported in this paper will be available in the Zenodo repository in the Methods in Research on Research (MiRoR) community (https://zenodo.org/communities/miror/?page=1&size=20)

Tables

Table 1. The 20 items to assess peer review report quality included in the survey

Labels	Items to assess PR report quality
Relevance	The reviewer comments on the relevance of the study
Originality	The reviewer comments on the originality of the study
Interpretation results	The reviewer comments on the interpretation of study results
Strengths and weaknesses (general)	The reviewer comments on the general strengths and weaknesses of the study
Strengths and weaknesses (methods)	The reviewer comments on the strengths and weaknesses of the study methods
Statistical methods	The reviewer comments on the appropriateness of the statistical methods
Methodological quality	The reviewer comments on the methodological quality (internal validity) of the study
Applicability and external validity	The reviewer comments on the applicability and external validity of the study results
Presentation and organization	The reviewer comments on the presentation and organization of the manuscript
Adherence to RG	The reviewer comments on the adherence of the manuscript to the reporting guidelines
Structure of reviewer's comms.	The reviewer´s comments are structured and organized
Clarity	The reviewer´s comments are clear and easy to read
Constructiveness	The reviewer´s comments are constructive
Detail/Thoroughness	The reviewer´s comments are detailed and thorough
Objectivity	The reviewer´s comments are objective
Fairness	The reviewer's comments are fair
Support by evidence	The reviewer's comments are evidence based
Knowledgeability	The reviewer knows and understands correctly the content of the manuscript
Tone	The reviewer uses a courteous tone
Timeliness	The reviewer completes the peer review report on time

Table 2. Survey participants' characteristics

Characteristics	Editors N=198	Authors N=248	Total N=446
Gender	**N=169**	**N=230**	**N=399**
Woman	46 (27.2%)	83 (36.1%)	129 (32.3%)
Man	121 (71.6%)	142 (61.7%)	263 (65.9%)
Other	2 (1.2%)	5 (2.2%)	7 (1.8%)
Age	**N=156**	**N=220**	**N=376**
<40	32 (20.5%)	71 (32.3%)	103 (27.4%)
41-50	29 (18.6%)	59 (26.8%)	88 (23.4%)
51-60	52 (33.3%)	37 (16.8%)	89 (23.7%)
>60	43 (27.6%)	53 (24.1%)	96 (25.5%)
Education	**N=169**	**N=230**	**N=399**
Bachelor Degree	4 (2.4%)	3 (1.3%)	7 (1.8%)
Master Degree	11 (6.5%)	20 (8.7%)	31 (7.8%)
PhD	107 (63.3%)	118 (51.3%)	225 (56.4%)
M.D. or equivalent	34 (20.1%)	76 (33.0%)	110 (27.6%)
Prefer not to answer	2 (1.2%)	1 (0.4%)	3 (0.8%)
Other	11 (6.5%)	12 (5.2%)	23 (5.8%)
Location journal/institution	**N=165**	**N=224**	**N=389**
Europe	132 (80.0%)	87 (38.8%)	219 (56.3%)
North America	23 (13.9%)	95 (42.4%)	118 (30.3%)
South America	2 (1.2%)	5 (2.2%)	7 (1.8%)
Africa	1 (0.6%)	1 (0.4%)	2 (0.5%)
Asia	3 (1.8%)	11 (4.9%)	14 (3.6%)
Australia	4 (2.4%)	25 (11.2%)	29 (7.5%)
Number of years of experience	**N=165**	**N=224**	**N=389**
<5 years	74 (44.8%)	36 (16.1%)	110 (28.3%)
6-10 years	46 (27.9%)	51 (22.8%)	97 (24.9%)
11-15 years	27 (16.4%)	34 (15.2%)	61 (15.7%)
16-20 years	7 (4.2%)	19 (8.5%)	26 (6.7%)
>20 years	11 (6.7%)	84 (37.5%)	95 (24.4%)

Table 3. Survey participants' comments on the importance and/or wording of the 20 items to assess peer review report quality

Themes	Definition	Codes	Examples
Dependencies	Theme including codes on how the importance of an item depends on different factors (e.g., type of study, paper quality, type of journal, etc.)	Dependency on the type of study (n=34)	*Depends on type of study. For systematic reviews of course fundamental. For other studies this will be more and more important for easier comparisons between studies and for quality improvement. It makes our work easier if the authors also compliance also improve*
		Dependency on the paper quality (n=20)	*This depends on the quality of the manuscript. Sometimes the quality is so low that a reviewer can highlight one or two major methodological flaws which are sufficient to reject.*
		Dependency on the type of journal (n=19)	*This depends on the journal's criteria*
		Dependency on the author's claim and impact of the study (n=7)	*this depends on the claims made*
Importance	Theme including codes on the importance (or not) of an item.	Importance of the item (n=43)	*This is absolutely key to the interpretation of the study. Unfortunately most reviewers, in my field, do not fully understand current (and correct) methods.*
		Importance of replication and conformation study (n=18)	*Not always important to be original study as some are trying to duplicate findings from previous studies.*
		Importance of perceptions, opinions and experience (n=14)	*But some comments will inevitably be opinion, regarding emphasis, values, writing style*
		Importance of a high quality review rather than on time review (n=13)	*Better to have a late high quality report than a moderate quality report on time.*
Improvements	Theme including codes on how an item is useful for both authors and editors in the peer review process.	Useful for authors and editors (n=21)	*It's important to make it easy for the editor and authors to understand the review, and for authors to respond.*
		Improving the manuscript (n=9)	*Important when it will help improve the quality of the*

			communication. Not necessary when it flows well.
		Avoiding exaggeration and misinterpretation (n=8)	*This is an area where the reviewer may have a valuable role in tempering an author's enthusiasm, hubris or bias.*
Item	Theme including codes on the characteristics of an item.	Related to other item (n=43)	*Yes, but it is confusing to separate this from the general strength and weaknesses. The question should be if the reviewer thinks that the message can (potentially) answer the research question.*
		Subjective item (n=22)	*Too subjective! What is relevant to one person of field could be totally not-relevant to another*
		Requirement (n=9)	*But it's an ethical requirement, and helps improve everyone's experience.*
Reviewer	Theme including codes on the expertise and characteristics of a peer reviewer.	Reviewer's expertise (n=148)	*Some reviewers know about methods and some about content. It would be ideal to always have both, but that is often not the case.*
		Impossibility to be totally objective (n=35)	*100% objectivity doesn't exist*
		Reviewer as an extra unpaid job (n=10)	*for the most part, reviews are done on a voluntary basis*
Responsibility	Theme including codes on the editor and/or author´s responsibility to assess an item.	Editor's responsibility (n=48)	*In my experience this is usually picked up by the Editors and Associate Editors rather than the reviewers.*
		Joint responsibility (n=24)	*I think this is the role of the editors as well as the reviewers.*
		Author's responsibility (n=6)	*Authors should already be doing this*
Structure and content	Theme including codes on the structure and content of a peer review report.	Straight to the critical points (n=14)	*Sometimes a succinct review is still helpful, if it cuts straight to the critical points. For example, if it is clear that a manuscript has major flaws, then a review*

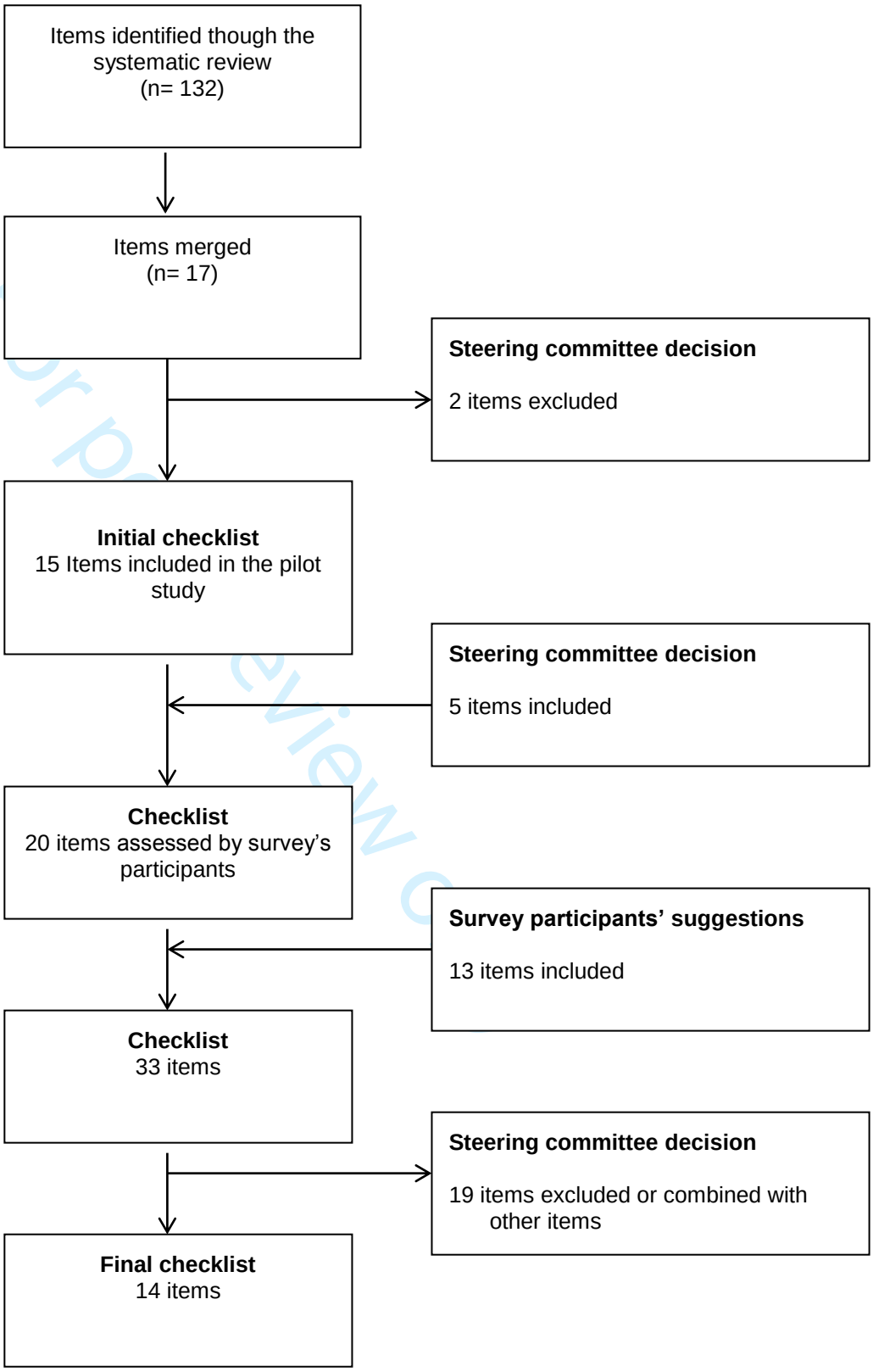
			that points out those flaws clearly and dispassionately would be very helpful. It would not necessarily need to delve into the finer details.
		Unnecessary to provide evidence to each comment (n=10)	*I don't think reviewers need to cite something for every point that they make.*
		Declaration of COI (n=8)	*Peer reviewers should disclose COI.*
		Standard structure of a review (n=7)	*I would suggest providing a template to reviewers.*
		Not necessary for all reviews (n=6)	*Reviews come in all lengths and vary in detail. It is helpful to have some reviewers provide detailed information but not necessary that all do so.*
Wording	Theme including codes on how to improve the wording of an item.	Wording of the item (n=110)	*Rather than "The reviewer's comments are evidence-based" I would suggest that the category should be: "The reviewer distinguishes between comments that are supported by evidence (and provides suitable citations) and those based on opinion or experience"*

Table 4. The ARCADIA tool

In the peer review report, did the reviewer comment on…		
Importance of the study	the contribution of the study to scientific knowledge?	☐ YES ☐ NO ☐ NA
	whether the relevant literature was accurately reviewed?	☐ YES ☐ NO ☐ NA
Robustness of the study methods	the soundness of the study methods (e.g., study design, outcomes, risk of bias)?	☐ YES ☐ NO ☐ NA
	the suitability of the statistical methods?	☐ YES ☐ NO ☐ NA
Interpretation and discussion of the study results	whether the study conclusions answer the research question(s) and correctly summarize the study results?	☐ YES ☐ NO ☐ NA
	whether the study limitations are acknowledged?	☐ YES ☐ NO ☐ NA
	the applicability and generalizability (external validity) of the study results?	☐ YES ☐ NO ☐ NA
Reporting and transparency of the manuscript	whether any major deviations from the study protocol are reported?	☐ YES ☐ NO ☐ NA
	whether the completeness of the reporting allows study reproducibility, by verifying the adherence of the manuscript to the corresponding RG?	☐ YES ☐ NO ☐ NA
	the presentation (e.g., quality of the written language, tables, figures, etc.) and organization of the manuscript?	☐ YES ☐ NO ☐ NA
	the availability of study data and material and whether the software works as indicated?	☐ YES ☐ NO ☐ NA
Were the peer reviewer´s comments…		
Characteristics of peer reviewer's comments	clear?	☐ YES ☐ NO
	constructive?	☐ YES ☐ NO
	objective and, if opportune, supported by evidence?	☐ YES ☐ NO

NA=Not applicable

Figure 1. Flowchart of items to include in a checklist to assess peer review report quality



Welcome to the survey!

Although the peer-review process plays a key role in research dissemination, only limited research has been conducted so far in this field.

The objective of this survey is to investigate the perspectives of biomedical editors and authors towards the **quality of peer-review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer review report in biomedical research.

Knowing your expertise, we would be very grateful if you could answer a few questions and share your opinion. The survey will take approximately **10 minutes** to complete. Your participation in this study is completely voluntary. If you decide to participate, all your answers will be de-identified and stored in a secured repository at Universitat Politècnica de Catalunya, Barcelona-Tech (Spain). The de-identified data from this study will be shared on Zenodo repository. In case you opt out of sharing your data, you will still be able to participate in the study.

This survey has received ethics approval from the Research Ethics Committee of the Universitat Politècnica de Catalunya, Barcelona-Tech (Spain).

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by Marie Skłodowska-Curie Action http://miror-ejd.eu/. The objective of MiRoR project is to train future generations of scientists in Research on Research, a new discipline aiming to promote research integrity increasing research value and reducing waste in health research.

This study is conducted by **Cecilia Superchi**, a PhD student at Universitat Politècnica de Catalunya, Barcelona-Tech and Université Paris Descartes, Sorbonne Paris Cité in collaboration with Prof. Darko Hren (University of Split), Prof. José Antonio Gonzalez (Universitat Politècnica de Catalunya) and Prof. Isabelle Boutron (Université Paris Descartes).

If you have any questions about this study or your rights as a participant, you may contact by email Cecilia Superchi, cecilia.superchi@upc.edu or Darko Hren, dhren@ffst.hr

Do you agree to take part in the study?

- ◯ Yes, I agree
- ◯ No, I do not agree

Do you agree to share your de-identified data?

- ◯ Yes, I agree
- ◯ No, I do not agree

Definition of peer-review report quality

The **quality of a peer-review report** could be defined as "to what extent the peer-review report helps editors to make a fair decision and authors to improve the quality of the submitted manuscript"

Do you agree with this definition?

- ◯ Yes
- ◯ No
- ◯ Partially

Please add your comments and ideas on how to improve the definition

Importance of the items to assess peer-review report quality

The following items have been identified in a systematic review as possible quality components of a peer-review report.

We are interested to know your opinion on the importance of these items, particularly whether the item should be included in a new tool assessing the quality of a peer-review report.

Please rate the IMPORTANCE of each item in assessing the quality of a peer-review report from 1 (not important) to 5 (very important).

We expect that for some items it will not be easy for you to make a clear decision about the importance of the item. In those cases we still invite you to offer your rating but you can elaborate on your decision. Furthermore we invite you to suggest potential improvements in wording of the items.

The reviewer comments on the **relevance of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **originality of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **interpretation of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **general strengths and weaknesses of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **strengths and weaknesses of study methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **appropriateness of the statistical methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **methodological quality (internal validity) of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **applicability and external validity of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **presentation and organization of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **adherence of the manuscript to the reporting guidelines**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **structured and organized**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **clear and easy to read**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer knows and understands correctly **the content of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **constructive**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **detailed and thorough**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer uses a **courteous tone**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer completes the peer review report **on time**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **evidence-based**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **fair**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **objective**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

New items to assess peer-review report quality

Are there any other items to assess the quality of a peer-review report that you think should be included?

Please list them.

Demographic characteristics

What is your gender?

- ◯ Female
- ◯ Male
- ◯ Prefer not to answer
- ◯ Other (please specify)

What is your age?

What is the highest level of education obtained?

- ◯ Bachelor Degree
- ◯ Master Degree
- ◯ PhD
- ◯ Other (please specify)
- ◯ MD or equivalent
- ◯ Prefer not to answer

Author's characteristics

What is your job title at your institution?

- ○ Researcher
- ○ Associate Professor
- ○ Assistant Professor
- ○ Professor
- ○ Other (please specify)

What type of institution are you affiliated at?

- ○ Private University
- ○ Public University
- ○ Research Centre
- ○ Other (please specify)

Where is the institution located?

- ○ Europe
- ○ Africa
- ○ North America
- ○ Asia
- ○ South America
- ○ Australia

How long have you been publishing scientific papers?

- ○ <5 years
- ○ 16-20 years
- ○ 6-10 years
- ○ >20 years
- ○ 11-15 years

Do you also work as biomedical editor?

- ○ Yes
- ○ No

Author's characteristics

Are you involved in making decisions on the manuscripts received by your journal?

- ◯ Yes
- ◯ No

Please check which of the following options you would be interested in

- [ ] I would be interested in receiving the results of the present study
- [ ] I would be interested in participating in the validation study of a new tool for assessing the quality of a peer-review report

Please write down your name and email address. Your data will be **EXCLUSIVELY** used for the option(s) which you have previously chosen.

Name	
Email Address	

Welcome to the survey!

Although the peer-review process plays a key role in research dissemination, only limited research has been conducted so far in this field.

The objective of this survey is to investigate the perspectives of biomedical editors and authors towards the **quality of peer-review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer review report in biomedical research.

Knowing your expertise, we would be very grateful if you could answer a few questions and share your opinion. The survey will take approximately **10 minutes** to complete. Your participation in this study is completely voluntary. If you decide to participate, all your answers will be de-identified and stored in a secured repository at Universitat Politècnica de Catalunya, Barcelona-Tech (Spain). The de-identified data from this study will be shared on Zenodo repository. In case you opt out of sharing your data, you will still be able to participate in the study.

This survey has received ethics approval from the Research Ethics Committee of the Universitat Politècnica de Catalunya, Barcelona-Tech (Spain).

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by Marie Skłodowska-Curie Action http://miror-ejd.eu/. The objective of MiRoR project is to train future generations of scientists in Research on Research, a new discipline aiming to promote research integrity increasing research value and reducing waste in health research.

This study is conducted by **Cecilia Superchi**, a PhD student at Universitat Politècnica de Catalunya, Barcelona-Tech and Université Paris Descartes, Sorbonne Paris Cité in collaboration with Prof. Darko Hren (University of Split), Prof. José Antonio Gonzalez (Universitat Politècnica de Catalunya) and Prof. Isabelle Boutron (Université Paris Descartes).

If you have any questions about this study or your rights as a participant, you may contact by email Cecilia Superchi, cecilia.superchi@upc.edu or Darko Hren, dhren@ffst.hr

Aren't you a biomedical editor? Please take part in the survey for **biomedical authors** following this link: https://www.surveymonkey.com/r/REPORT_QUALITY_AUTHORS

Do you agree to take part in the study?

- ◯ Yes, I agree
- ◯ No, I do not agree

Do you agree to share your de-identified data?

- ◯ Yes, I agree
- ◯ No, I do not agree

Definition of peer-review report quality

The **quality of a peer-review report** could be defined as "to what extent the peer-review report helps editors to make a fair decision and authors to improve the quality of the submitted manuscript"

Do you agree with this definition?

- ◯ Yes
- ◯ No
- ◯ Partially

Please add your comments and ideas on how to improve the definition

Importance of the items to assess peer-review report quality

The following items have been identified in a systematic review as possible quality components of a peer-review report.

We are interested to know your opinion on the importance of these items, particularly whether the item should be included in a new tool assessing the quality of a peer-review report.

Please rate the IMPORTANCE of each item in assessing the quality of a peer-review report from 1 (not important) to 5 (very important).

We expect that for some items it will not be easy for you to make a clear decision about the importance of the item. In those cases we still invite you to offer your rating but you can elaborate on your decision. Furthermore we invite you to suggest potential improvements in wording of the items.

The reviewer comments on the **relevance of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **originality of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **interpretation of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **general strengths and weaknesses of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **strengths and weaknesses of study methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **appropriateness of the statistical methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **methodological quality (internal validity) of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **applicability and external validity of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **presentation and organization of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **adherence of the manuscript to the reporting guidelines**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **structured and organized**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **clear and easy to read**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer knows and understands correctly **the content of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **constructive**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **detailed and thorough**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer uses a **courteous tone**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer completes the peer review report **on time**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **evidence-based**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **fair**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **objective**

Not important	Slighty important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

New items to assess peer-review report quality

Are there any other items to assess the quality of a peer-review report that you think should be included?

Please list them.

Demographic characteristics

What is your gender?

- ◯ Female
- ◯ Male
- ◯ Prefer not to answer
- ◯ Other (please specify)

What is your age?

What is the highest level of education obtained?

- ◯ Bachelor Degree
- ◯ Master Degree
- ◯ PhD
- ◯ Other (please specify)
- ◯ MD or equivalent
- ◯ Prefer not to answer

Editor's characteristics

What is your job title at your journal?

- ○ Editor in chief
- ○ Section editor
- ○ Associate editor
- ○ Deputy editor
- ○ Academic editor
- ○ Other (please specify)

Are you involved in making decisions on the manuscripts received by your journal?

- ○ Yes
- ○ No

At what type of journal do you currently working as editor?

- ○ General Journal
- ○ Specialty Journal

Where is the journal located?

- ○ Europe
- ○ Africa
- ○ North America
- ○ Asia
- ○ South America
- ○ Australia

How long have you been working as editor?

- ○ <5 years
- ○ 16-20 years
- ○ 6-10 years
- ○ >20 years
- ○ 11-15 years

Does your work inside or outside the journal include authoring scientific papers?

Yes

No

Please check which of the following options you would be interested in

- [ ] I would be interested in receiving the results of the present study
- [ ] I would be interested in participating in the validation study of a new tool for assessing the quality of a peer-review report

Please write down your name and email address. Your data will be **EXCLUSIVELY** used for the option(s) which you have previously chosen.

Name	
Email Address	

Supplementary file 2. Invitation email for corresponding authors and biomedical editors

From:
Cc:
To:
Subject: Academic Survey on Peer Review

Dear researcher,

As corresponding author of the article recently published in [CUSTOM 1], we would like to invite you to participate in an **academic survey**.

The objective of this survey is to investigate the perspectives of biomedical editors and authors on the **quality of peer-review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer-review report in biomedical research.

The survey will take approximately **10 minutes to complete**. Participation in this study is completely **voluntary** and you may withdraw at any time.

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No 676207 http://miror-ejd.eu/

We would be very grateful if you would take the time to complete our survey. **Your insights** as an author are **essential** to us.

If you have any questions, comments or queries please do not hesitate to contact us at cecilia.superchi@upc.edu or dhren@ffst.hr

We kindly thank you for your time, attention, and cooperation.

Sincerely,

Cecilia Superchi, PhD Student at Universitat Politècnica de Catalunya & Université Paris Descartes
Darko Hren, PhD, Prof. at University of Split
José Antonio Gonzalez, PhD, Prof. at Universitat Politècnica de Catalunya
Isabelle Boutron, MD, PhD, Prof. at Université Paris Descartes

From:
Cc:
To:
Subject: Academic Survey on Peer Review

Dear [Name] [Surname],

As [CUSTOM 1] at [CUSTOM 2], we would like to invite you to participate in an **academic survey on peer review.**

The objective of this survey is to investigate the perspectives of biomedical editors and authors on the **quality of peer-review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer-review report in biomedical research.

The survey will take approximately **10 minutes to complete**. Participation in this study is completely **voluntary** and you may withdraw at any time.

You are also encouraged to **forward the link** of the survey to your colleagues who may be interested in participating in this study https://www.surveymonkey.com/r/REPORT_QUALITY_EDITORS

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No 676207 http://miror-ejd.eu/

We would be very grateful if you would take the time to complete our survey. **Your insights** as a biomedical editor are **essential** to us.

If you have any questions, comments or queries, please do not hesitate to contact us at cecilia.superchi@upc.edu or dhren@ffst.hr

We thank you kindly for your time, attention, and cooperation.

Sincerely,

Cecilia Superchi, PhD Student at Universitat Politècnica de Catalunya & Université Paris Descartes
Darko Hren, PhD, Prof. at University of Split
José Antonio Gonzalez, PhD, Prof. at Universitat Politècnica de Catalunya
Isabelle Boutron, MD, PhD, Prof. at Université Paris Descartes

Supplementary file 3. Top 30-biomedical journals with the highest impact factors

Full Journal Title**	IF
New England Journal Of Medicine*	79.3
Lancet*	53.3
JAMA-Journal Of The American Medical Association	47.7
BMJ-British Medical Journal*	23.3
JAMA Internal Medicine	20.0
Annals Of Internal Medicine	19.4
Nature Reviews Disease Primers	16.1
Journal Of Cachexia Sarcopenia And Muscle	12.5
Plos Medicine	11.7
Bmc Medicine*	9.1
Mayo Clinic Proceedings*	7.2
Cochrane Database Of Systematic Reviews	6.8
Journal Of Internal Medicine	6.8
Canadian Medical Association Journal*	6.2
Journal Of Clinical Medicine*	5.6
American Journal Of Medicine*	5.1
Translational Research*	4.9
Annals Of Family Medicine*	4.5
Medical Journal Of Australia*	4.2
American Journal Of Preventive Medicine*	4.1
Amyloid-Journal Of Protein Folding Disorders	4.0
Journal Of General Internal Medicine*	4.0
Deutsches Arzteblatt International	3.9
Palliative Medicine	3.8
Preventive Medicine*	3.5
British Medical Bulletin	3.4
European Journal Of Internal Medicine*	3.3
British Journal Of General Practice*	3.3
Journal Of Pain And Symptom Management*	3.2
Qjm-An International Journal Of Medicine	3.2

* Journal reporting the corresponding author in the PubMed abstract.
**Source: InCites Journal Citation Reports 2017 under the category “Medicine, general and internal”.

Supplementary file 4. Complete participants characteristics

Characteristics	Editors N=165
Journal Role	
Editor-in-Chief	50 (30.3%)
Associate Editor	63 (38.2%)
Academic Editor	7 (4.2%)
Section Editor	6 (3.6%)
Deputy Editor	12 (7.3%)
Other (e.g. Statistical Editor, Patient Editor)	27 (16.4%)
Involvement in making decisions on the manuscript	
Yes	144 (87.3%)
No	21 (12.7%)
Type of Journal	
General Journal	39 (23.6%)
Specialty Journal	126 (76.4%)
Journal location	
Europe	132 (80.0%)
North America	23 (13.9%)
South America	2 (1.2%)
Africa	1 (0.6%)
Asia	3 (1.8%)
Australia	4 (2.4%)
Number of years of experience as editor	
<5 years	74 (44.8%)
6-10 years	46 (27.9%)
11-15 years	27 (16.4%)
16-20 years	7 (4.2%)
>20 years	11 (6.7%)
Authorship of scientific papers	
Yes	141 (85.5%)
No	24 (14.5%)

Characteristics	Authors N=224
Occupation	
Professor	63 (28.1%)
Associate Professor	31 (13.8%)
Assistant Professor	34 (15.2%)
Researcher	47 (21.0%)
Other (e.g. Lecturer, Postdoc, PhD)	49 (21.9%)
Type of Institution	
Public University	134 (59.8%)
Private University	33 (14.7%)
Research Centre	17 (7.6%)
Other (e.g. Hospital)	40 (17.9%)
Institution location	
Europe	87 (38.8%)
North America	95 (42.4%)
South America	5 (2.2%)
Africa	1 (0.4%)
Asia	11 (4.9%)
Australia	25 (11.2%)
Number of years of experience as author	
<5 years	36 (16.1%)
6-10 years	51 (22.8%)
11-15 years	34 (15.2%)
16-20 years	19 (8.5%)
>20 years	84 (37.5%)
Employment as biomedical editor	
Yes	63 (28.1%)
No	161 (71.9%)
Involvement in making decisions on the manuscript	
Yes	56 (88.9%)
No	7 (11.1%)

Supplementary file 5. Codebooks

The first codebook is about the suggestions made by survey participants on how to improve the definition of peer review report quality we provided.

The second codebook is about the comments made by survey participants on importance and/or wording of each item (n=20).

The third codebook is about the identification of new items to assess peer review report quality by survey participants.

Codebook 1. Suggestions on how to improve the definition of peer review report quality

Suggestions on how to improve the definition of peer review report quality (n= 87)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Assessment of different aspects of a study	Statements on the different aspects of a study that should be discussed in a peer review report	Accuracy of the study	NA	*Also helps ensure the accuracy of the content (at least in part)*
		Originality of the study	NA	*Should include recognition of what has already been done well where possible*
		Relevance of the study	NA	*Consider adding "the quality and value of the submitted manuscript." Or perhaps "quality and impact." The point is that some submitted work is high quality, but still not useful because it is just repeating prior work or answering an irrelevant question*
		Reproducibility of the study	NA	*The definition could broadly also take into account issues of reproducibility*
		Research integrity of the study	NA	*The definition should also include something about identifying plagiarism and conflict of interests*
		Robustness of the study	NA	*Not necessarily to improve the quality as that may be a secondary outcome. The peer review is to evaluate the scientific robustness of the research*

		Soundness of the study	NA	*Also - the extent to which the report critically assesses the soundness*
		Strengths and weaknesses of the study	NA	*I would include the extent to which the review accurately and clearly identifies weaknesses / limitations of the study ... but I don't disagree with your definition*
		Structure of the manuscript	NA	*A high quality peer review report may also identify the potential for a poorly structured paper to be revised into an impactful form*
		Study methodology and statistics of the study	NA	*Aspects such as rigorous statistical analysis and sampling/experimental design, degree of innovation, and the statement and testing of clear scientific hypothesis, should be addressed in a peer-review. Also, the technical issues and methodologies should be targeted during the peer-review process*
		Validity/trustworthiness of the study	NA	*What about the validity and trustworthiness of the findings?*
Consideration of journal´s policy	Statements on the consideration of the journal´s policy in writing a peer review report	NA	NA	*A good quality peer-review report takes into account journal policies and publication criteria while helping authors provide the best version of their work*
Irrelevant and ambiguous comments	Irrelevant and ambiguous comments for improving the provided definition of peer review report quality	NA	NA	*There is usually more than one report, so reportS*
Quality as a vague concept	Statements on the difficulty to define “quality”	Dependency on the type of journal and study	NA	*I would add at the end.... "based on a rubric specific to the type of article submitted"*
		Quality of research and quality of reporting	NA	*"Quality" is ambiguous. Relevant aspects of quality could include scientific validity (the extent to which the methods are adequate, the conclusions supported etc.) AND/OR reproducibility (the extent to which the study is described in sufficient detail that it could be reproduced). The former is a quality of the scientific study and the*

				latter is a quality of the text
Reviewer´s expertise	Statements on how the quality of a peer review report is related to the level of expertise of a reviewer	NA	NA	*The quality of the peer review also depends on how well the reviewer has understood the manuscript and the reviewer's level of expertise in the topic*
Reviewer's comments characteristics	Statements on the different characteristics of a peer review report	Clarity of the comments	NA	*I think the definition should include whether the reviewers have expressed themselves clearly and unambiguously*
		Constructiveness of the comments	NA	*Perhaps this is implied in the proposed definition, but you could mention that a high-quality peer review includes constructive criticism -- that is, not just an identification of flaws but suggestions for remedies*
		Fairness and impartiality of the comments	NA	*Add: 'is unbiased and competently-conducted'*
		Understanding correctly the content of the manuscript	NA	*The quality of the peer review also depends on how well the reviewer has understood the manuscript*
Role of external parties	Statements on the role of external parties in assessing the quality of a peer review report	NA	NA	*External parties should also play a role*
Scope of the peer review process	Statements on the scope of the peer review process	Different facets	NA	*Peer review has many facets*
		Ensuring accessibility to the readers	NA	*[..] and the accessibility to readers*
		Ensuring quality of science	NA	*I would like to insist on the role of peer-review to ensure the quality of the science presented in the manuscript*

		Evaluating rather than improving	NA	*The peer review is to evaluate the scientific robustness of the research*
		Independence	NA	*This definition does not capture the necessary independence of the peer review process*
		Related to decision making	NA	*This definition implies that editors' decisions can always override peer reviewers' appraisals. But an editor's appraisal of a paper should have equal weight to a peer reviewers appraisal*
		Transparency and critical appraisal	NA	*The focus of peer-review is transparency and critical appraisal. Peer review scope is broader than editorial decisions. Editorial decisions are a specific use case of peer review*
		Validation of the research	NA	*Peer review also helps to validate the research before publication, so the report also needs to do this to be effective*
		Variable process		*Peer review can be very variable; at its best it really improves the quality of papers. At its worst it is bullying and partial*
Timeliness of peer review process	Statements on the consideration of timeliness in defining the quality of a peer review report	NA	NA	*Timeliness should be included, less than 2 weeks is ideal*
Usefulness of the peer review report	Statements on the usefulness of a peer review report for authors and editors	Useful for authors	Improving manuscript quality	*It should be aimed at helping the authors improve the quality of their work.*
			More effective communication of research	*Peer review ideally contributes to effective communication through research publication, by exposing the author's work to the potential audience(s) for it and thus showing where readers stumble or identify limitations that need to be recognized. I think it would improve the definition if you could work in something about effective communication, which can be distinct from quality per se. More effective, clearer communication promotes more learning from the article and a stronger link to implementation. Also, quality should*

				probably be judged in terms of the purpose of the study; a delivery system study for example must provide more information on context in order to be useful than does a classical randomized trial such as of a pharmaceutical.
		Useful for editors	Filtering studies	*Also, 1) a check on poor research,*
			Decision to enhance the readership and citations	*helps editors to make a fair and informed decision that will enhance the readership and citations of the journal*
			Enabling fair decision	*I agree that they are primarily good to the extent that they help editors. "Fairness' is important, but enabling 'informed' and 'rigorous' decisions matter too.*
			Leading to incorrect decision	*Implies that the reviewer is making sensible suggestions, which may not be the case (and which an editor may not pick up on). E.g., a non-statistical reviewer commenting (incorrectly) on statistical methods and the editor is unaware if the comments are relevant/correct. The review could help the editor make a decision but it could be an incorrect decision.*
			Same weight for editors and reviewers appraisal	*This definition implies that editors' decisions can always override peer reviewers' appraisals. But an editor's appraisal of a paper should have equal weight to a peer reviewer's appraisal. This top down system can allow for bias.*
		Useful for both editors and authors	NA	*The point you suggest to help BOTH editor AND authors is a key element. When I ask for revision I provide new insights or suggestions to improve the quality and accuracy of a paper.*
Wording of the definition	Statements on how to improve the definition of the peer review report quality	Disagreement with the use of fair	NA	*This seems reasonable but I would leave out the word "fair" as I would assume that editors always aim to make fair decisions!*

Codebook 2. Comments on the importance and/or wording of each item

Relevance (n = 56)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author's responsibility	Statements on the author's responsibility to discuss the relevance of the study in the manuscript	NA	NA	*Relevance should be discussed by the authors and ultimately is decided by readers. I would expect reviewers to comment on the author's discussion on the relevance, and only exceptionally come with considerations of his own*
Contribution to the knowledge	Statements on the importance of the study as contribution to the scientific knowledge	NA	NA	*Relevance is important in the context of both the contribution to the knowledge base*
Dependency on the type of journal	Statements on how biomedical journals differently evaluate relevance of a study based on their own criteria	NA	NA	*Relevance also depends on the scope of the journal, and that is an editorial decision, opinion of the reviewer is not so important*
Editor's responsibility	Statements on the editor's responsibility to evaluate the relevance of the study	NA	NA	*This is largely an editorial decision*
Influencing editor's decision	Statements on how the relevance of a study can influence an editorial decision	NA	NA	*This is very helpful for the Editor to make a decision on the manuscript*
Readers as final judges	Statements on how readers are the final judges of the relevance of a study	NA	NA	*Relevance should be discussed by the authors and ultimately is decided by readers*

Related to other items	Statements on the link of the item with other items	NA	NA	*I took this to be the same question as the earlier one on applicability*
Reviewer's expertise	Statements on how the assessment of the relevance of a study depends on reviewer´s expertise	NA	NA	*Not important if you are a statistical reviewer, of a clinical article, without knowing the clinical area. BUT otherwise VERY IMPORTANT for clinical reviewers*
Subjective item	Statements on the subjective interpretation of the term "relevance"	External validity	NA	*Another aspect of relevance might relate to external validity or generalizability -- e.g., a lab study that does not have relevance to the real world*
		Future impact	NA	*Relevance may lie in the future, not in the present*
		Novelty	NA	*There are various aspects of "relevance" -- i.e., it might not be a significant contribution to the literature because the findings are not at all novel*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*Relevance to the broader field, or to general society?*

Originality (n= 56)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of journal	Statements on how biomedical journals differently evaluate the originality of a study based on their own criteria and policy	NA	NA	*Depends on the journal policy. More important when the reader is the client, less important when the author is the client*
Editor's responsibility	Statements on the editor's responsibility to evaluate the originality of the study	NA	NA	*Many journals mainly open which have different editorial policies now-a-days do not ask reviewers' to judge the originality. This is losing importance in open access era*
Importance of replication and confirmatory study	Statements on the importance of conducting replication and confirmatory studies	NA	NA	*Not always important to be original study as some are trying to duplicate findings from previous studies*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Slightly important item	NA	*This is only slightly important in that - once a study has been conducted*
		Important item	NA	*I find that important. The twentieth me too study is not relevant for the knowledge field*
Open access vs. subscription journal	Statements on how open access journals and subscription journals assess differently the originality of a study	NA	NA	*For our journal, this is very important, although I think for some others with a pay-for-publication model they place less importance on e.g. novelty*

Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Reviewer as not the best judge	NA	*Not all reviewers will necessarily be familiar with the literature in a particular area and may not be able to comment on originality of the study*
Wording of the item	Statements on how to improve the wording of the item and better define it	Impact rather than originality	NA	*Some less original studies can still be of significant value, so I prefer comments on impact to comments on originality*
		Originality as novelty	NA	*The Editor is usually aware of this, particularly Editors of journals with high impact factors, who are very keen to publish manuscripts reporting original/novel findings*
		Originality as what the study adds	NA	*I would prefer to think of this in terms of whether it really adds to our knowledge*

Interpretation of results (n= 33)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Avoiding exaggeration & misinterpretation and censoring divergent opinions	Statements on the importance of the item to avoid exaggeration and misinterpretation of study´s results	NA	NA	*This is an area where the reviewer may have a valuable role in tempering an author's enthusiasm, hubris or bias*
Conclusions supported by results, S&W and literature	Statements on the importance that study´s conclusions are supported by results, strengths and weaknesses and literature	NA	NA	*Interpretation of the findings should be judged by its coherence with findings and study limitations and strengths, and by its coherence with literature*
Contribution to the knowledge	Statements on the importance of the study as contribution to the scientific knowledge	NA	NA	*A judgement on the new contribution to knowledge*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Very important item	NA	*Interpretation of the results is crucial- it determines the message that is sent out. It is very important that reviewers pay attention to this interpretation*
Useful for readers	Statements on the uselessness of the item for the readers	NA	NA	*These are useless for the reader*
Liberty in the discussion section	Statements on the liberty of the authors to interpret study´s results in the discussion section	NA	NA	*As long as the results are crystal clear the authors can take some liberties in the discussion. As long as it is clear what is speculative*

Objective interpretation	Statements on the importance of the objectivity of the study´s interpretation	NA	NA	*As long as it is an objective interpretation without any confirmation bias*
Related to other items	Statements on the link of the item with other items	NA	NA	*See comments on methods above*
Related to study flaws	Statements on the importance of commenting on the item especially if there are major errors in the manuscript	NA		*This is important if there are any major flaws or if an alternative explanation for findings should be considered*
Related to study implications	Statements on the importance of commenting on the item for generating new hypothesis to test	NA	NA	*To me, this is the most important issue, for the point of papers is to generate new hypotheses to test. Unfortunately, in my field, editors often want to see the facts, but are wary about interpretations, probably about long-winded speculation in the past*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Statistics expertise	NA	*If the reviewer is experienced in statics can make a good interpretation of the results*
Rushed interpretation as common problem	Statements on the poor interpretation of the study´s results as common problem	NA	NA	*Discussion is the most important part of the manuscript. And sometimes it is a bit rushed by authors*
Scope of the PR process	Statements on the scope of the peer review process	Assisting editors to understand results	NA	*Editors aren't technical experts in every field. The PR process*

				is therefore important is assisting editors understanding the significance of results
Subjective item	Statements on the subjective interpretation of the term "interpretation of the results"	NA	NA	*All results are open to a variety of interpretations*
Wording of the item	Statements on how to improve the wording of the item and better define it	NA	NA	*The reviewer must comment on the discussion section, of which the interpretation is a part. But other elements (comparison with existing research etc.) is also important. I would replace 'interpretation' with 'discussion'*

Strengths and weaknesses (general) (n= 21)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author's responsibility	Statements on the author's responsibility to discuss the strengths and weaknesses of the study in the manuscript	NA	NA	*The paper should, the reviewer only should if the paper is missing something important*
Important when manuscript is overly long	Statements on the importance of the item when the manuscript is overly long	NA	NA	*Important when a manuscript is overly long*
Specificity of the comments	Statements on the importance of the specificity of the comments	NA	NA	*Specificity is more important*
Related to other items	Statements on the link of the item with other items	NA	NA	*This is repeated above for methods. So these two components overlap. However, I mark this as important*
Related to the study	Statements on the importance of commenting on the strengths and weaknesses of the study´s aims and study´s flaws	Related to study aims	NA	*I think this needs to be specifically related back to clear study aims and objectives (perhaps this is a separate category? If not I think it should be). Even a beautiful study design with great validity and statistics is rubbish if it doesn't allow you to answer your research questions! I always look to see if the analyses and interpretation address the goals of the study*

		Related to study flaws	NA	*This is important where there are issues*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Provided that the reviewer has the methodological skills to comment on methodological strengths and weaknesses.*
Taking into account reader's perspective	Statements on the importance of taking into account the reader´s perspective by peer reviewers	NA	NA	*Peer reviewers should take the reader perspective and ensure the manuscript is well balanced on these*
Including S&W in the general comments	Statements on including strengths and weaknesses in the general comments	NA	NA	*These should be clearly identified in the general comments*
Useful for editors	Statements on the importance of the item for editors for making an editorial choice	To make a decision	NA	*Important for deciding to accept or reject a manuscript*

Strengths and weaknesses (methods) (n= 29)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author's responsibility	Statements on author's responsibility to evaluate the item	NA	NA	*Authors should already be doing this*
Dependency on the methods quality	Statements on the importance of the item in relation to the quality of the methods of the study	NA	NA	*This really depends. If the methods are spurious, of course, this needs to be indicated*
Dependency on the type of journal and study	Statements on how the assessment of the item depends on the type of journal and study	NA	NA	*Methods are very important for our journal*
Focusing on the weaknesses	Statements on the importance of the item especially focusing on the weaknesses of the study	NA	NA	*Important, especially the weaknesses, where there is an obvious need*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Very important item	NA	*This is absolutely key to the interpretation of the study. Unfortunately most reviewers, in my field, do not fully understand current (and correct) methods*
Related to other items	Statements on the link of the item with other items	NA	NA	*Yes, but it is confusing to separate this from the general strength and weaknesses. The question should be if the reviewer thinks that the message can (potentially) answer the research question*

Reviewers' expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Providing the reviewer is suitably qualified to comment on the methods. In my experience, far too many reviewers comment on aspects of the Methods for which they are able to confidently comment on*
Subjective item	Statements on the subjective interpretation of the term "strengths and weakness (methods)"	To give indication on own stance	NA	*As long as the reviewer gives some indication of his/her own stance in determining strength and weakness*
Commenting on the transparency of methods used	Statements on the importance of commenting on the transparency of the methods used by the peer reviewers	NA	NA	*Authors should already be doing this. Is it necessary for reviewers to also state this, or is it redundant? Reviewers could comment on whether the authors have been transparent about the strengths and limitations*
Usefulness	Statements on the usefulness of the item for both editors and authors	Useful for authors	NA	*For the author if the strengths and weaknesses are not properly addressed in the paper*
		Useful for editors	NA	*This is important for the editor to make a decision*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*Appropriateness of methods (based on question) may be more important...unless there is a problem, then strengths and weaknesses becomes important*

Statistical methods (n=115)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*For some articles (e.g. RCTs, meta-analyses, and observational studies) assessment of the statistical methods is important. For other types of articles (reviews, commentaries, editorials) this is of less importance*
Editor's responsibility	Statements on the editor´s responsibility to determine the necessity of a statistical review for a study	Employment of statistical assessors by journals	Statistics sub-editor	*Journal needs to engage a statistics sub-editor for that*
		No familiarity with some methods	NA	*Sometimes editor may not be as familiar with certain statistical methods which makes it more difficult to do this*
		Statistical support to the reviewers by journals	NA	*Journal editors should provide statistical support to reviewers*
		To determine the necessity of statistical review by journals	NA	*But I think that a lot (?) of reviewers are not sufficiently capable to do so. It might be more appropriate that the editor determines the necessity of statistical review and explicitly asks the reviewer if he/she is capable to do so*

No consensus on appropriateness of statistical methods	Statements on no consensus on the appropriateness of methods by peer reviewers	Variety of appropriate methods	NA	*Often there are multiple "appropriate" methods. It is important to use one of the appropriate methods*
Optional component of quality	Statements on the item as option component of the quality	NA	NA	*This should be an optional or "where relevant" component of the quality*
Related to other item	Statements on the link of the item with other items	NA	NA	*This belongs to assessing the methods. Should not be a separate item because there is qualitative research*
Reviewers' expertise	Statements on how the assessment of the item depends on reviewer´s expertise	At least one reviewer	NA	*At least one of the reviewers should have reasonable statistical knowledge*
		Content expertise	NA	*Many reviewers may be subject matter experts but not necessarily experts in statistics*
		General methods reviewer	NA	*This applies to the methods in general, whether or not they are statistical. I think at least one reviewer needs to comment on methods, but not every reviewer. Articles may not use statistics but an expert on the methods should review the article*
		Inappropriate advice	NA	*Non-statisticians should not be encouraged to comment on the statistical methods*

		Statistical reviewer	NA	*Statistical reviewer should do this*
		Commenting on own statistical expertise	NA	*In my view, the reviewer should be required to state whether or not she/he has the expertise to evaluate the statistical methods properly*
Commenting on the use of statistical methods	Statements on the importance to comment of the appropriate use of the statistical methods by authors	NA	NA	*And the use of stat. methods (some methods are used incorrectly by authors)*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*We get a lot of qualitative work so the key to this question is the appropriateness of the methods and then specifics based on type of methods*

Methodological quality (n= 32)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*Where applicable it is important*
General comments	General statements	NA	NA	*But in any case, I think internal validity is very important -- if a study is claiming that there is a relation between two variables it should be on solid ground to do so*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Very important item	NA	*In my opinion this is the most important item*
Focusing more on methods than results	Statements on the importance of commenting on the methods	NA	NA	*In my opinion this is the most important item. I think a reviewer should primarily focus on methods rather than results*
Related to other items	Statements on the link of the item with other items	NA	NA	*I did not understand the difference between this question and the question "The reviewer comments on the strengths and weaknesses of study methods"*
Reporting of the study	Statements on the importance of good reporting for study reproducibility	NA	NA	*A description should be sufficient to repeat the study with a high likelihood to end up with the same results*

Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Acknowledgement lack expertise	NA	*Ideally yes. However good reviewers are also well aware of the limits of their own expertise. It is better for reviewers to acknowledge that they lack expertise in relation to a particular aspect of the methodology (this is OK - no one is expected to be an expert in everything) rather than attempt to critique something that is outside of their own scope of knowledge*
		At least one reviewer	NA	*At least one reviewer with expertise in methods should review the study*
		Technical vs. clinical expertise	NA	*When selecting reviewers one might choose some for technical expertise (methodology, statistics etc.) and others for clinical expertise/experience*
Making sure results are not biased	Statements on the importance of the study for making sure the study results are not biased	NA	NA	*Crucial to make sure the results are not biased*
Wording of the item	Statements on the wording of the item and how to improve it	Broad item	NA	*"Methodological quality" is a broad term that could apply to construct validity and statistical validity as well as internal validity*

Applicability and external validity (n= 37)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author´s responsibility	Statements on the responsibility of authors to comment on the applicability and external validity of the study providing sufficient information to the readers	NA	NA	*The paper should, not necessarily the reviewer*
Dependency on the practice of the reader	Statements on the difficulty to judge the importance of the item because it depends on the practice of the readers	NA	NA	*Difficult as would depend on the context of practice of the reader*
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*It depends on the specific topic of the study*
Editor's responsibility	Statements on the editor's responsibility to evaluate the item	NA	NA	*This is more 'scope, which is for the editor to decide. But help from a reviewer is appreciated*
Helping the editor to understand reproducibility of the study	Statements on the importance of the item to decide if a study can be reproduced	NA	NA	*This is important when reporting novel findings as it helps the Editor to decide if the results can be reproduced by another group*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Less important item	NA	*This is of lesser importance. There should be sufficient information included for any*

				readers to come to this conclusion themselves
Related to other items	Statements on the link of the item with other items	NA	NA	*Similar to my answer about interpretation*
Related to the paper	Statements on the claims made in the paper by the authors and impact of the study	Future research	NA	*Applicability might lie in the future, not in the present*
		Impact	NA	*This is important only in relation to the claims made in the paper about the impact and implications of a study*
		Related to the claim & content of the paper	NA	*This depends on the claims made*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Technical vs. clinical expertise	NA	*When selecting reviewers one might choose some for technical expertise (methodology, statistics etc.) and others for clinical expertise/experience*
Reviewers' comments characteristics	Statements on the different characteristics of a peer review report	Baring reviewers' opinion	NA	*This may be one area of the review where the reviewer can bring a personal opinion to bare. Does the reviewer think this is a useful paper?*
		Evidence based comments	NA	*Peer reviewers should provide citations (evidence) for their assessment. Simply saying that the results are not applicable to their practice is not enough*

		Tempering authors' enthusiasm	NA	*Similar to my answer about interpretation: this is an area where the reviewer may have a valuable role in tempering an author's enthusiasm, hubris or bias*
Subjective item	Statements on the subjective interpretation of the term "applicability and external validity"	NA	NA	*This can be very subjective and misleading*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*Applicability and external validity are two concepts, so this item is double-barrelled in not valid*

Presentation and organization (n= 45)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of journal (and its policy)	Statements on how biomedical journals differently evaluate the item based on their own criteria	Presence of copy editors in the journal	NA	*Depends a bit on whether journals have good copy editors*
		Taking into account the average reader of the journal	NA	*The reviewer needs to take into account the "average reader" of the journal - will they understand the paper?*
General comments	General statements	NA	NA	*Peer review is not an editorial exercise, but clarity and reproducibility are part of good science*
Useful for editors	Statements on the usefulness of the item for editors	NA	NA	*Because the readability is important to those who've not seen it before. Especially helpful when a handling editor is new, I think.*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Less important item	NA	*This is less important, because as long as the content is there, a reader should be able to make use of the paper, even if it requires more effort. But if the presentation and organisation is really bad, then it needs to be addressed*
		Presentation more important than organization	NA	*Presentation is important organization is not*

Improving the manuscript	Statements on the importance of the item to improve the quality of the manuscript	Clear recommendations	NA	*Yes, but in a way that provides the authors with clear recommendations on how to make improvements. Design flaws cannot always be addressed after the study, but issues with presentation and organization of the manuscript can*
		Communication	NA	*Important when it will help improve the quality of the communication. Not necessary when it flows well*
		Readability	NA	*Important because this impacts readability*
Not going into irrelevant comments	Statements on the importance of not making useless comments	NA	NA	*Important when it will help improve the quality of the communication. Not necessary when it flows well.*
Formatting minutiae	Statements on peer reviewers focusing on minutiae	NA	NA	*Some reviewers focus on formatting minutiae*
Related to reporting guidelines	Statements on the link of the item with reporting guidelines	NA	NA	*I find reviewer comments on the presentation and organization of the manuscript moderately important if the manuscript follows a check list (e.g. STROBE) and/or standard formatting, and if is easy to understand and follow*

Responsibility	Statements on editor, author or reviewer´s responsibility to evaluate the item	Joint responsibility	NA	*I think this is the role of the editors as well as the reviewers.*
		Editor's responsibility	NA	*Editors and editorial staff have a stronger role here.*
		Reviewer´s responsibility	NA	*I regularly make notes as to whether a section is better placed elsewhere in the document, and on sentence structure, and use and misuse of citations. I think this is an obligation that reviewers have to the author and the journal*
Subjective item	Statements on the subjective interpretation of the item	NA		*This is subjective and may vary between reviewers as log as general structure is preserved*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*The word "presentation" seems unclear. It could refer to the writing quality or to other factors*

Adherence to RG (n= 73)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Adherence to key points	Statements on the importance that a manuscript adherences on the key elements of a checklist	NA	NA	*I think whether a manuscript adheres to a specific item on a checklist is not that important. Adhering overall to the key elements needed to report is important*
Part of the PR process	Statements on the importance of checking the adherence of reporting guideline as part of the peer review process	NA	NA	*If it is widely accepted reporting guidelines like the Consort Guidelines I think that is an important part of peer review*
Author's responsibility	Statements on the author´s responsibility to follow reporting guidelines	Too demanding for authors	NA	*Some reviewers are too much strict on that*
Dependency on the type of journal	Statements on how the assessment of the item depends on the type of study	Consistent format	NA	*It would be great to have a consistent format and rubric to follow to increase comparability of manuscript and distress authors*
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*Depends on type of study. For systematic reviews of course fundamental. For other studies this will be more and more important for easier comparisons between studies and for quality improvement. It makes our work easier if the*

				authors also compliance also improve
Editor's responsibility	Statements on the editor's responsibility to evaluate the item	Joint responsibility	NA	*The editor can also take care of this aspect*
		Pre-review	NA	*I believe this is the editor's job pre-review*
		Reformatting articles	NA	*We accept manuscripts that have been formatted for other journals for peer review. Of course we move towards acceptance they need to be reformatted*
General comments	General statements	NA	NA	*Universal reporting guidelines, like CONSORT, can be expected that all reviewers would know*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item		*Essential*
Lack of awareness	Statements on the lack of complete awareness about reporting guidelines from respondents	NA	NA	*In my experience, reviewers know little about the reporting guidelines of the journal for which they are reviewing. I think reviewers should always be sent details of the key guidelines. Otherwise they make criticisms or suggest changes which are incompatible with the guideline of the journal*

Providing reporting guidelines	Statements on providing reporting guidelines to reviewers	NA	NA	*The reviewer should receive the reporting guidelines applicable to the manuscript under review*
Reviewer's responsibility	Statements on the reviewer´s responsibility to check if the manuscript adherences to reporting guideline	Less reviewer's responsibility	NA	*I see that this is less the reviewer's responsibility to be honest*
		Making easier for reviewers	NA	*Important for improving standards in reporting, but this should be made as easy for the reviewer as possible, because otherwise it can be too arduous*
		Discussion of the study's issues	NA	*Pointing out where the manuscript does not respect the guidelines is useful, but more important is discussing the issues themselves*
		Tedious for reviewers	NA	*When doing reviews, it is quite tedious to have to relate to difference reporting and formatting guidelines of particular journals*
Getting an accurate review	Statements on how reporting guidelines help delivering an accurate review	NA	NA	*That always irritating when authors do not follow the recommendations oto authors starting from pagination... which helps for delivering an accurate reviewing*

Unclear responsibility	Statements on the unclear responsibility of checking for the adherence of the manuscript to reporting guideline	NA	NA	*I am not sure whether this is the peer reviewers' or the editor's responsibility*
Utility of reporting guidelines	Statements on the uncertain utility of reporting guidelines	NA	NA	*I'm not convinced that reporting guidelines make that much difference, but they are certainly better than nothing*
Wording of the item	Statements on how to improve the wording of the item	Meaning of reporting guidelines	NA	*I simply do not know what this means. Which reporting guidelines?*

Structure of reviewer´s comments (n= 33)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Content and completeness are more important	Statements on the more importance of review´s content and completeness	NA	NA	*Completeness is more important than how structured*
Definition of structured and organized	Statements on how to define the item	NA	NA	*General comments (e.g. on style) followed by structured comments (line by line or section by section)*
Dependency on the structure of the manuscript	Statements on the importance of the item in relation to the structure of the manuscript	NA	NA	*Again - somewhat dependent on the structure of the manuscript that is being peer-reviewed*
General comments	General statements	NA	NA	*Peer review is not an editorial exercise, but clarity and reproducibility are part of good science.*
Useful for both authors and editors	Statements on the importance of the item in helping authors and editors	Making easier to answer	NA	*Makes it easier for the authors responding*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Not so important	NA	*It helps, but I'm not sure this is important enough to be assessed. Should covary strongly with other characteristics of the review*
Not related to meaningful content	Statements on the no relationship between a well-structured review and meaningful review's content	NA	NA	*Makes it easier to respond to but doesn't mean the review content is more or less meaningful*

Related to other item	Statements on the link of the item with other items	NA	NA	*Makes it easier to respond to comments if they are clear and easy to read*
Review reorganized by editors	Statements on the reorganization of a review by editors	NA	NA	*It is up to the editor to interpret the referee comments and make concrete recommendations or demands on the authors if needed.*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers	NA	NA	*Semantic point, be careful about asking too much from unpaid and unrewarded reviewers*
Standard structure of a review	Statements on the necessity to have a standard structure for reviews	Different perspective	NA	*Organised according to who's perspective...one person's structure is another's chaos?*
		More difficult for reviewers	NA	*But the more you set exacting standards for a review, the more difficult you make it for a reviewer. This is undoubtedly something to aim for, but reviewer time is an issue*
Time consuming to reorganize the review	Statements on the time lost in reorganizing the reviewer´s comments	NA	NA	*Otherwise time is lost in trying to reorganize and understand what the reviewer means*

Clarity (n= 26)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Editors can make the comments clearer	Statements on editor´s task to edit the reviewer´s comments	NA	NA	*Helpful but not essential since the editor can help make sense of them for authors*
General comments	General statements	NA	NA	*Peer review is not an editorial exercise, but clarity and reproducibility are part of good science*
Useful for authors and editors	Statements on the usefulness of a clear peer review report for both authors and editors	NA	NA	*Otherwise neither the editor nor the authors can use the review appropriately*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Clarity is important*
		Less important	NA	*As long as the authors can understand the meaning, it is more important that the paper is clear*
Not a marker of quality	Statements on not considering clarity as marker of quality	NA	NA	*To me, although this is essential, it is more of an expectation of the review, rather than a marker of quality*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers	NA	NA	*But also conscious that we're all writing reviews late at night and so sometimes the ideal 'slips'*

To avoid repeated cycles of PR	Statements on the importance of the item to avoid repeated cycles of PR	NA	NA	*Yes - to avoid repeated cycles of peer review*
Useful for authors	Statements on the usefulness of the item for authors	Authors can ask for further clarity	NA	*It should be acceptable for authors to query reviewers' comments and ask for further clarity*
		Easy to respond	NA	*Makes it easier to respond to comments if they are clear and easy to read*
		Making sure the comments are intended	NA	*It is necessary to improve the chances that the comments are taken as intended*
Wording of the item	Statements on how to improve the wording of the item	Disagreement on the wording easy to read	NA	*I think easy to understand may be a better way to say this. I'm not sure easy to read is as applicable in the age of the computer*

Constructiveness (n=46)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Clear guidance	Statements on the importance to give clear guidance on how to improve the manuscript	NA	NA	*Worth emphasising that they should, where appropriate, give clear guidance on how paper might be improved and not be derogatory*
Dependency on the paper quality	Statements on how the constructiveness of reviewer´s comments depends on paper quality	NA	NA	*I suppose there will be some submissions which are so poor, this will be difficult*
Extent of the comments	Statements on the consideration to what extent reviewer`s comments could be addressed	NA	NA	*I think that reviewers should also consider to what extent their comments can be addressed. For example, if it's a paper on a survey, it's not helpful for a reviewer to say that more people should be surveyed*
General comments	General statements	NA	NA	*The worst reviews are the ones where the reviewer just rambles on and does not provide something to respond to*
Importance of destructive comments	Statements on the importance of destructive comments	NA	NA	*Some bellicose reviews are pretty helpful*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*I would rank this as the most important*

Marker of quality	Statements on constructiveness as a marker of quality	NA	NA	*Constructiveness flags the reviewer's interest in improving the paper, so it is a marker of the likely value of their review*
Negativity of the comments	Statements on the total negativity of reviewer´s comments	NA	NA	*The reviewers' comments are important, however at times, without any reasons the comments are totally negative*
Not mandatory requirement	Statements on constructive comments as not a mandatory requirement	NA	NA	*They can be, but it's not mandatory. Some manuscripts shouldn't be published*
Not reviewers' responsibility	Statements on how reviewers should not rewrite the paper but be respectful	NA	NA	*It is not the reviewers' job to rewrite the paper or mentor the authors. However comments should always be respectful*
Related to author's experience	Statements on how constructive comments are related to the experience of authors	NA	NA	*It depends on the status of the author. A beginner in a field needs encouragement and support. An older expert who is talking rubbish deserves more direct language*
Related to recommendation	Statements on how constructive comments are also useful if the manuscript is rejected	NA	NA	*Important, even if the recommendation is to reject: the authors will probably submit elsewhere, the comments can be useful for them in order to improve the paper*
Related to the readership's interest	Statements on how constructive comments are related to the interest of readership	NA	NA	*Unless manuscript really not of interest to readership, then I would not expect a reviewer to*

				spend a lot of time essentially helping the authors
Related to other items	Statements on the link of the item with others	NA	NA	*Hands in hands with being courteous*
Sometimes difficult to be constructive	Statements on how comments are sometimes difficult to present in a constructive way	NA	NA	*I am not sure this reflects quality - valid concerns over methodology, results etc. are sometimes difficult to present in a constructive way. Clearly being constructive is preferable though*
Subjective term	Statements on the subjective interpretation of the term "constructiveness"	NA	NA	*I think this is a subjective term*
Usefulness for both authors and editors	Statements on how constructive comments are useful for both editors and authors	Useful for authors	NA	*Directly linked to helping the author improve the manuscript.*
		Useful for editors	NA	*In case the review aims to support the editor to offer a revision, constructiveness of the review is more relevant*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*I want respectful and helpful. Sometimes that is different than "constructive."*

Detail/Thoroughness (n= 62)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Accommodating reviewer's comments	Statements on how authors accommodate reviewer´s comments	NA	NA	*On the other hand, authors often spend a lot of time with accommodating reviewer comments that were maybe not that relevant to start with. So there is a limit to how detailed and thorough is still helpful and the authors should have the right to reject some of the requests*
Dependency on the paper quality	Statements on how detailed comments depends on the quality of the paper	Detailed when paper is inadequate	NA	*Sometimes, where a paper is clearly inadequate producing a detailed report is necessary*
Detailed but not useful review	Statements on how detailed comments are not always useful	NA	NA	*They can be detailed but not useful--for example, when they concentrate on grammar and spelling*
For improving or rejecting the manuscript	Statements on the importance of detailed review to improve or reject a manuscript	NA	NA	*Sometimes the length of the comments is greater than the length of the manuscript. Peer reviewers should provide positive suggestions how the paper can be improved or rejected. The forma of the comments should be the same as the format of the responses, e.g. what I propose and why*

Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*Most important!*
Inconsistency in length	Statements on how peer review report are inconsistent in length	NA	NA	*The most disconcerting thing about reviews is inconsistency - when one is five pages long, the other five lines*
Not always necessary	Statements on how detailed comments are not always necessary	NA	NA	*Ideally but not always necessarily*
Providing a justification	Statements on the importance of providing a justification in the comments	NA	NA	*This is a pet peeve of mine. Some reviewers say things like "it has been demonstrated that this method of analysis is flawed" without providing a reference, for instance*
Related to other items	Statements on the link of the item with others	NA	NA	*I would have thought clarity was a more important criteria then being detailed but agree about thoroughness*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers	NA	NA	*Reviewers' time is valuable*
Straight to the critical points	Statements on the importance of succinct comments	Detecting fatal flaws	NA	*Focusing on one major flaw is more important than recitating all the typos*
		Excessive details	NA	*But they can be too detailed leading to a report that is too long overwhelming the author with too many requested revisions*

		Expectation from the authors	NA	*the most important is that the reviewer clearly indicate what he/she is expecting from the authors*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*"detailed, thorough and clear" (or unambiguous)*

Objectivity (n= 62)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Citing own work	Statements on reviewers citing their own work in a peer review report	NA	NA	*Please can reviewers not cite their own work. This usually means they have approached the paper with bias*
Declaration of COI	Statements on the importance of reviewer's conflict of interest declaration	More important than be objective	NA	*Declarations of competing interest and bias are more important than the claim to be objective*
Dependency on the study type	Statements on how objective comments are related to the type of study	Related to study's quality	NA	*This is a fundamental principle, that the comments should be disinterested (i.e. not driven by the reviewer's self-interest) as this increases the chance of the comments relating to the paper's quality*
		Related to the novelty of the study	NA	*Moreover, reviewer's comments are strongly influenced by the reputation of the author and the novelty of the idea. The less known the author and the more novel the idea, the reviewer tends to be less objective*
Editor's objectivity rather than reviewer's objectivity	Statements on the importance of objectivity from editors rather than peer reviewers	NA	NA	*Editors are supposed to be objective, so reviewers can be opinionated if they wish*

Following a specific rubric	Statements on the importance to follow a specific rubric to guide comments by peer reviewers	NA	NA	*Reviewer should follow a specific rubric to guide comments and make revision manageable by author*
General comments	General statements	NA	NA	*This is one of the most critical elements of good peer review in my opinion but also one of the rarest things to find*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*This is a fundamental principle, that the comments should be disinterested (i.e. not driven by the reviewer's self-interest)*
Impossibility to be total objective	Statements on the impossibility for reviewers to be totally objective	Comments are subjective by definition	NA	*All reviews are subjective!*
		Desirable to express own opinion	Awareness of own experience	*A better expectation is that reviewers come to the role aware of their own backgrounds, culture, experiences, research and views on the topic will affect their assessment of the research*
		Reminding reviewers to be objective	NA	*As far as possible - reminding reviewers to be as objective as possible would be a good start*
		Subjective comments are helpful for both editors and authors	NA	*I think there are subjective comments that are still valuable to the authors and editors. For example, if the reviewer finds a section of the manuscript to be unclear, this is there subjective*

				opinion but can still help the authors re-assess and potentially improve that portion of the manuscript
Justification of the comments	Statements on the importance of substantiating the comments by peer reviewers	NA	NA	*It is important that the reviewer substantiates comments, and that the authors are able to respond in case of revising their manuscript, to the reviewer's comments*
Recruiting additional reviewers	Statements on recruitment of additional peer reviewers when the reviews are not objective	NA	NA	*When peer reviewers recommend citing own papers or clearly favour one treatment over the others, editors should recruit additional peer reviewers*
Related to authors' reputation	Statements on how peer reviewers are influenced by author's reputation	NA	NA	*Moreover, reviewer's comments are strongly influenced by the reputation of the author and the novelty of the idea. The less known the author and the more novel the idea, the reviewer tends to be less objective*
Related to other item	Statements on the link of the item with others	NA	NA	*Goes along with a courteous tone*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Every reviewer will always have their own perspective based on their expertise*
Wording of the item	Statements on how to improve the wording of the item	Unclear item	Difficult to define	*It would be very difficult to define this*

Fairness (n= 55)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Biases are unavoidable	Statements on how reviewer's biases are unavoidable	NA	NA	*Reviewer biases are a reality, but one should (where possible) recognize them and phrase criticism in that light*
Editor's responsibility	Statements on the editor's responsibility to be fair	NA	NA	*Fairness is the editor's responsibility to judge*
Fair depends on author's characteristics	Statements on how the reviewer´s comments are influenced by author's characteristics	NA	NA	*It is well known that reviewers comments are not fair in terms of the location, ethnicity and gender of the authors*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*Fairness is extremely important*
Importance to back up opinions	Statements on the importance of backing up opinions by peer reviewers	NA	NA	*Crucial that a reviewer backs up their opinion with evidence from the paper/published literature*
Justification based on the paper quality	Statements on how justification of reviewers is based on the quality of the paper	NA	NA	*As per above, the reviewers comments must be justified based on the quality of the article rather than on their personal views*
Recognizing COI	Statements on recognition of reviewer´s conflict of interest	NA	NA	*More specifically, Editors should identify if the referee*

				has potential conflict of interest especially if he/she can have a conflict of interest working on the same field or topics. We all know such
Related to other items	Statements on the link of the item with others	NA	NA	*Objective, evidence-based, fair etc. are highly correlated*
Reviewers' perspectives	Statements on the importance of having peer reviewers with different perspective	NA	NA	*We often seek reviewers with different perspectives, so the entire editorial review team is constructed to be fair*
Subjective item	Statements on the subjective interpretation of the term "fairness"	NA	NA	*Fair, of course, is subjective*
Wording of the item	Statements on how to improve the wording of the item	Simplistic way to assess quality	NA	*Donald Trump sees the world in terms of fair or unfair. I think this is too simple a view*
		Unclear item	Difficult to define	*How do you define 'fair'?*
			Difficult to measure	*How would you measure "fairness" of a review?*

Support by evidence (n= 69)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Context dependency	Statements on how the importance of item depends on the context	NA	NA	*Completely depends on the context; sometimes common sense can suffice but other times evidence-based critiques are necessary to show authors and editors why something needs changing*
Dependency on the type of study	Statements on how comments supports by evidence depends on the type of study	NA	NA	*Depends on the type of study*
Editor's responsibility	Statements on the editor's responsibility to determine if reviewer´s comments are relevant or not	NA	NA	*It is the editor's role to determine whether they are relevant or not*
Especially for supporting criticism	Statements on the importance of supporting criticism using evidence	NA	NA	*Comments, especially criticisms, should be supported by citations wherever possible; subjective criticism ("I prefer such and such a method ...") is not constructive*
Helpful when there are disagreements	Statements on the particular importance of the item especially when there are disagreements	NA	NA	*Particularly if authors objections are rooted in disagreements with their own opinions or collaborators opinions*
Importance of perceptions, opinions and experiences	Statements on the importance of perceptions, opinions and	Especially for too innovative manuscript	NA	*The exceptions are when the manuscript is too innovative or*

	experience of a reviewer in assessing a paper			*groundbreaking. In these case, knowledge and expertise to identify a possible major breakthrough is of utmost importance*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Very important. As an author and a frequent reviewer I have seen reviewer comments which are anything from completely wrong to simple statements of opinion without any evidence-base. But in my experience, what I would regard as quality reviews, up to date and accurate, are exceptionally rare*
		Not important item	NA	*Do you mean that they offer citations for their comments? If so, that's probably not important*
More information needed by authors	Statements on the necessity to get more information by authors	NA	NA	*Sometimes comments may be based on a hunch -- and more information from authors may be needed*
Unnecessary to provide evidence to each comments	Statements on how it is unnecessary to provide evidence for each comment	NA	NA	*I don't think reviewers need to cite something for every point that they make*
Related to other items	Statements on the link of the item with others	NA	NA	*If you mean, is it an objective review, then it is important*

Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Based on the reviewer's knowledge and experience*
Selecting use of evidence	Statements on providing the use of no selective evidence	NA	NA	*Provided the use of evidence is not selective*
Difficulty of backing up all comments	Statements on the difficulty of backing up each comments	NA	NA	*Wild claims may need back up but it is unrealistic to expect reviewers to back up everything they say. A quicker system for raising and addressing queries would be a better response*
Wording of the item	Statements on how to improve the wording of the item	Difficult to measure	NA	*But I'm not sure how you judge this*
		Providing citations	NA	*Do you mean that they offer citations for their comments?*
		Unclear item	NA	*I don't really understand*

Knowledgeability (n= 57)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
CoI between reviewers and authors	Statements on possible conflicts of interest between peer reviewers and authors	NA	NA	*Although this is very important it can create a conflict of interest as the authors and reviewers may be involved in the same field of research and this could result in a degree of bias for or against the research described in the manuscript*
General comments	General statements	NA	NA	*Reviewers should be able to commit time and effort to the process and be held accountable to the commitment.*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Obviously this is a key requirement*
Related to other items	Statements on the link of the item with others	NA	NA	*Peer reviewers should have understanding of research methodology as well*
Responsibility	Statements on editor or author´s responsibility to evaluate the item	Author's responsibility	NA	*Failures in this can be about whether the authors have communicated their work clearly*
		Editor's responsibility	NA	*They have to try to understand it - but sometimes they do not. This is where the editor must cast a critical eye to ensure the*

				reviewer has grasped the essence of the article.
Review as guide for editors	Statements on how a review is a guidance for editors		NA	*Extremely important. Nothing more annoying to an author than realising the reviewer has not fully read the paper. Also crucial if the review is to provide fair guidance for editors*
Reviewer as disadvantaged position	Statements on how the peer reviews is often in a disadvantaged position		NA	*The reviewer is often at a disadvantage as he/she is given limited information on which to make a decision on whether to accept or reject the offer to review*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers		NA	*Yes, but see issue above about late night reviewing*
Reviewers as readers proxy	Statements on how the reviewer acts as a proxy for the reader		NA	*The peer reviewer acts as a proxy for the reader, so a basic understanding of the manuscript's content is important*
Reviewers' expertise	Statements on different reviewer´s expertise	Assessment reviewers' expertise	NA	*This is very important. I've long thought that one of the review criteria should be self-ratings of the reviewer's expertise in the substantive and methodological aspects of the article*

		Declaration of competence by reviewers	NA	*Reviewers should declare their competence in the subject of the manuscript*
		More reviewers	NA	*Some reviewers know about methods and some about content. It would be ideal to always have both, but that is often not the case*
		Understanding also research methodology	NA	*Peer reviewers should have understanding of research methodology as well.*
Wording of the item	Statements on how to improve the wording of the item	Confusing item	NA	*I found this question confusing. Are you asking if the reviewer is competent to evaluate the content of the manuscript?*
		Difficult to assess	NA	*Not sure how you would know if the reviewer knows and understands correctly the content of the manuscript*
		Disagreement with the wording	NA	*"Knows" and "understands" are distinct concepts and should not be combined here*

Timeliness (n= 49)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Better quality rather than on time	Statements on how a high quality review is more important than an on-time review	NA	NA	*But it is better to wait a while and have a high-quality review than to receive a quick, superficial and/or unfair review.*
Depends on the delay	Statements on how the importance of the item depends on the type of delay	NA	NA	*A few additional days of delay is not a major issue, while months of delay are*
Dependency on the type of journal	Statements on how biomedical journals differently evaluate the item	NA	NA	*Less important for pre-prints or F1000Research*
Difficult for editors	Statements on how long delay can cause difficulties to the editor	NA	NA	*Difficult for the editor if the delay is too long (or, worse, need to find another reviewer)*
Feasible and flexible deadlines	Statements on the importance to provide reasonable deadlines	Tendency to give short deadlines	NA	*Important, but there seems to be a trend among the editors to get reviews done in shorter amounts of time. Reminders are very helpful, but also some flexibility*
General comments	General statements	NA	NA	*The peer review process needs to be helpful for getting quality research into the public domain in a timely manner. It should not be a road block*
Golden rule	Statements on how to be on time is the golden rule	NA	NA	*It's the Golden Rule - it's just polite to be on time! Do as you would be done by etc.*

Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Very important for us as we try to provide a rapid response to the outcome of papers*
Journal´s reputation rather than good science	Statements on how the journal´s reputation is more important than good science	NA	NA	*Far too many editors now are asking for reviews to be complete in too little time. This is being done for the good of the journal's reputation, not for the good of science*
More time does not mean more quality	Statements on how giving more time does not means having more quality	NA	NA	*Increasing time for revision doesn't add anything to the quality*
Nor related to the quality of PR process	Statements on how the item is not related to the quality of the entire peer review process	NA	NA	*Again not sure this contributes directly to the quality of the peer review process but is important in terms of ensuring that publication timetables can be adhered to*
Orthogonal factors to review quality	Statements on orthogonal factors to review quality	NA	NA	*Availability, timeliness, reliability (reviewing when says will review) are orthogonal to review quality.*
Related to reviewer´s professionalism	Statements on how to be on time is related to reviewer´s professionalism	NA	NA	*This is more on the professionalism of the reviewer rather than the quality of the review*
Reviewer as extra unpaid job	Statements on the voluntary job of reviewers	Demanding work	NA	*There are heavy demands on our time. I find to carry out a quality review on a paper which needs*

				expert reviewing to achieve the two aims laid out initially, that it takes a minimum of half a day, and often a day to do a good job
		Difficult to find reviewers	NA	*We are all busy! biggest issue is finding someone to say yes in first place*
		Little delay	NA	*On time is pretty important but a little delay really is not a problem. I'm speaking as a journal editor... I always have plenty to do so a week delay is probably fine and will not cause authors too much pain*
		Reasonable time for the reviewer	NA	*But are the deadlines reasonable?*
Scope of the peer review	Statements on the scope of the peer review process	NA	NA	*Of course, with the caveat that peer review is voluntary and usually being fitted in around other work activities*
Time given by the journal	Statements on how journals give different deadlines	NA	NA	*Depends on how much time the journal gives. 2 weeks is not enough!*
Wording of the item	Statements on how to improve the wording of the item	Unclear item	NA	*Important to define what "on time" means. For example, is one day late a problem? Or a week late if the authors tell you?*

Tone (n= 40)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Academia as though environment	Statements on the aggressiveness and competition in academia	NA	NA	*There is enough competition and aggressiveness in academia without having to receive reviews which are rude or condescending.*
Dependency on the author´s work	Statements on how the tone used by reviewers depends on the author´s work	NA	NA	*Depends on the author, and how much rubbish is being put forward*
Dependency on the paper quality	Statements on how the tone depends on paper quality	NA	NA	*And how much rubbish is being put forward*
Editor's responsibility	Statements on the editor´s responsibility to tone down the peer review reports	Removal comments by editors	NA	*Editor can tone down or edit out obnoxious comments, but it is better not to have to do this*
Golden rule	Statements on how to use a courteous tone is the golden rule	NA	NA	*Golden Rule again. Hiding rudeness behind anonymity is odious*
Hiding behind anonymity	Statements on how peer reviewers use anonymity to hide rudeness	NA	NA	*All too often, peer reviewers use the anonymity of the review process as an excuse to be rude and scathing in a way they would be unlikely adopt in person*
Impact of a rude review	Statements on how rude reviews can impact authors	NA	NA	*And if authors receive discourteous comments, this really does 'sour' the process*

				and put people off, especially new researchers
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*This is important. Especially to keep his/her nerves when looking at the first revised version and notice that the revision is not answering the queries and comments! Should also provide authors the keys to improve the paper and answer politely to referees...*
Not always a necessary requirement	Statements on how courteous tone is not always a necessary requirement	NA	NA	*This is nice, but not totally necessary*
Related to cultural differences	Statements on how courteous tone is culturally bound	NA	NA	*I think this is important, but courtesy is culturally bound*
Related to other items	Statements on the link of the item with others	NA	NA	*This relates to constructiveness, above*
Requirement	Statements on the requirement to use a courteous tone in a peer review report	NA	NA	*But it's an ethical requirement, and helps improve everyone's experience*
Review quality is important than courteous tone	Statements on how an higher quality review is more important than a courteous review	NA	NA	*Would you rather be treated by skilled, but rude, surgeon, or by a courteous flop?*
Useful for authors	Statements on the usefulness of the item for the authors	NA	NA	*The reviewer's aim should be to give comments that make the next version of the*

				manuscript better (whether or not it is accepted for that specific journal)
Wording of the item	Statements on how to improve the wording of the item	Difficult to define	NA	*It is impossible to define 'courteous' so I doubt that this is operationalisable*

Codebook 3. Identification of new items to assess peer review report quality

Identification of new items to assess peer review report quality (n=152)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Characteristics of reviewer's comments	Statements on the characteristics of the comments made by a peer reviewer	Clarity	Clarity of the expected changes	*Being clear about the changes they want to see (vs. vague comments about weaknesses - what would most strengthen)*
			Clarity of the language	*The peer review report should be comprehensive and written clearly. It should not be ambiguous.*
		Constructiveness	NA	*Constructive attitude should include orientation and suggestion to authors to improve the manuscript.*
		Detailed	NA	*Specific details for concerns and suggestions for improvement are the keys for me. I need substantive concerns.*
		Evidence-based	NA	*A good peer-review report also includes references*
		Fairness/Unbiased	NA	*The comments should not only be fair, but also unbiased.*
		Specificity	NA	*I think that the challenge is that the comments are context*

				specific. Reviews are helpful when they identify something that is a challenge or something that can be done better. These rely upon the context.
		Structure of the peer review report	Additional comment to the editor	*Provides additional comments to the editor that provides context to the reviewer's assessment.*
			Explicit recommendation	*The reviewer makes an explicit recommendation about what decision to make – i.e., "reject", "revise and re-review", "accept", etc.*
			Initial summary	*The reviewer should begin her report with a short synthesis of the study (goals and main results)*
			Length of comments	*Length, very short peer review reports make me suspicious that they have even read the paper in enough detail. E.g. I once received a peer review "report" that had a single sentence along the lines of: "Good methods and results".*
		Tone	NA	*Candid*
Related to peer review process	General statements on the peer review process	Anonymity _PR process type	NA	*Reviewers and authors should be anonymous on both sides*

		Dependency on the type of journal	NA	*Reviewers should understand the nature of the journal that the manuscript had been submitted to - and should be prepared to state if he / she thinks the paper is not appropriate or relevant to that journal's readership.*
		Disclosure of reviewer's COI	Editor's task	*Conflict of interest. We often can tell that the journal has sent our paper to a reviewer who will not be objective in their review, and sometimes even when we've asked the editor to not use a particular reviewer. Editors have an obligation to insure a fair review, and often they do not. In these instances, the outcome is a foregone conclusion*
			Requests motivated by reviewer's COI	*The reviewer does not make requests that seem to be motivated by a competitive attitude or a conflict of interest.*
			Reviewer's publication record	*Publication record of the reviewer*
		Editor's responsibility	Balanced and fair decision	*Editor' decisions should also be balanced and fair, especially when reviewing are*

				discordant... Editors should also read papers... .
			Filtering reviews	*Editors need to protect authors from poor reviewers.*
		High profit of scientific publishing industry	NA	*The scientific publishing industry makes very high profits, in fact it is the most profitable "legal business model" among all economic activities.*
		Peer reviewers' training	NA	*A good reviewer needs to be trained: should be important to organize courses*
		Poor quality of the second review	NA	*Completion of a second review after the first draft-this is often poorly done*
		Quality scale	NA	*It might be helpful to consider one of the research quality ratings scales that are used in quantitative reviews.*
		Review quality as usefulness to make an editor's decision	NA	*The quality of a peer-review report that an author receives is partially determined by what the editors contribute to it before sending it to the author and how quickly they use it to make a decision. Too many don't send enough feedback, especially when two or more reviewers disagree.*

				This leads to three, four, or more back-and-forth "reviews" where reviewers are trapped in a cycle of disagreement and the editors won't make any significant contributions or a decision to resolve the disagreement.
		Reviewer as unpaid extra job	NA	*Reviewers are scientists that perform a professional service for the scientific publishing industry that in the vast majority of the cases is not paid.*
		Reviewer's final choice	Difference between major and minor revisions	*Always a grey area between Major revisions and Minor revisions that foxes a reviewer*
			Explanation choice	*If the reviewer makes a recommendation, e.g. accept or reject, they must provide a reason why. A review that just says "accept", "good work", is not a valid peer review.*
		Reviewer's recognition	Professional evaluators by publishers	*Alternatively, publishers may turn to "professional evaluators", who they may find in consultancy firms (KPMG, McKinsey, etc.) and pay their fees.....*
			Rewards for reviewers	*Payment for the reviewers should be considered, and this*

				would also increase the quality of the evaluations
		Saturation of the system	Professional evaluators by publishers	*So far reviewers are working "pro bono", and with the tremendous growth in the number of journals and the overall increase of the scientific activity worldwide (think just of the soaring number of papers coming from China in recent years) the system is becoming saturated, and reviewers becoming fed up*
		Scope of review	NA	*The number of items listed is a factor. If too many and in too minute detail, the article could be suppressed by the sheer workload of trying to address the comments*
		Weighting reviewer's comments	NA	*It is important for the ae to weight reviewers comments - some are rubbish and can be disregarded*
Related to the study	Statements on different aspects of a study that should be commented in a peer review report	About references	Suggesting relevant references	*Including references not known to the author*
		Addressing study's aims	NA	*I think the 'does this study address its stated aims' issue that I raised in my earlier responses is very important*

		Adherence to ethical guidelines	NA	*Comment on the study's adherence to ethical guidelines*
		Appendices	NA	*Use of supplemental material/appendices when appropriate (e.g., sensitivity analyses)*
		Applicability of the study	NA	*And general applicability*
		Authors' contribution and acknowledgments	NA	*Clearly articulate the role of every team member, and their contribution to the study. For evidence syntheses, require librarian involvement and give them authorship, the same with statisticians. Everyone in the team, without whose knowledge the study would not be possible, sound, or complete, should be acknowledged.*
		Context of the study	NA	*Puts the study in appropriate context*
		Data availability and software	NA	*Referees check the data availability and if new software actually works*
		Data quality	NA	*Quality of the data is most important*
		Ensuring disclosure of COI	NA	*Conflict of interests could be included*

		Ensuring inclusion of data sharing statements	NA	*Reviewers should ensure data sharing statements are included*
		Ensuring language quality	NA	*Comment on readability*
		Ethics	NA	*Ethical considerations of research*
		Importance of methods	NA	*Perhaps reviewing upto methods and evaluating the study that way is worth more consideration.*
		Literature is adequately reviewed	Most recent research	*Reviewer rating of whether The authors discuss the most recent relevant research on the topic*
		Originality	NA	*The added value of the study to what is already known.*
		Potential impact	NA	*Potential impact of study*
		Presentation (tables and figures)	NA	*And appropriateness of accompanying visual aids (graphs, tables e.t.c.).*
		Publication study's protocol and deviation from it	NA	*Whether a protocol was lodged in publication or on an independent site e.g. OSF and whether it matches the paper and if not, if reporting of deviations is transparent.*

		Relevance	NA	*relevance*
		Replicability/Reproducibility	NA	*Whether the study can be replicated on current methods whether limitations are acknowledged (this was covered actually I think)*
		Study conclusions	NA	*And finally if the conclusion answers the research question.*
		Study introduction	NA	*If the in introduction leads to the research question*
		Study limitations	NA	*Whether limitations are acknowledged*
		Study theoretical framework	NA	*Logic of the theoretical framework*
		Study weaknesses	NA	*Reviewer comments on the limitations of the study*
		Suggestions for future studies	NA	*Suggestions for future studies*
Reviewer's expertise		Considering reviewer's expertise	NA	*I have experienced vastly different qualities of reviews, so I think that the reviewer's expertise in the area of the article needs to be considered.*
		Focus on the points suggested by editor	NA	*Suggestions from editor re: which points/themes to focus on*

		Knowing dimensions not assessed by reviewers	NA	*What the reviewer feels they cannot comment on (e.g. is outside their expertise)*
		Rating or commenting on own level of expertise	NA	*The reviewer should state those aspects of the study for which they have limited knowledge*
		Reviewer's type	NA	*Items need to be tailored for whether the reviewer is a stats (methodology) reviewer or clinical expert*

NA= not available

Supplementary file 6. New items suggested by survey participants

New items	Example
1. Adherence to ethical guidelines	*"Comment on the study's adherence to ethical guidelines"*
2. Author's contribution and acknowledgements	*"Clearly articulate the role of every team member, and their contribution to the study. For evidence syntheses, require librarian involvement and give them authorship, the same with statisticians. Everyone in the team, without whose knowledge the study would not be possible, sound, or complete, should be acknowledged."*
3. Data availability and software	*"Referees check the data availability and if new software actually works"*
4. Disclosure of COI	*"Conflict of interests could be included"*
5. Data sharing statements	*"Reviewers should ensure data sharing statements are included"*
6. Study protocol	*"Whether a protocol was lodged in publication or on an independent site e.g., OSF and whether it matches the paper and if not, if reporting of deviations is transparent."*
7. Addressing study aims	*"I think the 'does this study address its stated aims' issue that I raised in my earlier responses is very important"*
8. Study introduction	*"If the in introduction leads to the research question"*
9. Study limitations	*"Whether limitations are acknowledged"*
10. Study conclusion	*"And finally if the conclusion answers the research question."*
11. Theoretical framework	*"Logic of the theoretical framework"*
12. Relevant literature	*"Reviewer rating of whether The authors discuss the most recent relevant research on the topic"*
13. Reproducibility	*"Whether the study can be replicated on current methods"*

Supplementary file 7. Explanations of the items included in the ARCADIA tool

Domain 1: Importance of the study

Item 1.a Contribution

A study can contribute to scientific knowledge in many ways: it can be a novel or confirmatory study with little or great impact on society and/or the research community. The contribution of a study is therefore not only associated to its novelty. Studies also need to be replicated in order to verify the validity of their results. The peer reviewer should discuss the importance of the study's research question.

Item 1.b Relevant literature

The peer reviewer should check if the authors reviewed the relevant research related to the study's topic in order to situate the study within the context of the existing literature.

Domain 2: Robustness of the study methods

Item 2.a Study methods

The peer reviewer should evaluate the soundness of the study methods, such as the selection of the study design, assessment of the risk of bias, etc., to understand whether the methods were appropriate to the study's aims, as well as if they were properly used and reported.

Item 2.b Statistical methods

Data can be analysed in many ways, but the only appropriate statistical models are those that fit well with the study design and the characteristics of the variables. The peer reviewer with expertise in statistics should assess whether or not the study followed a suitable statistical procedure, as well as if they were correctly conducted and reported. A reviewer should clearly state if she/he is qualified to review the statistics of a study.

Domain 3: Interpretation and discussion of the study results

Item 3.a Study conclusions

The reviewer should verify if the conclusions answer the research question(s) and correctly summarize the study results.

Item 3.b Study limitations

The reviewer should check if the weaknesses of the study are correctly identified and discussed in order to interpret the validity of the research.

Item 3.c Applicability and generalizability

The reviewer should comment on the applicability and generalizability of the study results. Applicability and generalizability are two underlying concepts of external validity [1]. The first concerns how "the results from a sample can be extended to the population from which the sample was drawn", while the second how "the inferences drawn from study participants can be used in the care of patients drawn from any populations" [1].

Domain 4: Reporting and transparency of the manuscript

Item 4.a Study protocol

Public access to study protocols is important to increase transparency and reduce waste of biomedical research. In the case of previous publication and/or inclusion as an additional file of a study protocol, the reviewer should verify that the major deviations from it are reported in the manuscript.

Item 4.b Reporting

The reviewer should comment if the reporting of the study is clear, complete and transparent enough for facilitating its reproducibility by verifying the adherence of the manuscript to the corresponding reporting guideline. The Enhancing the Quality and Transparency of Health Research (EQUATOR) Network provides a toolkit to be used during the peer review process for selecting the appropriate reporting guideline [2].

Item 4.c Presentation and organization

The reviewer should discuss the quality of the written language used in the manuscript, as well as of how the study results (tables, figures, etc.) are presented.

Item 4.d Data availability and software

When applicable, the reviewer should ensure that the study data and materials are available and the software work as indicated.

Domain 5: Characteristics of the reviewer's comments

Item 5.a Clarity

A peer review report should be clear, succinct and well organized in order to be understood correctly by editors and authors.

Item 5.b Constructiveness

A peer review report should contain constructive and polite comments that allow the authors to improve the quality of their work.

Item 5.c Objectivity

Comments provided in a peer review report should be as objective as possible and, if considered appropriate, include references to support the reviewer's statements.

References

1. Murad MH, Katabi A, Benkhadra R, Montori VM. External validity, generalisability, applicability and directness: a brief primer. BMJ Evid-Based Med. 2018;23(1):17–9.
2. EQUATOR Network. Welcome to our toolkit for peer reviewing health research! [Internet]. Available from: https://www.equator-network.org/toolkits/peer-reviewing-research/

BMJ Open

The development of ARCADIA: a tool for assessing the quality of peer review reports in biomedical research

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-035604.R1
Article Type:	Original research
Date Submitted by the Author:	27-Apr-2020
Complete List of Authors:	Superchi, Cecilia; Barcelona-Tech, Statistics and Operations Research Department; Université de Paris 1 Hren, Darko; University of Split, School of Humanities and Social Sciences Blanco, David; Barcelona Tech, Statistics and Operations Research Department; Université de Paris 1 Rius, Roser; Barcelona-Tech, Statistics and Operations Research Department Recchioni, Alessandro; Springer Nature Boutron, Isabelle; Université de Paris 1; Hôpital Hôtel-Dieu de Paris, Centre d'épidémiologie Clinique González, José Antonio; Barcelona-Tech, Statistics and Operations Research Department
Primary Subject Heading:	Medical publishing and peer review
Secondary Subject Heading:	Research methods, Epidemiology
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, PUBLIC HEALTH, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The development of ARCADIA: a tool for assessing the quality of peer review reports in biomedical research

Authors: Cecilia Superchi[1,2], Darko Hren[3], David Blanco[1,2], Roser Rius[1], Alessandro Recchioni[4], Isabelle Boutron[2,5†], José Antonio González[1†]

[†]Equal contributors

Affiliations:

[1] Statistics and Operations Research Department, Barcelona-Tech, UPC, c/ Jordi Girona 1, C5- 213, 08034, Barcelona, Spain

[2] Université de Paris, CRESS, INSERM, INRA, F-75004 Paris, France

[3] Chair of Psychology, Faculty of Humanities and Social Sciences, University of Split, Split, Croatia

[4] Springer Nature, Heidelberger Platz 3, 14197 Berlin, Germany

[5] Centre d'épidémiologie Clinique, Hôpital Hôtel-Dieu, Assistance Publique Hôpitaux de Paris (APHP), F-75004 Paris, France

Corresponding author:

Cecilia Superchi, Statistics and Operations Research Department, Barcelona-Tech, UPC, c/ Jordi Girona 1, C5- 213, 08034, Barcelona, Spain. cecilia.superchi@upc.edu, +34 631654317

Abstract

Objective: To develop a tool to assess the quality of peer review reports in biomedical research.

Methods: We conducted an online survey intended for biomedical editors and authors. The survey aimed to 1) determine if participants endorse the proposed definition of peer review report quality; 2) identify the most important items to include in the final version of the tool; and 3) identify any missing items. Participants rated on a 5-point scale whether an item should be included in the tool and they were also invited to comment on the importance and wording of each item. Principal component analysis (PCA) was performed to examine items redundancy and a general inductive approach was used for qualitative data analysis.

Results: A total of 446 biomedical editors and authors participated in the survey. Participants were mainly male (65.9%), middle-aged (mean=50.3, SD=13) and with PhD degrees (56.4%). The majority of participants (84%) agreed on the definition of peer review report quality we proposed. The 20 initial items included in the survey questionnaire were generally highly rated with a mean score ranging from 3.38 (SD=1.13) to 4.60 (SD=0.69) (scale 1 to 5). Participants suggested 13 items that were not included in the initial list of items. A steering committee composed of five members with different expertise discussed the selection of items to include in the final version of the tool. The final checklist includes 14 items encompassed in five domains (Importance of the study, Robustness of the study methods, Interpretation and discussion of the study results, Reporting and transparency of the manuscript, Characteristics of peer reviewer´s comments).

Conclusion: ARCADIA tool could be used regularly by editors to evaluate the reviewers' work, and also as an outcome when evaluating interventions to improve the peer review process.

Words count: (abstract: 290, word limit: 300), (main text: 4015, word limit: 4000), 4 tables, 1 figure, 1 web application, 7 supplementary files

Keywords: Peer review, Report, Checklist, Quality control, Survey

Strengths and limitations of this study

- ARCADIA constitutes the first tool that has been systematically developed to assess the quality of peer review reports.
- Its development is based on an exhaustive review of the literature and on empirical data from a large and heterogeneous sample of both biomedical editors and authors.
- The majority of editors and authors were from Europe and North America, which may limit the generalizability of the results.
- ARCADIA has not yet been validated.

Background

Editorial peer review stands as the gateway to scientific publication. The process was established to ensure that research papers are vetted by independent experts before they are published, although it is recently being increasingly questioned due to beliefs that it is flawed [1,2] . Despite efforts over the last 30 years to "make peer review scientific", its impact is still considered suboptimal [3].

Peer reviewers, who are the pivotal actors in this process, are requested to write a review report evaluating the submitted manuscript. A peer review report helps authors improve the quality of their manuscripts, and it also helps editors make an informed decision about the outcome of the manuscript. However, evidence shows that these peer review reports are often of poor quality [4,5].

Tools for assessing the quality of peer review reports have been proposed, of which we have conducted a systematic review and identified 24 tools: 23 scales and 1 checklist [6]. However, none reported any definition of peer review report quality, only one described the scale development, and 10 provided measures of reliability and validity. Further, the development and validation process resulted from a small consensus of people, and the concepts evaluated by these tools were quite heterogeneous.

In 2016, Bruce et al. published a review evaluating the impact of interventions to improve the quality of the peer review process [5]. The authors showed that it is essential to clarify the outcomes (such as, for example, the quality of peer review reports), which should be used in randomized controlled trials to evaluate these interventions.

A validated tool is direly needed to clearly define the quality of a peer review report in biomedical research. This tool could be used regularly by editors to evaluate the reviewers' work, and also as an outcome when evaluating interventions to improve the peer review process. In the present study, we report on the development of a new tool to assess peer review reports in biomedical research.

Methods

The study was approved by the Research Committee of the Governing Council of the Universitat Politècnica de Catalunya, Barcelona Tech, Spain (Reference: EC 02, Date: 02/05/2018).

Steering committee

We formed a steering committee of five members (CS, DH, AR, IB and JAG), whose expertise include clinical epidemiology, biostatistics, social science and editorial peer review. The steering committee agreed on how to define peer review report quality; they agreed on the survey questionnaire based on the results of a previous systematic review [6]; they interpreted the results of the survey; and they agreed on the final version of the tool.

Defining the tool's objective

The tool aims to assess the quality of peer review reports in biomedical research. We defined the quality of a peer review report as "the extent to which a peer review report helps editors make a fair decision and authors improve the quality of the submitted manuscript".

Generating the items

A systematic review allowed the identification of 24 tools, aimed at assessing the quality of peer review reports [6]. We extracted 132 items from such tools. After removing the redundant items, we obtained 17 items. We then eliminated two items and incorporated five new ones that met our definition of peer review report quality, after piloting the survey questionnaire and discussing with the steering committee. Overall, 20 items were identified to assess peer review report quality (Table 1).

Survey

We conducted an online survey of editors and authors in order to: 1) determine if they endorse the proposed definition of peer review report quality; 2) identify the most important items to include in the final tool; and 3) identify any new items that should be included.

Survey questionnaire

The questionnaire was constructed using the online survey software SurveyMonkey [7]. It was structured into four main parts and included both open and multiple-choice questions. First, the participants were asked to agree ("yes/no/partially") on the definition we provided for peer review report quality. They were also invited to add any comments or ideas on how to improve the

definition. Second, they were asked to rate the importance of the 20 items for assessing the quality of peer review reports we identified. Their responses were based on a 1–5 Likert scale (1 being not important and 5 very important). In particular, we asked the participants if the item should be included in a tool for assessing the quality of peer review reports. Moreover, they were invited to comment on the importance and wording of each item. In order to eliminate the question order effect, the items appeared in random order for each respondent. Third, the participants were invited to suggest any additional items missing that they considered important for assessing the quality of peer review reports. Finally, the questionnaire included nine demographic questions related to sex, age, education level, job title, referring institution and job experience as biomedical editor and/or author. We developed two versions of the questionnaire because biomedical editors and authors were recruited differently, despite the fact that some of them could play both roles (see Supplementary file 1). The two versions were structured in the same way; they only differed in some questions related to the demographic characteristics. The questionnaire was piloted among six experienced scientific editors and authors, followed by a subsequent revision based on their feedback.

Participants and recruitment strategy

We targeted biomedical editors and authors using a purposive sampling approach to recruit a heterogeneous sample of information-rich cases [8].

Biomedical editors

By means of standardized email, we invited two groups of editors to participate in the survey: 586 biomedical editors from 43 journals in the BMJ Publishing group; and 478 editors from 235 journals identified in a previous cross-sectional bibliometric study [9] (see Supplementary file 2). The survey was also distributed to 27 editors from 48 journals in BMC (part of Springer Nature), using internal email, and to members of the European Association of Science Editors (EASE) through their newsletter. In the invitation email and newsletter, the editors were encouraged to forward the survey to colleagues who might be interested in issues related to peer review. This recruitment strategy, known as snowballing, allowed us to identify "information-rich key informants" among biomedical editors [8]. On the first page of the survey, participants were informed that the collected data would be anonymous, and they were further asked if they would agree to share their de-identified data in an open access repository. Two reminder emails were sent to non-respondents. Finally, the survey was promoted on Twitter and on the EASE blog [10] and Methods in Research on Research (MiRoR) [11] websites.

Authors

Searching the top 30-biomedical journals with the highest impact factors, we identified 4396 corresponding authors of articles that reported original research and which were published in Medline between February 1 and October 31 2018 (see Supplementary file 3). We used the R package easyPubMed to extract the email contacts [12]. The corresponding authors received a standardized email that explained the purpose of the study and included a link to the survey (see Supplementary file 2). The first page of the survey informed participants that the data were collected anonymously and also asked if they would agree to share their de-identified data in an open access repository. Two reminder emails were sent to non-respondents.

We did not use a snowballing strategy to recruit authors. However, since the survey directed to biomedical editors was promoted on Twitter by different users who sometimes did not provide thorough instructions, we included in the first page of the survey, also the link to the questionnaire addressed to authors. This was done so that a researcher, who was not an editor and mistakenly opened the link to the survey questionnaire, was still able to participate to the study as biomedical author.

Data analysis

We described the demographic data in terms of frequencies and percentages. The importance of the 20 items to assess peer review report quality is described in means and proportions of editors or authors who rated the importance of the items from 1 to 5. The items were also sorted according to the mean raking of all participants and either editors or authors. We also calculated Pearson correlations among items. The calculations and graphical representations were all obtained using the statistical software R 3.5.0 [13].

Principal component analysis of quantitative data

We conducted a principal component analysis (PCA) to examine item redundancy among the 20 items to assess peer review report included in the survey. PCA is a multivariate statistical technique used to reduce the number of variables in a dataset to a smaller number of dimensions [14]. The new dimensions (or *principal components*) are mutually independent and are determined by choosing the directions that explain the most variation in the data. The first principal component (PC1) accounts for the largest possible variance in the data, and each succeeding PC accounts for decreasing amounts of the remaining. This exploratory analysis helps reveal simple underlying structures in complex datasets. We performed PCA using the R package FactoMineR [15].

Inductive content analysis of qualitative data

We used a general inductive approach for qualitative data analysis. In particular, we followed the five steps of inductive analysis proposed by David R. Thomas: 1) Preparation of raw data files; 2) Close reading of text; 3) Creation of codes; 4) Overlapping coding and uncoded text; 5) Continuing revision and refinement of themes system [16]. In the third phase, two investigators (CS and DB) created independently the initial codes from the responses of the first 100 participants for each open-ended question. In order to ensure consistency and credibility, the initial codes were discussed with a third investigator (DH) and a codebook was developed and was used for analysing the remaining responses. In case new codes were successively created from the remaining responses, the emerging codes were added to the codebook and applied to entire dataset. Two investigators (CS and DH) reviewed and refined the codebook and further clustered the codes into major themes. We used the software NVivo V.12 for data management and analysis [17].

Selecting items

The steering committee reviewed all items and, ultimately, drafted and refined the final version of the tool. Based on the participants' qualitative and quantitative answers, redundant items were combined, existing items were modified and/or expanded on, and new items proposed by survey participants were added.

Patient or public involvement

Patients and members of the public were not involved in the study.

Results

Participants

Between November 7 2018 and February 4 2019, 198 biomedical editors and 248 authors completed the survey. Of the 1134-biomedical editors and 3633 corresponding authors invited via email, 89 (7.8%) and 238 (6.5%) completed the survey, respectively. In addition, 109 editors and 10 authors completed the survey using the web link.

Participants were mainly male (263/399, 65.9%) with a PhD degree (225/399, 56.4%), and their ages were equally distributed across ranges (mean=50.3, SD=13). They were mainly located in Europe (219/389, 56.3%) and North America (118/389, 30.3%). More than half of the editors had work experience of more than 5 years (91/165, 55.2%), while over one-third of the authors had

work experience of more than 20 years (84/224, 37.5%) (see Table 2). Editors were mainly associate editors (63/165, 38.2%) and editors in chief (50/165, 30.3%), primarily involved in making decisions on the submitted manuscripts (144/165, 87.3%). Most of them worked in specialty journals (126/165, 76.4%) and they were used to contribute as authors in scientific papers (141/165, 85.5%). The corresponding authors were mainly professors (63/224, 28.1%), but also PhD students, postdocs or lecturers (49/224, 21.9%) or researchers (47/224, 21%). The majority of them worked in public universities (134/224, 59.8%) and they were not employed as editor (161/224, 71.9%) in biomedical journals. Among those who also work as biomedical editors (63/224, 28.1%), 88.9% of them are involved in making decision on the manuscript (Supplementary file 4).

Definition of peer review report quality

Overall 84% (362/431) participants, precisely 85% (160/188) editors and 83% (202/243) authors, agreed on the definition of peer review report quality that we provided in the survey. The definition was slightly modified to take into account participants comments (Supplementary file 5). The quality of a peer review report is now defined as "the extent to which a peer review report helps, first, editors make an informed and unbiased decision about the manuscripts' outcome and, second, authors improve the quality of the submitted manuscript".

Quantitative results

We created a web application that is publicly available at https://www-eio.upc.edu/redir/ReportQuality. Through the application, the readers can easily access and explore the quantitative results of the survey.

Rating the importance of items

The items were generally highly rated, with a mean score ranging from 3.38 (SD=1.13) to 4.60 (SD=0.69). All the items were scored 4 or 5 by >50% of the participants (see web application). The three items rated as the most important were: 1) *Knowledgeability*; 2) *Methodological quality*; and 3) *Fairness*. The three least important items were: 1) *Originality*, 2) *Presentation and organization*; and 3) *Adherence to RG*.

A peer review report aims to help authors improve their submitted manuscripts and assist editors in taking editorial decisions. Due to this dual objective, we compared editors' and authors' mean

scores in order to investigate whether any difference is found in their perceptions regarding the importance of the 20 items that assess peer review report quality. We found little discrepancy in the mean scores between biomedical editors and authors, with only two items indicating any difference: 1) *Timeliness* and 2) *Detail/Thoroughness*. The *Timeliness* of the peer review report was considered more important to authors than to editors (respectively, in the 12th and 16th rank positions). Meanwhile, editors rated the *Detail/Thoroughness* of the reviewer's comments higher than did authors (respectively, in the 11th and 16th rank positions).

Correlations among items

Overall, we found relatively weak positive correlations among items. The largest positive correlations were found between *Relevance* and *Originality*, and between *Fairness* and *Objectivity* (r = 0.55 and 0.43, respectively).

Principal Component Analysis

The first principal component (PC1) accounted for 22.1% of data variability. The next two dimensions (PC2 and PC3) accounted for 38.5% of the cumulative variability and contributed gradually, that is, they increased at only small increments. PC1 was positively correlated to all items (or variables), and it showed correlations higher than 0.4 —which is the figure commonly used as a threshold reference for factor loadings — for 16 out of 20 items (see web application). These results illustrate that the data variance was not concentrated in a few components but distributed across all of them; hence, reducing the number of items is not recommended, since this would imply an important loss of data information.

The study of the supplementary variables did not reveal any differences between authors and editors in terms of items rating. However, we found that female participants above the age of 55 years old generally provided higher rating for the items, compared to younger male participants.

Qualitative results

Comments on importance and/or wording of items

Out of 446 survey participants, 267 (59.9 %) made at least one comment on the importance and/or wording of the items. Based on the initial coding of the comments, we were able to identify eight general themes that they addressed: Peer reviewer; Wording; Importance; Dependency; Responsibility; Item; Structure and content; and Improvement. Table 3 reports the eight themes

together with their definition and the most frequent codes (n>5), with example quotes. The entire codebook is found in Supplementary file 5.

New items

Participants suggested 13 items that were not included in the initial list of items. These items are listed in Supplementary file 6. The entire codebook is found in Supplementary file 5.

Steering committee meeting

The steering committee met on the 19/07/2019 to discuss the selection of items to include in the final version of the tool. Their decisions were based on the participants' quantitative and qualitative answers. The flow of the items is summarized in Figure 1.

The items *Relevance* and *Originality* were merged into a new item named *Contribution* (of the study). This decision was based on the high positive correlation found between the two items (0.55) and on the participants' opinions. Furthermore, participants suggested in their comments that the item *Relevance* was "*highly subjective*", because "*each reviewer's decision on relevance reflects what is relevant to them, which may not reflect relevance to the journal*". They also believed that the *Originality* of a study is not always an important aspect for comments in a peer review report, because some manuscripts "*are trying to duplicate findings from previous studies*". They therefore suggested reformulating the two items by asking the reviewer what the study "*adds to our knowledge*".

The steering committee decided to include the item *Interpretation of results* as a domain of the tool instead of a single item, changing the name into *Interpretation and discussion of the study results*. This decision resulted from the addition of two new items (*Conclusions* and *Limitations*), based on the suggestions of survey participants. The domain *Interpretation and discussion of the study results* now encompasses three items: 1) *Study conclusions*; 2) *Study limitations and* 3) *Applicability and generalizability.*

Overall, survey participants believed that the items *Strengths and weaknesses (general)* and *Strengths and weaknesses (methods)* were "*confusing to separate*". Additionally, the steering committee agreed that *Strengths and weaknesses (methods) and Methodological quality* were also redundant; thus, it was ultimately decided to merge the three items into a new item named *Study methods.*

The items *Objectivity* and *Fairness* were merged because of both the moderate correlation between them (0.43) and the participants' opinions. Participants suggested that the total objectivity of the reviewer's comments is not possible because "*all decisions contain some personal biases and subjectivity*" and they also believed that the term fairness was "*very subjective*" and difficult to define. Additionally, the steering committee agreed to also combine these two items into *Supported by evidence*. The committee finally decided to merge all three items into *Objectivity*, and this was defined as "comments provided in a peer review report should be as objective as possible and, if considered appropriate, include references to support the reviewer's statements".

The steering committee agreed to merge *Structure of reviewer's comments* and *Clarity*, because participants considered both important for making the peer review report easy "*to read for both editors and authors*". Moreover, participants suggested that the *Detail/Thoroughness* of a peer review report was mostly associated with the quality of a manuscript, because in certain occasions a study can be so poorly conducted that "*a reviewer can highlight one or two major methodological flaws*" without conducting a detailed review. They therefore believed that a detailed report is not "*always necessary*" and instead preferred a succinct report that "*cuts straight to the critical points*". Taking into account the participants' opinions, the steering committee finally decided to include a single item named *Clarity*, which is defined as "a peer review report should be clear, succinct and well organized in order to be understood correctly by editors and authors".

The items *Tone* and *Constructiveness* were merged into *Constructiveness*, which is defined as "a peer review report should contain constructive and polite comments that allow the authors to improve the quality of their work". This decision was based on the participants' opinions that "*the comments should be polite and constructive*".

The item *Adherence to RG* and the new item *Reproducibility* suggested by survey participants were merged into *Reporting* based on the steering committee decision. The item *Reporting* was defined as "the reviewer should comment if the reporting of the study is clear, complete and transparent enough for facilitating its reproducibility by verifying the adherence of the manuscript to the corresponding reporting guideline."

The items *Timeliness* and *Knowledgeability* were not included in the final version of the tool. Survey participants suggested that *Timeliness* was not "*directly tied to review quality*" because

"*some of the best reviews come in past the deadline*". Furthermore, the steering committee agreed that the item *Knowledgeability* was generally difficult to assess, because it implied that anyone using the tool would have enough competence to evaluate the reviewer's knowledge and expertise. Five new items suggested by survey participants (*Data availability*, *Study protocol*, *Study conclusions*, *Study limitations* and *Relevant literature*) were finally included in the tool.

The ARCADIA tool

The ARCADIA (**A**ssessment of **R**eview reports with a **C**hecklist **A**vailable to e**DI**tors and **A**uthors) tool was finally developed. The tool is a checklist that includes five domains and 14 items (Table 4). Brief explanations of the items included in the five domains are provided in Supplementary file 7.

Discussion

This study resulted in a checklist of items to assess the quality of peer review reports in biomedical research. The checklist constitutes the first tool that has been systematically developed to assess the quality of peer review reports.

The checklist is simple, applicable to any biomedical field, and consists of five domains covering 14 items, each of which is phrased as a question. Each item should be ticked as 'Yes' or 'No'. An item could be also checked 'NA' if it is not covered in the study (e.g., there are no data or other materials attached to the manuscript) and/or the peer reviewer is not qualified to comment on that specific aspect (e.g., statistical methods). The ARCADIA tool has several strengths. It is the first tool ever developed based on an exhaustive review of the literature [6] and on empirical data from a large sample of both biomedical editors and authors. Further, it is the only tool that clearly defines the quality of peer review reports, as its definition was based on the perspectives of 446 authors and editors.

To develop the tool, we recruited a large sample of biomedical editors and authors with varying experience and backgrounds. We found the percentage of female participants who took part in the survey to be quite low (129/399, 32.3%). This is in line with evidence showing that gender equity in academic medicine careers remains far behind [18]. Moreover, we recruited corresponding authors (who are usually first authors) from the top 30 biomedical journals. Evidence also shows that

women are underrepresented as first authors among biomedical journals with high impact factors [19].

Overall, we did not find any differences between authors and editors in terms of item rating by conducting PCA. Only two items, *Timeliness* and *Detail/thoroughness,* presented a difference according to the separate mean score rankings of authors and editors. Timeliness was considered more important for authors and this could be justified by the fact that authors are usually more interested in receiving decisions about their manuscript as soon as possible. Whereas, editors rated *detail/thoroughness* as more important to them, given thorough and detailed peer review reports help them make a better editorial decision on any given manuscript.

The present study also has some limitations. The survey questionnaire included some open-ended questions, which allowed participants to voluntarily express their opinions. However, we were not able to inquire further to clarify and verify some information provided by the study's participants. Therefore, the interpretation of some information could be affected by the perception of the three investigators who conducted the qualitative analysis. Additionally, since participants could comment voluntarily on the importance and wording of each item, the number of comments among items differed greatly. Furthermore, the majority of editors and authors were from Europe and North America, which may limit the generalizability of the results. This result may be due to the recruitment strategy we used, especially to identify biomedical editors. Although we also utilized a snowballing strategy, we mainly contacted editors through European biomedical journals. Finally, the present study reports on the first version of the ARCADIA tool, which has not yet been validated.

Implications

The tool is a general checklist available to all biomedical editors and authors. It could be regularly used by editors to evaluate the reviewers' work, and it can also be used as an outcome when evaluating interventions in order to improve the peer review process.

Conclusions

ARCADIA is the first checklist that has been systematically developed to assess the quality of peer review reports. It is based on the perspectives of a large and heterogeneous sample of biomedical editors and authors. Our plans for future work are to validate the ARCADIA tool.

Figure caption: Flowchart of items to include in a checklist to assess the quality of peer review reports

Abbreviations: ARCADIA: Assessment of Review reports with a Checklist Available to eDItors and Authors; EASE: European Association of Science Editors; EQUATOR: Enhancing the Quality and Transparency of Health Research; MiRoR: Methods in Research on Research; NA: Not Applicable; PC1: First principal component; PCA: Principal component analysis; RG: Reporting guidelines

Acknowledgements

We thank Patrick Bossuyt, David Moher, Elizabeth Moylan, Albert Selva O'Callaghan, Alessandro Recchioni and José Jiménez Villa for piloting the survey questionnaire. We thank all people who participated in the survey.

References

1. Stahel PF, Moore EE. Peer review for biomedical publications: we can improve the system. BMC Med. 2014 Dec;12(1):179.
2. Smith R. Peer review: a flawed process at the heart of science and journals. 2006;99:5.
3. Rennie, Drummond. Make peer review scientific. Nature. 2016;
4. Jefferson T, Rudin M, Brodney Folse S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. Cochrane Methodology Review Group, editor. Cochrane Database Syst Rev [Internet]. 2007 Apr 18 [cited 2019 Jun 20]; Available from: http://doi.wiley.com/10.1002/14651858.MR000016.pub3
5. Bruce R, Chauvin A, Trinquart L, Ravaud P, Boutron I. Impact of interventions to improve the quality of peer review of biomedical journals: a systematic review and meta-analysis. BMC Med. 2016;14(1):85.
6. Superchi C, González JA, Solà I, Cobo E, Hren D, Boutron I. Tools used to assess the quality of peer review reports: a methodological systematic review. BMC Med Res Methodol. 2019;19(1):48.
7. SurveyMonkey Inc. SurveyMonkey. In San Mateo, California, USA; Available from: www.surveymonkey.com
8. Putton, M. Purposeful sampling. In: Qualitative evaluation and research methods. Sage; p. 169–86.
9. Sharp MK, Tokalić R, Gómez G, Wager E, Altman DG, Hren D. A cross-sectional bibliometric study showed suboptimal journal endorsement rates of STROBE and its extensions. J Clin Epidemiol. 2019 Mar;107:42–50.
10. EASE Blog. Biomedical editors survey on peer review [Internet]. 2018. Available from: https://ese-bookshelf.blogspot.com/2018/11/biomedical-editors-survey-on-peer-review.html
11. MiRoR. Biomedical editors survey on peer review [Internet]. 2018. Available from: http://miror-ejd.eu/2018/11/12/biomedical-editors-survey-on-peer-review/
12. Fantini,Damiano. Search and Retrieve Scientific Publication Records from PubMed [Internet]. 2019. Available from: https://www.data-pulse.com/dev_site/easypubmed/
13. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; Available from: https://www.R-project.org/
14. Jolliffe, I.T. Principal Component Analysis. 2nd ed. Springer; 2002.

15. Husson F, Josse J, Le S, Mazet J. FactoMineR: A Package for Multivariate Analysis. J Stat Softw. 2008;25(1):1–18.
16. Thomas, David R. A general inductive approach for qualitative data analysis. AJE. 2013;27.
17. QSR International. NVivo [Internet]. Available from: https://www.qsrinternational.com/nvivo/home
18. Bates C, Gordon L, Travis E, Chatterjee A, Chaudron L, Fivush B, et al. Striving for Gender Equity in Academic Medicine Careers: A Call to Action. Acad Med. 2016 Aug;91(8):1050–2.
19. Filardo G, da Graca B, Sass DM, Pollock BD, Smith EB, Martinez MA-M. Trends and comparison of female first authorship in high impact medical journals: observational study (1994-2014). BMJ. 2016 Mar 2;i847.

Contributors: All authors provided intellectual contributions to the development of this manuscript. CS, DH, IB and JAG jointly contributed to the study conception, design and interpretation of data. CS conducted the survey. CS, RR and JAG conducted the quantitative analysis and created the web application. CS, DB and DH conducted the qualitative analysis. CS, DH, AR, IB and JAG formed the steering committee. CS led the writing of the manuscript. IB and JAG led the supervision of the manuscript preparation. All authors provided detailed comments on earlier drafts and approved the final manuscript.

Funding: This project was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement no 676207. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscripts.

Competing interests: The authors declared no competing interests.

Patient consent for publication: Not required.

Ethics approval: The study was approved by the Research Committee of the Governing Council of the Universitat Politècnica de Catalunya, Barcelona-Tech (Spain).

Provenance and peer review: Not commissioned; externally peer reviewed.

Data availability statement: The dataset supporting the conclusions of the research reported in this paper will be available in the Zenodo repository in the Methods in Research on Research (MiRoR) community (https://zenodo.org/communities/miror/?page=1&size=20)

Tables

Table 1. The 20 items to assess peer review report quality included in the survey

Labels	Items to assess PR report quality
Relevance	The reviewer comments on the relevance of the study
Originality	The reviewer comments on the originality of the study
Interpretation results	The reviewer comments on the interpretation of study results
Strengths and weaknesses (general)	The reviewer comments on the general strengths and weaknesses of the study
Strengths and weaknesses (methods)	The reviewer comments on the strengths and weaknesses of the study methods
Statistical methods	The reviewer comments on the appropriateness of the statistical methods
Methodological quality	The reviewer comments on the methodological quality (internal validity) of the study
Applicability and external validity	The reviewer comments on the applicability and external validity of the study results
Presentation and organization	The reviewer comments on the presentation and organization of the manuscript
Adherence to RG	The reviewer comments on the adherence of the manuscript to the reporting guidelines
Structure of reviewer's comms.	The reviewer´s comments are structured and organized
Clarity	The reviewer´s comments are clear and easy to read
Constructiveness	The reviewer´s comments are constructive
Detail/Thoroughness	The reviewer´s comments are detailed and thorough
Objectivity	The reviewer´s comments are objective
Fairness	The reviewer's comments are fair
Support by evidence	The reviewer's comments are evidence based
Knowledgeability	The reviewer knows and understands correctly the content of the manuscript
Tone	The reviewer uses a courteous tone
Timeliness	The reviewer completes the peer review report on time

Table 2. Survey participants' characteristics

Characteristics	Editors N=198	Authors N=248	Total N=446
Gender	**N=169**	**N=230**	**N=399**
Woman	46 (27.2%)	83 (36.1%)	129 (32.3%)
Man	121 (71.6%)	142 (61.7%)	263 (65.9%)
Other	2 (1.2%)	5 (2.2%)	7 (1.8%)
Age	**N=156**	**N=220**	**N=376**
<40	32 (20.5%)	71 (32.3%)	103 (27.4%)
41-50	29 (18.6%)	59 (26.8%)	88 (23.4%)
51-60	52 (33.3%)	37 (16.8%)	89 (23.7%)
>60	43 (27.6%)	53 (24.1%)	96 (25.5%)
Education	**N=169**	**N=230**	**N=399**
Bachelor Degree	4 (2.4%)	3 (1.3%)	7 (1.7%)
Master Degree	11 (6.5%)	20 (8.7%)	31 (7.8%)
PhD	107 (63.3%)	118 (51.3%)	225 (56.4%)
M.D. or equivalent	34 (20.1%)	76 (33.0%)	110 (27.6%)
Prefer not to answer	2 (1.2%)	1 (0.4%)	3 (0.7%)
Other	11 (6.5%)	12 (5.2%)	23 (5.8%)
Location journal/institution	**N=165**	**N=224**	**N=389**
Europe	132 (80.0%)	87 (38.8%)	219 (56.3%)
North America	23 (14.0%)	95 (42.4%)	118 (30.3%)
South America	2 (1.2%)	5 (2.2%)	7 (1.8%)
Africa	1 (0.6%)	1 (0.4%)	2 (0.5%)
Asia	3 (1.8%)	11 (5.0%)	14 (3.6%)
Australia	4 (2.4%)	25 (11.2%)	29 (7.5%)
Number of years of experience	**N=165**	**N=224**	**N=389**
<5 years	74 (44.8%)	36 (16.1%)	110 (28.3%)
6-10 years	46 (27.9%)	51 (22.7%)	97 (24.9%)
11-15 years	27 (16.4%)	34 (15.2%)	61 (15.7%)
16-20 years	7 (4.2%)	19 (8.5%)	26 (6.7%)
>20 years	11 (6.7%)	84 (37.5%)	95 (24.4%)

Table 3. Survey participants' comments on the importance and/or wording of the 20 items to assess peer review report quality

Themes	Definition	Codes	Examples
Dependencies	Theme including codes on how the importance of an item depends on different factors (e.g., type of study, paper quality, type of journal, etc.)	Dependency on the type of study (n=34)	*Depends on type of study. For systematic reviews of course fundamental. For other studies this will be more and more important for easier comparisons between studies and for quality improvement. It makes our work easier if the authors also compliance also improve*
		Dependency on the paper quality (n=20)	*This depends on the quality of the manuscript. Sometimes the quality is so low that a reviewer can highlight one or two major methodological flaws, which are sufficient to reject.*
		Dependency on the type of journal (n=19)	*This depends on the journal's criteria*
		Dependency on the author's claim and impact of the study (n=7)	*This depends on the claims made*
Importance	Theme including codes on the importance (or not) of an item.	Importance of the item (n=43)	*This is absolutely key to the interpretation of the study. Unfortunately most reviewers, in my field, do not fully understand current (and correct) methods.*
		Importance of replication and conformation study (n=18)	*Not always important to be original study as some are trying to duplicate findings from previous studies.*
		Importance of perceptions, opinions and experience (n=14)	*But some comments will inevitably be opinion, regarding emphasis, values, writing style*
		Importance of a high quality review rather than on time review (n=13)	*Better to have a late high quality report than a moderate quality report on time.*
Improvements	Theme including codes on how an item is useful for both authors and editors in the peer review process.	Useful for authors and editors (n=21)	*It's important to make it easy for the editor and authors to understand the review, and for authors to respond.*
		Improving the manuscript (n=9)	*Important when it will help improve the quality of the*

			communication. Not necessary when it flows well.
		Avoiding exaggeration and misinterpretation (n=8)	*This is an area where the reviewer may have a valuable role in tempering an author's enthusiasm, hubris or bias.*
Item	Theme including codes on the characteristics of an item.	Related to other item (n=43)	*Yes, but it is confusing to separate this from the general strength and weaknesses. The question should be if the reviewer thinks that the message can (potentially) answer the research question.*
		Subjective item (n=22)	*Too subjective! What is relevant to one person of field could be totally not-relevant to another*
		Requirement (n=9)	*But it's an ethical requirement, and helps improve everyone's experience.*
Reviewer	Theme including codes on the expertise and characteristics of a peer reviewer.	Reviewer's expertise (n=148)	*Some reviewers know about methods and some about content. It would be ideal to always have both, but that is often not the case.*
		Impossibility to be totally objective (n=35)	*100% objectivity doesn't exist*
		Reviewer as an extra unpaid job (n=10)	*For the most part, reviews are done on a voluntary basis*
Responsibility	Theme including codes on the editor and/or author´s responsibility to assess an item.	Editor's responsibility (n=48)	*In my experience this is usually picked up by the Editors and Associate Editors rather than the reviewers.*
		Joint responsibility (n=24)	*I think this is the role of the editors as well as the reviewers.*
		Author's responsibility (n=6)	*Authors should already be doing this*
Structure and content	Theme including codes on the structure and content of a peer review report.	Straight to the critical points (n=14)	*Sometimes a succinct review is still helpful, if it cuts straight to the critical points. For example, if it is clear that a manuscript has major flaws, then a review*

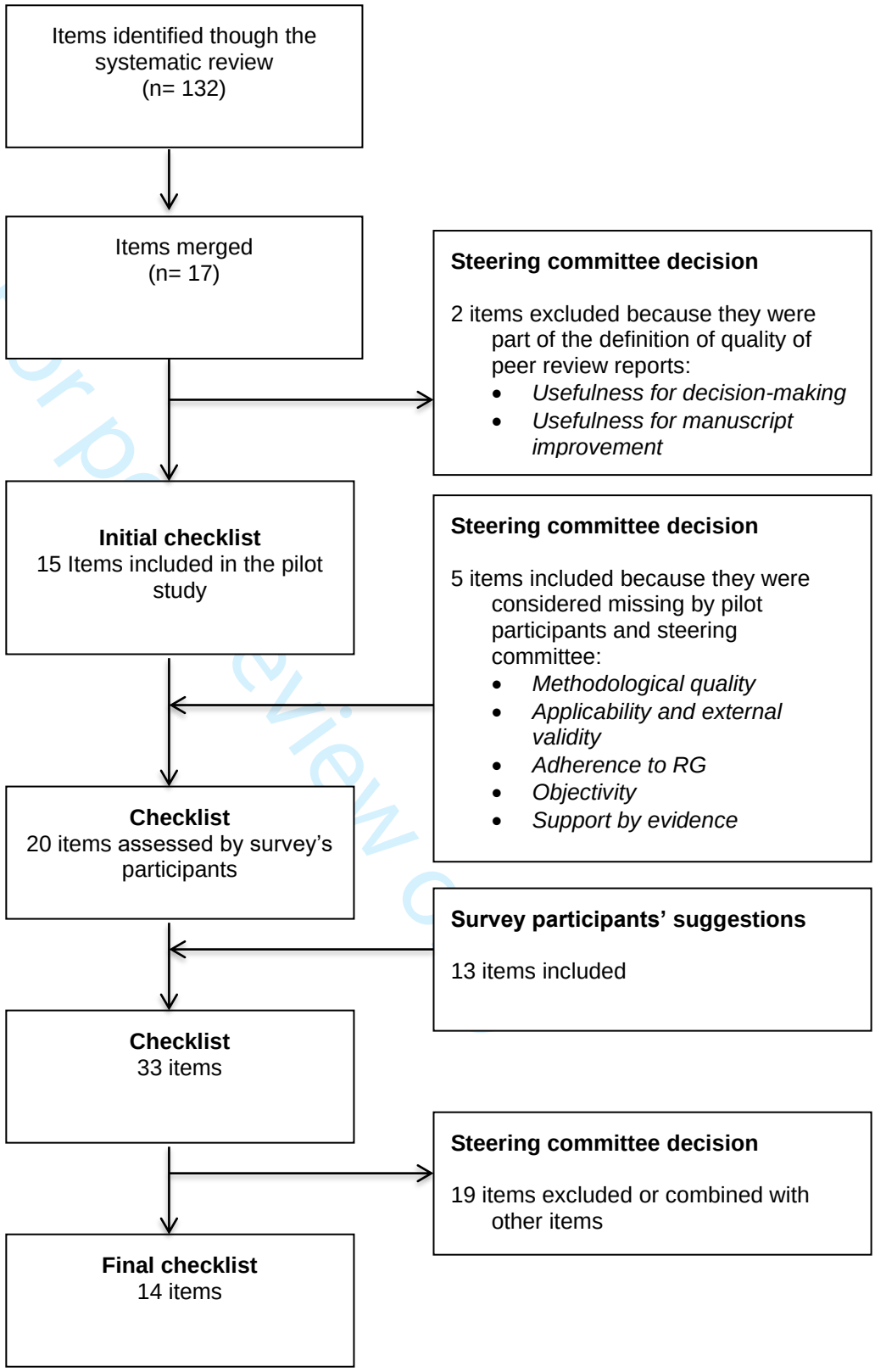
			that points out those flaws clearly and dispassionately would be very helpful. It would not necessarily need to delve into the finer details.
		Unnecessary to provide evidence to each comment (n=10)	*I don't think reviewers need to cite something for every point that they make.*
		Declaration of COI (n=8)	*Peer reviewers should disclose COI.*
		Standard structure of a review (n=7)	*I would suggest providing a template to reviewers.*
		Not necessary for all reviews (n=6)	*Reviews come in all lengths and vary in detail. It is helpful to have some reviewers provide detailed information but not necessary that all do so.*
Wording	Theme including codes on how to improve the wording of an item.	Wording of the item (n=110)	*Rather than "The reviewer's comments are evidence-based" I would suggest that the category should be: "The reviewer distinguishes between comments that are supported by evidence (and provides suitable citations) and those based on opinion or experience"*

Table 4. The ARCADIA tool

In the peer review report, did the reviewer comment on…		
Importance of the study	the contribution of the study to scientific knowledge?	☐ YES ☐ NO ☐ NA
	whether the relevant literature was accurately reviewed?	☐ YES ☐ NO ☐ NA
Robustness of the study methods	the soundness of the study methods (e.g., study design, outcomes, risk of bias)?	☐ YES ☐ NO ☐ NA
	the suitability of the statistical methods?	☐ YES ☐ NO ☐ NA
Interpretation and discussion of the study results	whether the study conclusions answer the research question(s) and correctly summarize the study results?	☐ YES ☐ NO ☐ NA
	whether the study limitations are acknowledged?	☐ YES ☐ NO ☐ NA
	the applicability and generalizability (external validity) of the study results?	☐ YES ☐ NO ☐ NA
Reporting and transparency of the manuscript	whether any major deviations from the study protocol are reported?	☐ YES ☐ NO ☐ NA
	whether the completeness of the reporting allows study reproducibility, by verifying the adherence of the manuscript to the corresponding RG?	☐ YES ☐ NO ☐ NA
	the presentation (e.g., quality of the written language, tables, figures, etc.) and organization of the manuscript?	☐ YES ☐ NO ☐ NA
	the availability of study data and material?	☐ YES ☐ NO ☐ NA
Were the peer reviewer´s comments…		
Characteristics of peer reviewer's comments	clear?	☐ YES ☐ NO
	constructive?	☐ YES ☐ NO
	objective and, if opportune, supported by evidence?	☐ YES ☐ NO

NA=Not applicable

Figure 1. Flowchart of items to include in a checklist to assess peer review report quality



Supplementary file 1. Survey questionnaires

1.1. Survey questionnaire for authors

Welcome to the survey!

Although the peer review process plays a key role in research dissemination, only limited research has been conducted so far in this field.

The objective of this survey is to investigate the perspectives of biomedical editors and authors towards the **quality of peer review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer review report in biomedical research.

Knowing your expertise, we would be very grateful if you could answer a few questions and share your opinion. The survey will take approximately **10 minutes** to complete. Your participation in this study is completely voluntary. If you decide to participate, all your answers will be de-identified and stored in a secured repository at Universitat Politècnica de Catalunya, Barcelona-Tech (Spain). The de-identified data from this study will be shared on Zenodo repository. In case you opt out of sharing your data, you will still be able to participate in the study.

This survey has received ethics approval from the Research Ethics Committee of the Universitat Politècnica de Catalunya, Barcelona-Tech (Spain).

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by Marie Skłodowska-Curie Action http://miror-ejd.eu/. The objective of MiRoR project is to train future generations of scientists in Research on Research, a new discipline aiming to promote research integrity increasing research value and reducing waste in health research.

This study is conducted by **Cecilia Superchi**, a PhD student at Universitat Politècnica de Catalunya, Barcelona-Tech and Université Paris Descartes, Sorbonne Paris Cité in collaboration with Prof. Darko Hren (University of Split), Prof. José Antonio Gonzalez (Universitat Politècnica de Catalunya) and Prof. Isabelle Boutron (Université Paris Descartes).

If you have any questions about this study or your rights as a participant, you may contact by email Cecilia Superchi, cecilia.superchi@upc.edu or Darko Hren, dhren@ffst.hr

Do you agree to take part in the study?

- ◯ Yes, I agree
- ◯ No, I do not agree

Do you agree to share your de-identified data?

- ◯ Yes, I agree
- ◯ No, I do not agree

Definition of peer review report quality

The **quality of a peer review report** could be defined as "to what extent the peer review report helps editors to make a fair decision and authors to improve the quality of the submitted manuscript"

Do you agree with this definition?

- ◯ Yes
- ◯ No
- ◯ Partially

Please add your comments and ideas on how to improve the definition

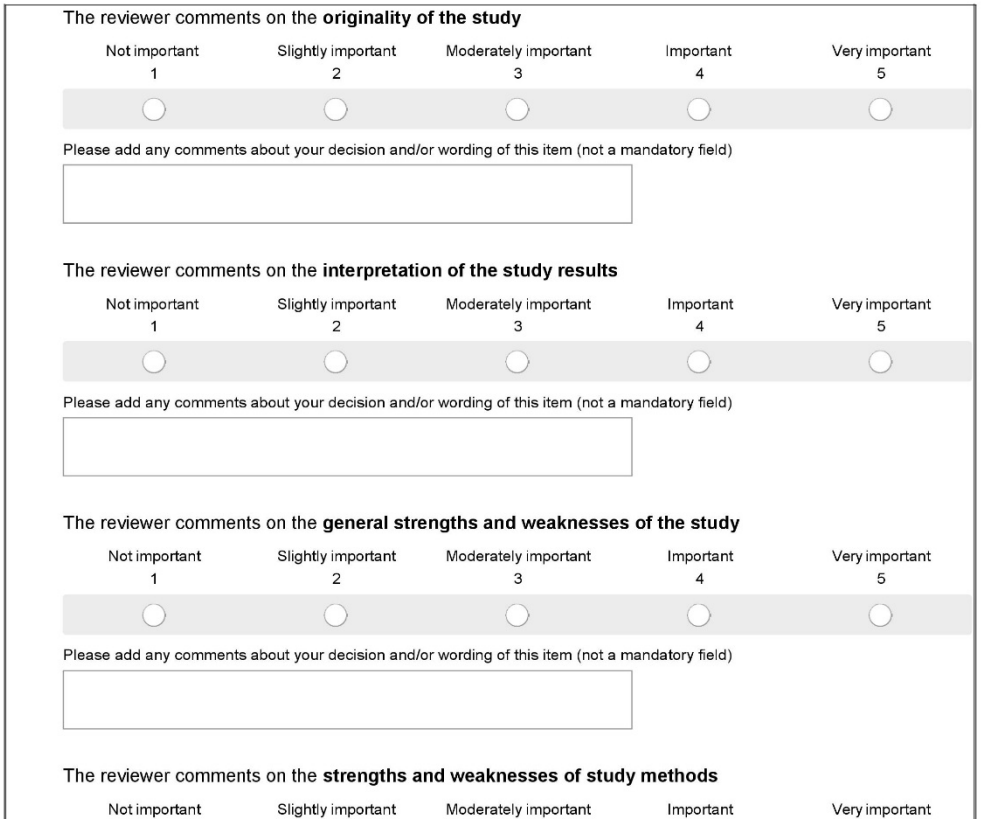
Importance of the items to assess peer review report quality

The following items have been identified in a systematic review as possible quality components of a peer review report.

We are interested to know your opinion on the importance of these items, particularly whether the item should be included in a new tool assessing the quality of a peer review report.

Please rate the IMPORTANCE of each item in assessing the quality of a peer review report from 1 (not important) to 5 (very important).

We expect that for some items it will not be easy for you to make a clear decision about the importance of the item. In those cases we still invite you to offer your rating but you can elaborate on your decision. Furthermore we invite you to suggest potential improvements in wording of the items.



The reviewer comments on the **originality of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **interpretation of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **general strengths and weaknesses of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **strengths and weaknesses of study methods**

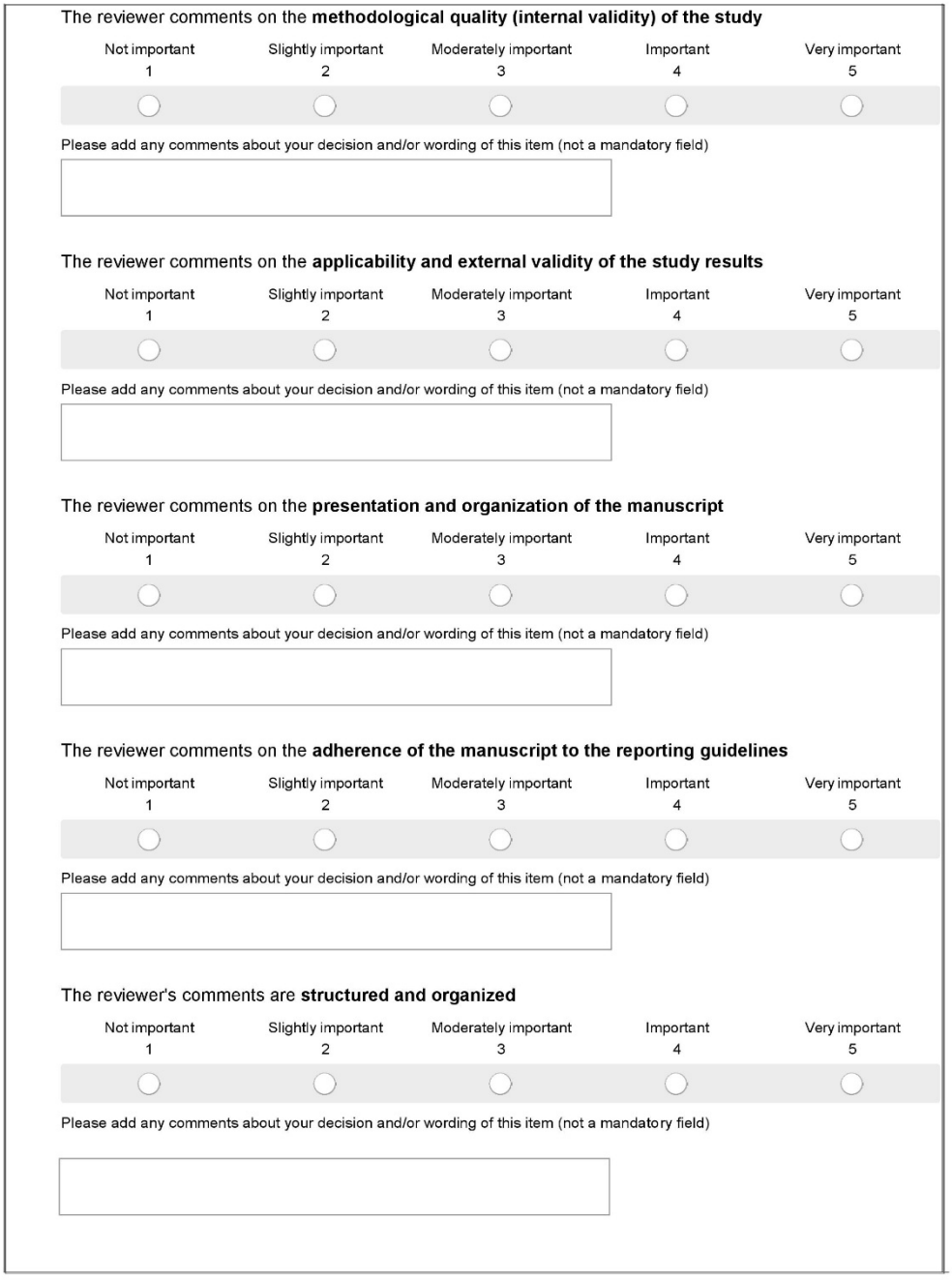
Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **appropriateness of the statistical methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)



The reviewer comments on the **methodological quality (internal validity) of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **applicability and external validity of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **presentation and organization of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **adherence of the manuscript to the reporting guidelines**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **structured and organized**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **clear and easy to read**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer knows and understands correctly **the content of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **constructive**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **detailed and thorough**

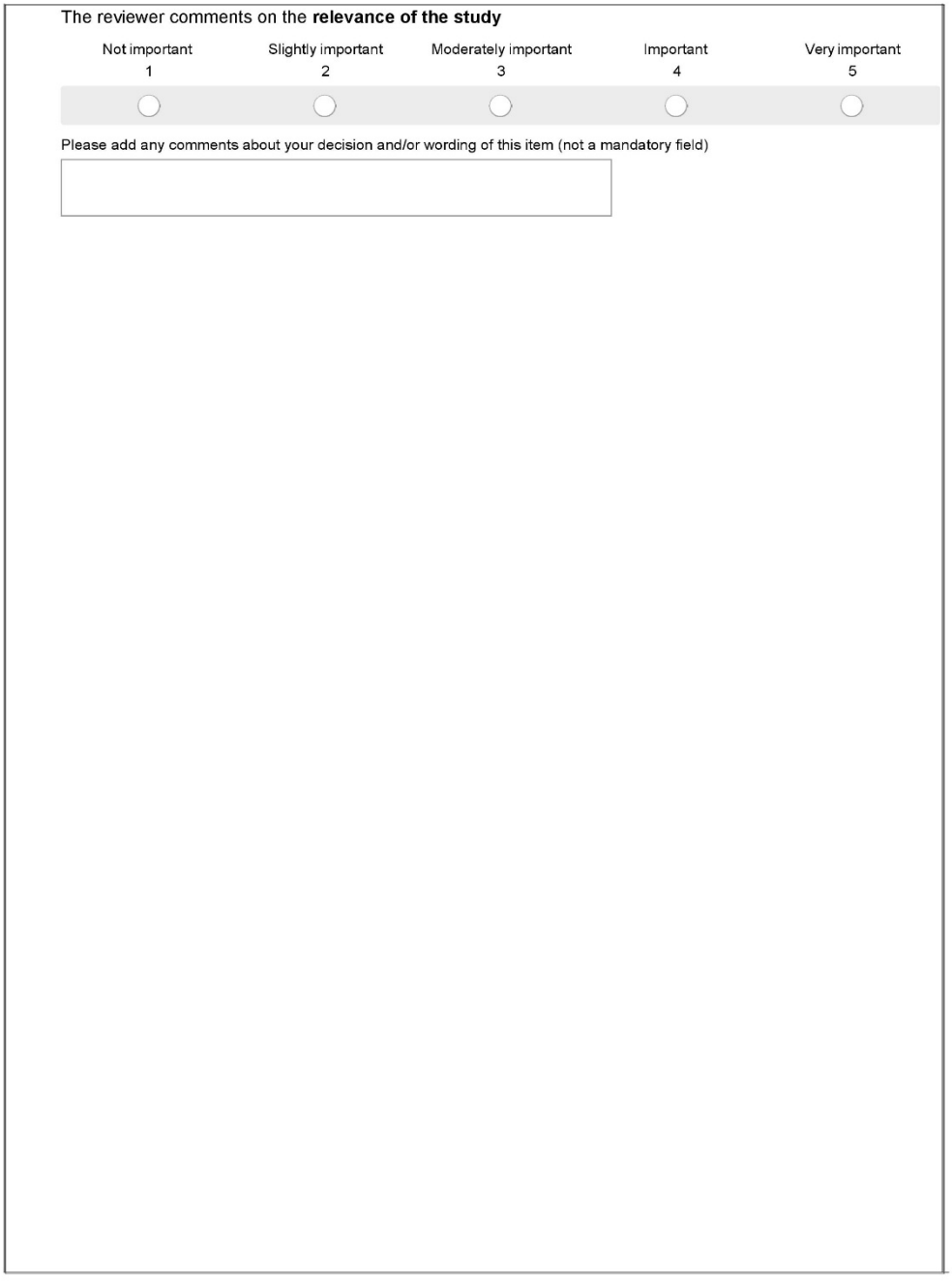
Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer uses a **courteous tone**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)


The reviewer comments on the **relevance of the study**
Not important
1
Slightly important
2
Moderately important
3
Important
4
Very important
5
Please add any comments about your decision and/or wording of this item (not a mandatory field)

New items to assess peer review report quality

Are there any other items to assess the quality of a peer review report that you think should be included?

Please list them.

Demographic characteristics

What is your gender?

- ○ Woman
- ○ Man
- ○ Prefer not to answer
- ○ Other (please specify)

What is your age?

What is the highest level of education obtained?

- ○ Bachelor Degree
- ○ Master Degree
- ○ PhD
- ○ MD or equivalent
- ○ Prefer not to answer
- ○ Other (please specify)

Author´s characteristics

What is your job title at your institution?

- ○ Researcher
- ○ Assistant Professor
- ○ Associate Professor
- ○ Professor
- ○ Other (please specify)

What type of institution are you affiliated at?

- ○ Private University
- ○ Public University
- ○ Research Centre
- ○ Other (please specify)

Where is the institution located?

- ○ Europe
- ○ North America
- ○ South America
- ○ Africa
- ○ Asia
- ○ Australia

MiRoR Methods in Research on Research

Author´s characteristics

How long have you been publishing scientific papers?

- ◯ <5 years
- ◯ 6-10 years
- ◯ 11-15 years
- ◯ 16-20 years
- ◯ >20 years

Do you also work as biomedical editor?

- ◯ Yes
- ◯ No

Are you involved in making decisions on the manuscripts received by your journal?

- ◯ Yes
- ◯ No

MiRoR
Methods in
Research on Research

Study results and next step

Please check which of the following options you would be interested in

- ◯ I would be interested in receiving the results of the present study
- ◯ I would be interested in participating in the validation study of a new tool for assessing the quality of a peer-review report

Please write down your name and email address. Your data will be **EXCLUSIVELY** used for the option(s) which you have previously chosen.

Name

Email address

1.2. Survey questionnaire for biomedical editors

Welcome to the survey!

Although the peer review process plays a key role in research dissemination, only limited research has been conducted so far in this field.

The objective of this survey is to investigate the perspectives of biomedical editors and authors towards the **quality of peer review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer review report in biomedical research.

Knowing your expertise, we would be very grateful if you could answer a few questions and share your opinion. The survey will take approximately **10 minutes** to complete. Your participation in this study is completely voluntary. If you decide to participate, all your answers will be de-identified and stored in a secured repository at Universitat Politècnica de Catalunya, Barcelona-Tech (Spain). The de-identified data from this study will be shared on Zenodo repository. In case you opt out of sharing your data, you will still be able to participate in the study.

This survey has received ethics approval from the Research Ethics Committee of the Universitat Politècnica de Catalunya, Barcelona-Tech (Spain).

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by Marie Skłodowska-Curie Action http://miror-ejd.eu/. The objective of MiRoR project is to train future generations of scientists in Research on Research, a new discipline aiming to promote research integrity increasing research value and reducing waste in health research.

This study is conducted by **Cecilia Superchi**, a PhD student at Universitat Politècnica de Catalunya, Barcelona-Tech and Université Paris Descartes, Sorbonne Paris Cité in collaboration with Prof. Darko Hren (University of Split), Prof. José Antonio Gonzalez (Universitat Politècnica de Catalunya) and Prof. Isabelle Boutron (Université Paris Descartes).

If you have any questions about this study or your rights as a participant, you may contact by email Cecilia Superchi, cecilia.superchi@upc.edu or Darko Hren, dhren@ffst.hr

Aren't you a biomedical editor? Please take part in the survey for biomedical authors following this link:
https://www.surveymonkey.com/r/REPORT_QUALITY_AUTHORS

Do you agree to take part in the study?	Do you agree to share your de-identified data?
◯ Yes, I agree	◯ Yes, I agree
◯ No, I do not agree	◯ No, I do not agree

MiRoR Methods in Research on Research

Definition of peer review report quality

The **quality of a peer review report** could be defined as "to what extent the peer review report helps editors to make a fair decision and authors to improve the quality of the submitted manuscript"

Do you agree with this definition?

- ◯ Yes
- ◯ No
- ◯ Partially

Please add your comments and ideas on how to improve the definition

Importance of the items to assess peer review report quality

The following items have been identified in a systematic review as possible quality components of a peer review report.

We are interested to know your opinion on the importance of these items, particularly whether the item should be included in a new tool assessing the quality of a peer review report.

Please rate the IMPORTANCE of each item in assessing the quality of a peer review report from 1 (not important) to 5 (very important).

We expect that for some items it will not be easy for you to make a clear decision about the importance of the item. In those cases we still invite you to offer your rating but you can elaborate on your decision. Furthermore we invite you to suggest potential improvements in wording of the items.

The reviewer comments on the **originality of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **interpretation of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **general strengths and weaknesses of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **strengths and weaknesses of study methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **appropriateness of the statistical methods**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **methodological quality (internal validity) of the study**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **applicability and external validity of the study results**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **presentation and organization of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer comments on the **adherence of the manuscript to the reporting guidelines**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **structured and organized**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **clear and easy to read**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer knows and understands correctly **the content of the manuscript**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **constructive**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer's comments are **detailed and thorough**

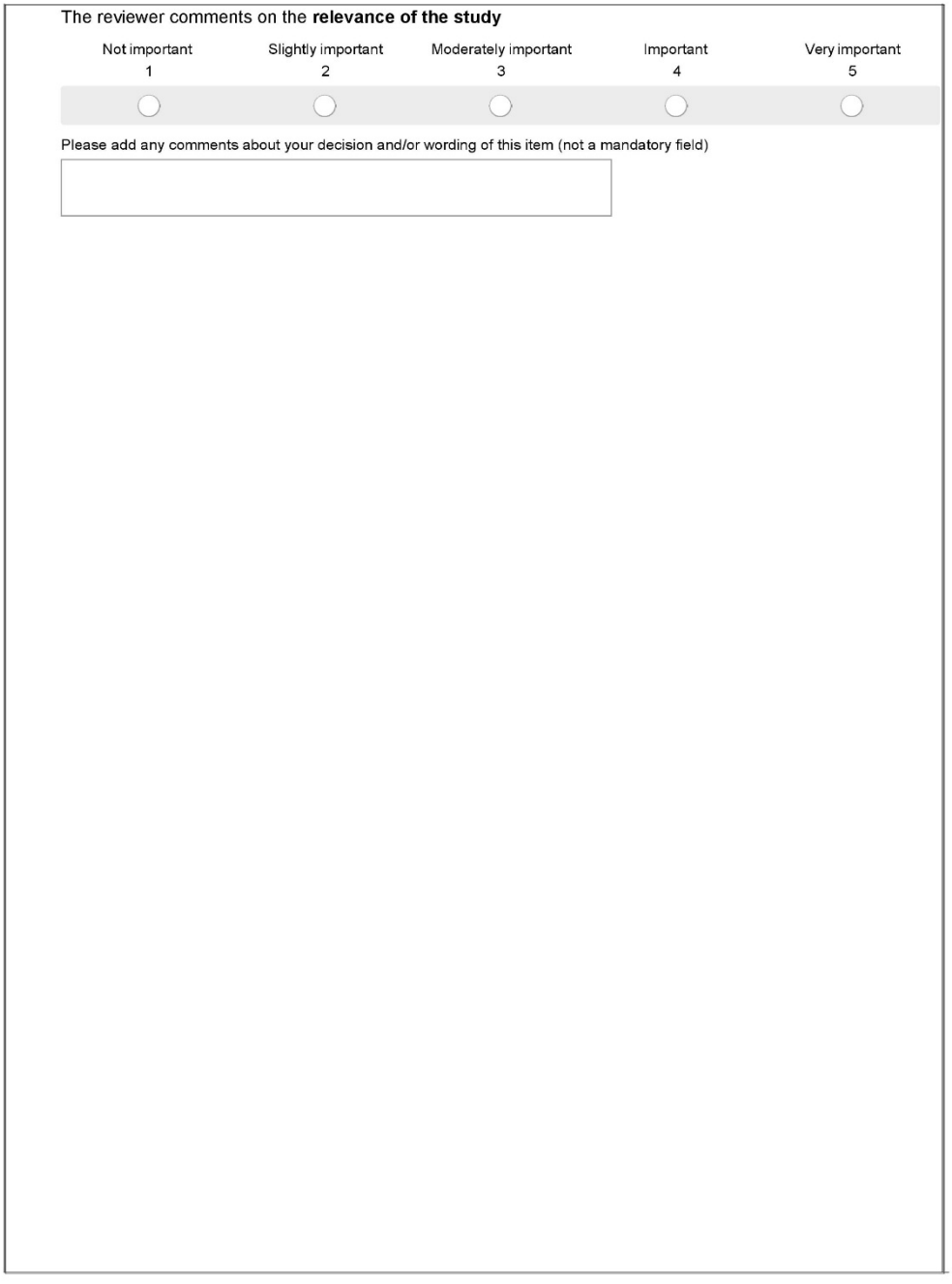
Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)

The reviewer uses a **courteous tone**

Not important	Slightly important	Moderately important	Important	Very important
1	2	3	4	5
○	○	○	○	○

Please add any comments about your decision and/or wording of this item (not a mandatory field)


The reviewer comments on the **relevance of the study**
Not important
1
Slightly important
2
Moderately important
3
Important
4
Very important
5
Please add any comments about your decision and/or wording of this item (not a mandatory field)

New items to assess peer review report quality

Are there any other items to assess the quality of a peer review report that you think should be included?

Please list them.

Demographic characteristics

What is your gender?

- ○ Woman
- ○ Man
- ○ Prefer not to answer
- ○ Other (please specify)

What is your age?

What is the highest level of education obtained?

- ○ Bachelor Degree
- ○ Master Degree
- ○ PhD
- ○ MD or equivalent
- ○ Prefer not to answer
- ○ Other (please specify)

Editor´s characteristics

What is your job title at your journal?

- ○ Editor in chief
- ○ Associate editor
- ○ Academic editor
- ○ Section editor
- ○ Deputy editor
- ○ Other (please specify)

Are you involved in making decisions on the manuscripts received by your journal?

- ○ Yes
- ○ No

At what type of journal do you currently working as editor?

- ○ General Journal
- ○ Specialty Journal

Editor´s characteristics

Where is the journal located?

- ○ Europe
- ○ North America
- ○ South America
- ○ Africa
- ○ Asia
- ○ Australia

How long have you been working as editor?

- ○ <5 years
- ○ 6-10 years
- ○ 11-15 years
- ○ 16-20 years
- ○ >20 years

Does your work inside or outside the journal include authoring scientific papers?

- ○ Yes
- ○ No

Study results and next step

Please check which of the following options you would be interested in

- ◯ I would be interested in receiving the results of the present study
- ◯ I would be interested in participating in the validation study of a new tool for assessing the quality of a peer-review report

Please write down your name and email address. Your data will be **EXCLUSIVELY** used for the option(s) which you have previously chosen.

Name

Email address

Supplementary file 2. Invitation email for corresponding authors and biomedical editors

From:
Cc:
To:
Subject: Academic Survey on Peer Review

Dear researcher,

As corresponding author of the article recently published in [CUSTOM 1], we would like to invite you to participate in an **academic survey**.

The objective of this survey is to investigate the perspectives of biomedical editors and authors on the **quality of peer-review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer-review report in biomedical research.

The survey will take approximately **10 minutes to complete**. Participation in this study is completely **voluntary** and you may withdraw at any time.

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No 676207 http://miror-ejd.eu/

We would be very grateful if you would take the time to complete our survey. **Your insights** as an author are **essential** to us.

If you have any questions, comments or queries please do not hesitate to contact us at cecilia.superchi@upc.edu or dhren@ffst.hr

We kindly thank you for your time, attention, and cooperation.

Sincerely,

Cecilia Superchi, PhD Student at Universitat Politècnica de Catalunya & Université Paris Descartes
Darko Hren, PhD, Prof. at University of Split
José Antonio Gonzalez, PhD, Prof. at Universitat Politècnica de Catalunya
Isabelle Boutron, MD, PhD, Prof. at Université Paris Descartes

From:
Cc:
To:
Subject: Academic Survey on Peer Review

Dear [Name] [Surname],

As [CUSTOM 1] at [CUSTOM 2], we would like to invite you to participate in an **academic survey on peer review.**

The objective of this survey is to investigate the perspectives of biomedical editors and authors on the **quality of peer-review reports**. We hope this work will help us to develop a new tool to assess the quality of a peer-review report in biomedical research.

The survey will take approximately **10 minutes to complete**. Participation in this study is completely **voluntary** and you may withdraw at any time.

You are also encouraged to **forward the link** of the survey to your colleagues who may be interested in participating in this study https://www.surveymonkey.com/r/REPORT_QUALITY_EDITORS

This study is part of the **Methods in Research on Research** (MiRoR) project, a joint doctoral training programme in the field of clinical research funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No 676207 http://miror-ejd.eu/

We would be very grateful if you would take the time to complete our survey. **Your insights** as a biomedical editor are **essential** to us.

If you have any questions, comments or queries, please do not hesitate to contact us at cecilia.superchi@upc.edu or dhren@ffst.hr

We thank you kindly for your time, attention, and cooperation.

Sincerely,

Cecilia Superchi, PhD Student at Universitat Politècnica de Catalunya & Université Paris Descartes
Darko Hren, PhD, Prof. at University of Split
José Antonio Gonzalez, PhD, Prof. at Universitat Politècnica de Catalunya
Isabelle Boutron, MD, PhD, Prof. at Université Paris Descartes

Supplementary file 3. Top 30-biomedical journals with the highest impact factors

Full Journal Title**	IF
New England Journal Of Medicine*	79.3
Lancet*	53.3
JAMA-Journal Of The American Medical Association	47.7
BMJ-British Medical Journal*	23.3
JAMA Internal Medicine	20.0
Annals Of Internal Medicine	19.4
Nature Reviews Disease Primers	16.1
Journal Of Cachexia Sarcopenia And Muscle	12.5
Plos Medicine	11.7
Bmc Medicine*	9.1
Mayo Clinic Proceedings*	7.2
Cochrane Database Of Systematic Reviews	6.8
Journal Of Internal Medicine	6.8
Canadian Medical Association Journal*	6.2
Journal Of Clinical Medicine*	5.6
American Journal Of Medicine*	5.1
Translational Research*	4.9
Annals Of Family Medicine*	4.5
Medical Journal Of Australia*	4.2
American Journal Of Preventive Medicine*	4.1
Amyloid-Journal Of Protein Folding Disorders	4.0
Journal Of General Internal Medicine*	4.0
Deutsches Arzteblatt International	3.9
Palliative Medicine	3.8
Preventive Medicine*	3.5
British Medical Bulletin	3.4
European Journal Of Internal Medicine*	3.3
British Journal Of General Practice*	3.3
Journal Of Pain And Symptom Management*	3.2
Qjm-An International Journal Of Medicine	3.2

* Journal reporting the corresponding author in the PubMed abstract.
**Source: InCites Journal Citation Reports 2017 under the category “Medicine, general and internal”.

Supplementary file 4. Complete participants characteristics

Characteristics	Editors N=165
Journal Role	
Editor-in-Chief	50 (30.3%)
Associate Editor	63 (38.2%)
Academic Editor	7 (4.2%)
Section Editor	6 (3.6%)
Deputy Editor	12 (7.3%)
Other (e.g. Statistical Editor, Patient Editor)	27 (16.4%)
Involvement in making decisions on the manuscript	
Yes	144 (87.3%)
No	21 (12.7%)
Type of Journal	
General Journal	39 (23.6%)
Specialty Journal	126 (76.4%)
Authorship of scientific papers	
Yes	141 (85.5%)
No	24 (14.5%)

Characteristics	Authors N=224
Occupation	
Professor	63 (28.1%)
Associate Professor	31 (13.8%)
Assistant Professor	34 (15.2%)
Researcher	47 (21.0%)
Other (e.g. Lecturer, Postdoc, PhD)	49 (21.9%)
Type of Institution	
Public University	134 (59.8%)
Private University	33 (14.7%)
Research Centre	17 (7.6%)
Other (e.g. Hospital)	40 (17.9%)
Employment as biomedical editor	
Yes	63 (28.1%)
No	161 (71.9%)
Involvement in making decisions on the manuscript	
Yes	56 (88.9%)
No	7 (11.1%)

Supplementary file 5. Codebooks

The first codebook is about the suggestions made by survey participants on how to improve the definition of peer review report quality we provided.

The second codebook is about the comments made by survey participants on importance and/or wording of each item (n=20).

The third codebook is about the identification of new items to assess peer review report quality by survey participants.

Codebook 1. Suggestions on how to improve the definition of peer review report quality

Suggestions on how to improve the definition of peer review report quality (n= 87)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Assessment of different aspects of a study	Statements on the different aspects of a study that should be discussed in a peer review report	Accuracy of the study	NA	*Also helps ensure the accuracy of the content (at least in part)*
		Originality of the study	NA	*Should include recognition of what has already been done well where possible*
		Relevance of the study	NA	*Consider adding "the quality and value of the submitted manuscript." Or perhaps "quality and impact." The point is that some submitted work is high quality, but still not useful because it is just repeating prior work or answering an irrelevant question*
		Reproducibility of the study	NA	*The definition could broadly also take into account issues of reproducibility*
		Research integrity of the study	NA	*The definition should also include something about identifying plagiarism and conflict of interests*
		Robustness of the study	NA	*Not necessarily to improve the quality as that may be a secondary outcome. The peer review is to evaluate the scientific robustness of the research*

		Soundness of the study	NA	*Also - the extent to which the report critically assesses the soundness*
		Strengths and weaknesses of the study	NA	*I would include the extent to which the review accurately and clearly identifies weaknesses / limitations of the study ... but I don't disagree with your definition*
		Structure of the manuscript	NA	*A high quality peer review report may also identify the potential for a poorly structured paper to be revised into an impactful form*
		Study methodology and statistics of the study	NA	*Aspects such as rigorous statistical analysis and sampling/experimental design, degree of innovation, and the statement and testing of clear scientific hypothesis, should be addressed in a peer-review. Also, the technical issues and methodologies should be targeted during the peer-review process*
		Validity/trustworthiness of the study	NA	*What about the validity and trustworthiness of the findings?*
Consideration of journal´s policy	Statements on the consideration of the journal´s policy in writing a peer review report	NA	NA	*A good quality peer-review report takes into account journal policies and publication criteria while helping authors provide the best version of their work*
Irrelevant and ambiguous comments	Irrelevant and ambiguous comments for improving the provided definition of peer review report quality	NA	NA	*There is usually more than one report, so reportS*
Quality as a vague concept	Statements on the difficulty to define “quality”	Dependency on the type of journal and study	NA	*I would add at the end.... "based on a rubric specific to the type of article submitted"*
		Quality of research and quality of reporting	NA	*"Quality" is ambiguous. Relevant aspects of quality could include scientific validity (the extent to which the methods are adequate, the conclusions supported etc.) AND/OR reproducibility (the extent to which the study is described in sufficient detail that it could be reproduced). The former is a quality of the scientific study and the*

				latter is a quality of the text
Reviewer´s expertise	Statements on how the quality of a peer review report is related to the level of expertise of a reviewer	NA	NA	*The quality of the peer review also depends on how well the reviewer has understood the manuscript and the reviewer's level of expertise in the topic*
Reviewer's comments characteristics	Statements on the different characteristics of a peer review report	Clarity of the comments	NA	*I think the definition should include whether the reviewers have expressed themselves clearly and unambiguously*
		Constructiveness of the comments	NA	*Perhaps this is implied in the proposed definition, but you could mention that a high-quality peer review includes constructive criticism -- that is, not just an identification of flaws but suggestions for remedies*
		Fairness and impartiality of the comments	NA	*Add: 'is unbiased and competently-conducted'*
		Understanding correctly the content of the manuscript	NA	*The quality of the peer review also depends on how well the reviewer has understood the manuscript*
Role of external parties	Statements on the role of external parties in assessing the quality of a peer review report	NA	NA	*External parties should also play a role*
Scope of the peer review process	Statements on the scope of the peer review process	Different facets	NA	*Peer review has many facets*
		Ensuring accessibility to the readers	NA	*[..] and the accessibility to readers*
		Ensuring quality of science	NA	*I would like to insist on the role of peer-review to ensure the quality of the science presented in the manuscript*

		Evaluating rather than improving	NA	*The peer review is to evaluate the scientific robustness of the research*
		Independence	NA	*This definition does not capture the necessary independence of the peer review process*
		Related to decision making	NA	*This definition implies that editors' decisions can always override peer reviewers' appraisals. But an editor's appraisal of a paper should have equal weight to a peer reviewers appraisal*
		Transparency and critical appraisal	NA	*The focus of peer-review is transparency and critical appraisal. Peer review scope is broader than editorial decisions. Editorial decisions are a specific use case of peer review*
		Validation of the research	NA	*Peer review also helps to validate the research before publication, so the report also needs to do this to be effective*
		Variable process		*Peer review can be very variable; at its best it really improves the quality of papers. At its worst it is bullying and partial*
Timeliness of peer review process	Statements on the consideration of timeliness in defining the quality of a peer review report	NA	NA	*Timeliness should be included, less than 2 weeks is ideal*
Usefulness of the peer review report	Statements on the usefulness of a peer review report for authors and editors	Useful for authors	Improving manuscript quality	*It should be aimed at helping the authors improve the quality of their work.*
			More effective communication of research	*Peer review ideally contributes to effective communication through research publication, by exposing the author's work to the potential audience(s) for it and thus showing where readers stumble or identify limitations that need to be recognized. I think it would improve the definition if you could work in something about effective communication, which can be distinct from quality per se. More effective, clearer communication promotes more learning from the article and a stronger link to implementation. Also, quality should*

				probably be judged in terms of the purpose of the study; a delivery system study for example must provide more information on context in order to be useful than does a classical randomized trial such as of a pharmaceutical.
		Useful for editors	Filtering studies	*Also, 1) a check on poor research,*
			Decision to enhance the readership and citations	*helps editors to make a fair and informed decision that will enhance the readership and citations of the journal*
			Enabling fair decision	*I agree that they are primarily good to the extent that they help editors. "Fairness' is important, but enabling 'informed' and 'rigorous' decisions matter too.*
			Leading to incorrect decision	*Implies that the reviewer is making sensible suggestions, which may not be the case (and which an editor may not pick up on). E.g., a non-statistical reviewer commenting (incorrectly) on statistical methods and the editor is unaware if the comments are relevant/correct. The review could help the editor make a decision but it could be an incorrect decision.*
			Same weight for editors and reviewers appraisal	*This definition implies that editors' decisions can always override peer reviewers' appraisals. But an editor's appraisal of a paper should have equal weight to a peer reviewer's appraisal. This top down system can allow for bias.*
		Useful for both editors and authors	NA	*The point you suggest to help BOTH editor AND authors is a key element. When I ask for revision I provide new insights or suggestions to improve the quality and accuracy of a paper.*
Wording of the definition	Statements on how to improve the definition of the peer review report quality	Disagreement with the use of fair	NA	*This seems reasonable but I would leave out the word "fair" as I would assume that editors always aim to make fair decisions!*

Codebook 2. Comments on the importance and/or wording of each item

Relevance (n = 56)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author's responsibility	Statements on the author's responsibility to discuss the relevance of the study in the manuscript	NA	NA	*Relevance should be discussed by the authors and ultimately is decided by readers. I would expect reviewers to comment on the author's discussion on the relevance, and only exceptionally come with considerations of his own*
Contribution to the knowledge	Statements on the importance of the study as contribution to the scientific knowledge	NA	NA	*Relevance is important in the context of both the contribution to the knowledge base*
Dependency on the type of journal	Statements on how biomedical journals differently evaluate relevance of a study based on their own criteria	NA	NA	*Relevance also depends on the scope of the journal, and that is an editorial decision, opinion of the reviewer is not so important*
Editor's responsibility	Statements on the editor's responsibility to evaluate the relevance of the study	NA	NA	*This is largely an editorial decision*
Influencing editor's decision	Statements on how the relevance of a study can influence an editorial decision	NA	NA	*This is very helpful for the Editor to make a decision on the manuscript*
Readers as final judges	Statements on how readers are the final judges of the relevance of a study	NA	NA	*Relevance should be discussed by the authors and ultimately is decided by readers*

Related to other items	Statements on the link of the item with other items	NA	NA	*I took this to be the same question as the earlier one on applicability*
Reviewer's expertise	Statements on how the assessment of the relevance of a study depends on reviewer´s expertise	NA	NA	*Not important if you are a statistical reviewer, of a clinical article, without knowing the clinical area. BUT otherwise VERY IMPORTANT for clinical reviewers*
Subjective item	Statements on the subjective interpretation of the term "relevance"	External validity	NA	*Another aspect of relevance might relate to external validity or generalizability -- e.g., a lab study that does not have relevance to the real world*
		Future impact	NA	*Relevance may lie in the future, not in the present*
		Novelty	NA	*There are various aspects of "relevance" -- i.e., it might not be a significant contribution to the literature because the findings are not at all novel*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*Relevance to the broader field, or to general society?*

Originality (n= 56)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of journal	Statements on how biomedical journals differently evaluate the originality of a study based on their own criteria and policy	NA	NA	*Depends on the journal policy. More important when the reader is the client, less important when the author is the client*
Editor's responsibility	Statements on the editor's responsibility to evaluate the originality of the study	NA	NA	*Many journals mainly open which have different editorial policies now-a-days do not ask reviewers' to judge the originality. This is losing importance in open access era*
Importance of replication and confirmatory study	Statements on the importance of conducting replication and confirmatory studies	NA	NA	*Not always important to be original study as some are trying to duplicate findings from previous studies*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Slightly important item	NA	*This is only slightly important in that - once a study has been conducted*
		Important item	NA	*I find that important. The twentieth me too study is not relevant for the knowledge field*
Open access vs. subscription journal	Statements on how open access journals and subscription journals assess differently the originality of a study	NA	NA	*For our journal, this is very important, although I think for some others with a pay-for-publication model they place less importance on e.g. novelty*

Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Reviewer as not the best judge	NA	*Not all reviewers will necessarily be familiar with the literature in a particular area and may not be able to comment on originality of the study*
Wording of the item	Statements on how to improve the wording of the item and better define it	Impact rather than originality	NA	*Some less original studies can still be of significant value, so I prefer comments on impact to comments on originality*
		Originality as novelty	NA	*The Editor is usually aware of this, particularly Editors of journals with high impact factors, who are very keen to publish manuscripts reporting original/novel findings*
		Originality as what the study adds	NA	*I would prefer to think of this in terms of whether it really adds to our knowledge*

Interpretation of results (n= 33)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Avoiding exaggeration & misinterpretation and censoring divergent opinions	Statements on the importance of the item to avoid exaggeration and misinterpretation of study´s results	NA	NA	*This is an area where the reviewer may have a valuable role in tempering an author's enthusiasm, hubris or bias*
Conclusions supported by results, S&W and literature	Statements on the importance that study´s conclusions are supported by results, strengths and weaknesses and literature	NA	NA	*Interpretation of the findings should be judged by its coherence with findings and study limitations and strengths, and by its coherence with literature*
Contribution to the knowledge	Statements on the importance of the study as contribution to the scientific knowledge	NA	NA	*A judgement on the new contribution to knowledge*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Very important item	NA	*Interpretation of the results is crucial- it determines the message that is sent out. It is very important that reviewers pay attention to this interpretation*
Useful for readers	Statements on the uselessness of the item for the readers	NA	NA	*These are useless for the reader*
Liberty in the discussion section	Statements on the liberty of the authors to interpret study´s results in the discussion section	NA	NA	*As long as the results are crystal clear the authors can take some liberties in the discussion. As long as it is clear what is speculative*

Objective interpretation	Statements on the importance of the objectivity of the study´s interpretation	NA	NA	*As long as it is an objective interpretation without any confirmation bias*
Related to other items	Statements on the link of the item with other items	NA	NA	*See comments on methods above*
Related to study flaws	Statements on the importance of commenting on the item especially if there are major errors in the manuscript	NA		*This is important if there are any major flaws or if an alternative explanation for findings should be considered*
Related to study implications	Statements on the importance of commenting on the item for generating new hypothesis to test	NA	NA	*To me, this is the most important issue, for the point of papers is to generate new hypotheses to test. Unfortunately, in my field, editors often want to see the facts, but are wary about interpretations, probably about long-winded speculation in the past*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Statistics expertise	NA	*If the reviewer is experienced in statics can make a good interpretation of the results*
Rushed interpretation as common problem	Statements on the poor interpretation of the study´s results as common problem	NA	NA	*Discussion is the most important part of the manuscript. And sometimes it is a bit rushed by authors*
Scope of the PR process	Statements on the scope of the peer review process	Assisting editors to understand results	NA	*Editors aren't technical experts in every field. The PR process*

				is therefore important is assisting editors understanding the significance of results
Subjective item	Statements on the subjective interpretation of the term "interpretation of the results"	NA	NA	*All results are open to a variety of interpretations*
Wording of the item	Statements on how to improve the wording of the item and better define it	NA	NA	*The reviewer must comment on the discussion section, of which the interpretation is a part. But other elements (comparison with existing research etc.) is also important. I would replace 'interpretation' with 'discussion'*

Strengths and weaknesses (general) (n= 21)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author's responsibility	Statements on the author's responsibility to discuss the strengths and weaknesses of the study in the manuscript	NA	NA	*The paper should, the reviewer only should if the paper is missing something important*
Important when manuscript is overly long	Statements on the importance of the item when the manuscript is overly long	NA	NA	*Important when a manuscript is overly long*
Specificity of the comments	Statements on the importance of the specificity of the comments	NA	NA	*Specificity is more important*
Related to other items	Statements on the link of the item with other items	NA	NA	*This is repeated above for methods. So these two components overlap. However, I mark this as important*
Related to the study	Statements on the importance of commenting on the strengths and weaknesses of the study´s aims and study´s flaws	Related to study aims	NA	*I think this needs to be specifically related back to clear study aims and objectives (perhaps this is a separate category? If not I think it should be). Even a beautiful study design with great validity and statistics is rubbish if it doesn't allow you to answer your research questions! I always look to see if the analyses and interpretation address the goals of the study*

		Related to study flaws	NA	*This is important where there are issues*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Provided that the reviewer has the methodological skills to comment on methodological strengths and weaknesses.*
Taking into account reader's perspective	Statements on the importance of taking into account the reader´s perspective by peer reviewers	NA	NA	*Peer reviewers should take the reader perspective and ensure the manuscript is well balanced on these*
Including S&W in the general comments	Statements on including strengths and weaknesses in the general comments	NA	NA	*These should be clearly identified in the general comments*
Useful for editors	Statements on the importance of the item for editors for making an editorial choice	To make a decision	NA	*Important for deciding to accept or reject a manuscript*

Strengths and weaknesses (methods) (n= 29)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author's responsibility	Statements on author's responsibility to evaluate the item	NA	NA	*Authors should already be doing this*
Dependency on the methods quality	Statements on the importance of the item in relation to the quality of the methods of the study	NA	NA	*This really depends. If the methods are spurious, of course, this needs to be indicated*
Dependency on the type of journal and study	Statements on how the assessment of the item depends on the type of journal and study	NA	NA	*Methods are very important for our journal*
Focusing on the weaknesses	Statements on the importance of the item especially focusing on the weaknesses of the study	NA	NA	*Important, especially the weaknesses, where there is an obvious need*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Very important item	NA	*This is absolutely key to the interpretation of the study. Unfortunately most reviewers, in my field, do not fully understand current (and correct) methods*
Related to other items	Statements on the link of the item with other items	NA	NA	*Yes, but it is confusing to separate this from the general strength and weaknesses. The question should be if the reviewer thinks that the message can (potentially) answer the research question*

Reviewers' expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Providing the reviewer is suitably qualified to comment on the methods. In my experience, far too many reviewers comment on aspects of the Methods for which they are able to confidently comment on*
Subjective item	Statements on the subjective interpretation of the term "strengths and weakness (methods)"	To give indication on own stance	NA	*As long as the reviewer gives some indication of his/her own stance in determining strength and weakness*
Commenting on the transparency of methods used	Statements on the importance of commenting on the transparency of the methods used by the peer reviewers	NA	NA	*Authors should already be doing this. Is it necessary for reviewers to also state this, or is it redundant? Reviewers could comment on whether the authors have been transparent about the strengths and limitations*
Usefulness	Statements on the usefulness of the item for both editors and authors	Useful for authors	NA	*For the author if the strengths and weaknesses are not properly addressed in the paper*
		Useful for editors	NA	*This is important for the editor to make a decision*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*Appropriateness of methods (based on question) may be more important...unless there is a problem, then strengths and weaknesses becomes important*

Statistical methods (n=115)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*For some articles (e.g. RCTs, meta-analyses, and observational studies) assessment of the statistical methods is important. For other types of articles (reviews, commentaries, editorials) this is of less importance*
Editor's responsibility	Statements on the editor´s responsibility to determine the necessity of a statistical review for a study	Employment of statistical assessors by journals	Statistics sub-editor	*Journal needs to engage a statistics sub-editor for that*
		No familiarity with some methods	NA	*Sometimes editor may not be as familiar with certain statistical methods which makes it more difficult to do this*
		Statistical support to the reviewers by journals	NA	*Journal editors should provide statistical support to reviewers*
		To determine the necessity of statistical review by journals	NA	*But I think that a lot (?) of reviewers are not sufficiently capable to do so. It might be more appropriate that the editor determines the necessity of statistical review and explicitly asks the reviewer if he/she is capable to do so*

No consensus on appropriateness of statistical methods	Statements on no consensus on the appropriateness of methods by peer reviewers	Variety of appropriate methods	NA	*Often there are multiple "appropriate" methods. It is important to use one of the appropriate methods*
Optional component of quality	Statements on the item as option component of the quality	NA	NA	*This should be an optional or "where relevant" component of the quality*
Related to other item	Statements on the link of the item with other items	NA	NA	*This belongs to assessing the methods. Should not be a separate item because there is qualitative research*
Reviewers' expertise	Statements on how the assessment of the item depends on reviewer´s expertise	At least one reviewer	NA	*At least one of the reviewers should have reasonable statistical knowledge*
		Content expertise	NA	*Many reviewers may be subject matter experts but not necessarily experts in statistics*
		General methods reviewer	NA	*This applies to the methods in general, whether or not they are statistical. I think at least one reviewer needs to comment on methods, but not every reviewer. Articles may not use statistics but an expert on the methods should review the article*
		Inappropriate advice	NA	*Non-statisticians should not be encouraged to comment on the statistical methods*

		Statistical reviewer	NA	*Statistical reviewer should do this*
		Commenting on own statistical expertise	NA	*In my view, the reviewer should be required to state whether or not she/he has the expertise to evaluate the statistical methods properly*
Commenting on the use of statistical methods	Statements on the importance to comment of the appropriate use of the statistical methods by authors	NA	NA	*And the use of stat. methods (some methods are used incorrectly by authors)*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*We get a lot of qualitative work so the key to this question is the appropriateness of the methods and then specifics based on type of methods*

Methodological quality (n= 32)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*Where applicable it is important*
General comments	General statements	NA	NA	*But in any case, I think internal validity is very important -- if a study is claiming that there is a relation between two variables it should be on solid ground to do so*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Very important item	NA	*In my opinion this is the most important item*
Focusing more on methods than results	Statements on the importance of commenting on the methods	NA	NA	*In my opinion this is the most important item. I think a reviewer should primarily focus on methods rather than results*
Related to other items	Statements on the link of the item with other items	NA	NA	*I did not understand the difference between this question and the question "The reviewer comments on the strengths and weaknesses of study methods"*
Reporting of the study	Statements on the importance of good reporting for study reproducibility	NA	NA	*A description should be sufficient to repeat the study with a high likelihood to end up with the same results*

Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Acknowledgement lack expertise	NA	*Ideally yes. However good reviewers are also well aware of the limits of their own expertise. It is better for reviewers to acknowledge that they lack expertise in relation to a particular aspect of the methodology (this is OK - no one is expected to be an expert in everything) rather than attempt to critique something that is outside of their own scope of knowledge*
		At least one reviewer	NA	*At least one reviewer with expertise in methods should review the study*
		Technical vs. clinical expertise	NA	*When selecting reviewers one might choose some for technical expertise (methodology, statistics etc.) and others for clinical expertise/experience*
Making sure results are not biased	Statements on the importance of the study for making sure the study results are not biased	NA	NA	*Crucial to make sure the results are not biased*
Wording of the item	Statements on the wording of the item and how to improve it	Broad item	NA	*"Methodological quality" is a broad term that could apply to construct validity and statistical validity as well as internal validity*

Applicability and external validity (n= 37)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Author´s responsibility	Statements on the responsibility of authors to comment on the applicability and external validity of the study providing sufficient information to the readers	NA	NA	*The paper should, not necessarily the reviewer*
Dependency on the practice of the reader	Statements on the difficulty to judge the importance of the item because it depends on the practice of the readers	NA	NA	*Difficult as would depend on the context of practice of the reader*
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*It depends on the specific topic of the study*
Editor's responsibility	Statements on the editor's responsibility to evaluate the item	NA	NA	*This is more 'scope, which is for the editor to decide. But help from a reviewer is appreciated*
Helping the editor to understand reproducibility of the study	Statements on the importance of the item to decide if a study can be reproduced	NA	NA	*This is important when reporting novel findings as it helps the Editor to decide if the results can be reproduced by another group*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Less important item	NA	*This is of lesser importance. There should be sufficient information included for any*

				readers to come to this conclusion themselves
Related to other items	Statements on the link of the item with other items	NA	NA	*Similar to my answer about interpretation*
Related to the paper	Statements on the claims made in the paper by the authors and impact of the study	Future research	NA	*Applicability might lie in the future, not in the present*
		Impact	NA	*This is important only in relation to the claims made in the paper about the impact and implications of a study*
		Related to the claim & content of the paper	NA	*This depends on the claims made*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	Technical vs. clinical expertise	NA	*When selecting reviewers one might choose some for technical expertise (methodology, statistics etc.) and others for clinical expertise/experience*
Reviewers' comments characteristics	Statements on the different characteristics of a peer review report	Baring reviewers' opinion	NA	*This may be one area of the review where the reviewer can bring a personal opinion to bare. Does the reviewer think this is a useful paper?*
		Evidence based comments	NA	*Peer reviewers should provide citations (evidence) for their assessment. Simply saying that the results are not applicable to their practice is not enough*

		Tempering authors' enthusiasm	NA	*Similar to my answer about interpretation: this is an area where the reviewer may have a valuable role in tempering an author's enthusiasm, hubris or bias*
Subjective item	Statements on the subjective interpretation of the term "applicability and external validity"	NA	NA	*This can be very subjective and misleading*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*Applicability and external validity are two concepts, so this item is double-barrelled in not valid*

Presentation and organization (n= 45)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Dependency on the type of journal (and its policy)	Statements on how biomedical journals differently evaluate the item based on their own criteria	Presence of copy editors in the journal	NA	*Depends a bit on whether journals have good copy editors*
		Taking into account the average reader of the journal	NA	*The reviewer needs to take into account the "average reader" of the journal - will they understand the paper?*
General comments	General statements	NA	NA	*Peer review is not an editorial exercise, but clarity and reproducibility are part of good science*
Useful for editors	Statements on the usefulness of the item for editors	NA	NA	*Because the readability is important to those who've not seen it before. Especially helpful when a handling editor is new, I think.*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Less important item	NA	*This is less important, because as long as the content is there, a reader should be able to make use of the paper, even if it requires more effort. But if the presentation and organisation is really bad, then it needs to be addressed*
		Presentation more important than organization	NA	*Presentation is important organization is not*

Improving the manuscript	Statements on the importance of the item to improve the quality of the manuscript	Clear recommendations	NA	*Yes, but in a way that provides the authors with clear recommendations on how to make improvements. Design flaws cannot always be addressed after the study, but issues with presentation and organization of the manuscript can*
		Communication	NA	*Important when it will help improve the quality of the communication. Not necessary when it flows well*
		Readability	NA	*Important because this impacts readability*
Not going into irrelevant comments	Statements on the importance of not making useless comments	NA	NA	*Important when it will help improve the quality of the communication. Not necessary when it flows well.*
Formatting minutiae	Statements on peer reviewers focusing on minutiae	NA	NA	*Some reviewers focus on formatting minutiae*
Related to reporting guidelines	Statements on the link of the item with reporting guidelines	NA	NA	*I find reviewer comments on the presentation and organization of the manuscript moderately important if the manuscript follows a check list (e.g. STROBE) and/or standard formatting, and if is easy to understand and follow*

Responsibility	Statements on editor, author or reviewer´s responsibility to evaluate the item	Joint responsibility	NA	*I think this is the role of the editors as well as the reviewers.*
		Editor's responsibility	NA	*Editors and editorial staff have a stronger role here.*
		Reviewer´s responsibility	NA	*I regularly make notes as to whether a section is better placed elsewhere in the document, and on sentence structure, and use and misuse of citations. I think this is an obligation that reviewers have to the author and the journal*
Subjective item	Statements on the subjective interpretation of the item	NA		*This is subjective and may vary between reviewers as log as general structure is preserved*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*The word "presentation" seems unclear. It could refer to the writing quality or to other factors*

Adherence to RG (n= 73)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Adherence to key points	Statements on the importance that a manuscript adherences on the key elements of a checklist	NA	NA	*I think whether a manuscript adheres to a specific item on a checklist is not that important. Adhering overall to the key elements needed to report is important*
Part of the PR process	Statements on the importance of checking the adherence of reporting guideline as part of the peer review process	NA	NA	*If it is widely accepted reporting guidelines like the Consort Guidelines I think that is an important part of peer review*
Author's responsibility	Statements on the author´s responsibility to follow reporting guidelines	Too demanding for authors	NA	*Some reviewers are too much strict on that*
Dependency on the type of journal	Statements on how the assessment of the item depends on the type of study	Consistent format	NA	*It would be great to have a consistent format and rubric to follow to increase comparability of manuscript and distress authors*
Dependency on the type of study	Statements on how the assessment of the item depends on the type of study	NA	NA	*Depends on type of study. For systematic reviews of course fundamental. For other studies this will be more and more important for easier comparisons between studies and for quality improvement. It makes our work easier if the*

				authors also compliance also improve
Editor's responsibility	Statements on the editor's responsibility to evaluate the item	Joint responsibility	NA	*The editor can also take care of this aspect*
		Pre-review	NA	*I believe this is the editor's job pre-review*
		Reformatting articles	NA	*We accept manuscripts that have been formatted for other journals for peer review. Of course we move towards acceptance they need to be reformatted*
General comments	General statements	NA	NA	*Universal reporting guidelines, like CONSORT, can be expected that all reviewers would know*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item		*Essential*
Lack of awareness	Statements on the lack of complete awareness about reporting guidelines from respondents	NA	NA	*In my experience, reviewers know little about the reporting guidelines of the journal for which they are reviewing. I think reviewers should always be sent details of the key guidelines. Otherwise they make criticisms or suggest changes which are incompatible with the guideline of the journal*

Providing reporting guidelines	Statements on providing reporting guidelines to reviewers	NA	NA	*The reviewer should receive the reporting guidelines applicable to the manuscript under review*
Reviewer's responsibility	Statements on the reviewer´s responsibility to check if the manuscript adherences to reporting guideline	Less reviewer's responsibility	NA	*I see that this is less the reviewer's responsibility to be honest*
		Making easier for reviewers	NA	*Important for improving standards in reporting, but this should be made as easy for the reviewer as possible, because otherwise it can be too arduous*
		Discussion of the study's issues	NA	*Pointing out where the manuscript does not respect the guidelines is useful, but more important is discussing the issues themselves*
		Tedious for reviewers	NA	*When doing reviews, it is quite tedious to have to relate to difference reporting and formatting guidelines of particular journals*
Getting an accurate review	Statements on how reporting guidelines help delivering an accurate review	NA	NA	*That always irritating when authors do not follow the recommendations oto authors starting from pagination... which helps for delivering an accurate reviewing*

Unclear responsibility	Statements on the unclear responsibility of checking for the adherence of the manuscript to reporting guideline	NA	NA	*I am not sure whether this is the peer reviewers' or the editor's responsibility*
Utility of reporting guidelines	Statements on the uncertain utility of reporting guidelines	NA	NA	*I'm not convinced that reporting guidelines make that much difference, but they are certainly better than nothing*
Wording of the item	Statements on how to improve the wording of the item	Meaning of reporting guidelines	NA	*I simply do not know what this means. Which reporting guidelines?*

Structure of reviewer´s comments (n= 33)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Content and completeness are more important	Statements on the more importance of review´s content and completeness	NA	NA	*Completeness is more important than how structured*
Definition of structured and organized	Statements on how to define the item	NA	NA	*General comments (e.g. on style) followed by structured comments (line by line or section by section)*
Dependency on the structure of the manuscript	Statements on the importance of the item in relation to the structure of the manuscript	NA	NA	*Again - somewhat dependent on the structure of the manuscript that is being peer-reviewed*
General comments	General statements	NA	NA	*Peer review is not an editorial exercise, but clarity and reproducibility are part of good science.*
Useful for both authors and editors	Statements on the importance of the item in helping authors and editors	Making easier to answer	NA	*Makes it easier for the authors responding*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Not so important	NA	*It helps, but I'm not sure this is important enough to be assessed. Should covary strongly with other characteristics of the review*
Not related to meaningful content	Statements on the no relationship between a well-structured review and meaningful review's content	NA	NA	*Makes it easier to respond to but doesn't mean the review content is more or less meaningful*

Related to other item	Statements on the link of the item with other items	NA	NA	*Makes it easier to respond to comments if they are clear and easy to read*
Review reorganized by editors	Statements on the reorganization of a review by editors	NA	NA	*It is up to the editor to interpret the referee comments and make concrete recommendations or demands on the authors if needed.*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers	NA	NA	*Semantic point, be careful about asking too much from unpaid and unrewarded reviewers*
Standard structure of a review	Statements on the necessity to have a standard structure for reviews	Different perspective	NA	*Organised according to who's perspective...one person's structure is another's chaos?*
		More difficult for reviewers	NA	*But the more you set exacting standards for a review, the more difficult you make it for a reviewer. This is undoubtedly something to aim for, but reviewer time is an issue*
Time consuming to reorganize the review	Statements on the time lost in reorganizing the reviewer´s comments	NA	NA	*Otherwise time is lost in trying to reorganize and understand what the reviewer means*

Clarity (n= 26)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Editors can make the comments clearer	Statements on editor´s task to edit the reviewer´s comments	NA	NA	*Helpful but not essential since the editor can help make sense of them for authors*
General comments	General statements	NA	NA	*Peer review is not an editorial exercise, but clarity and reproducibility are part of good science*
Useful for authors and editors	Statements on the usefulness of a clear peer review report for both authors and editors	NA	NA	*Otherwise neither the editor nor the authors can use the review appropriately*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Clarity is important*
		Less important	NA	*As long as the authors can understand the meaning, it is more important that the paper is clear*
Not a marker of quality	Statements on not considering clarity as marker of quality	NA	NA	*To me, although this is essential, it is more of an expectation of the review, rather than a marker of quality*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers	NA	NA	*But also conscious that we're all writing reviews late at night and so sometimes the ideal 'slips'*

To avoid repeated cycles of PR	Statements on the importance of the item to avoid repeated cycles of PR	NA	NA	*Yes - to avoid repeated cycles of peer review*
Useful for authors	Statements on the usefulness of the item for authors	Authors can ask for further clarity	NA	*It should be acceptable for authors to query reviewers' comments and ask for further clarity*
		Easy to respond	NA	*Makes it easier to respond to comments if they are clear and easy to read*
		Making sure the comments are intended	NA	*It is necessary to improve the chances that the comments are taken as intended*
Wording of the item	Statements on how to improve the wording of the item	Disagreement on the wording easy to read	NA	*I think easy to understand may be a better way to say this. I'm not sure easy to read is as applicable in the age of the computer*

Constructiveness (n=46)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Clear guidance	Statements on the importance to give clear guidance on how to improve the manuscript	NA	NA	*Worth emphasising that they should, where appropriate, give clear guidance on how paper might be improved and not be derogatory*
Dependency on the paper quality	Statements on how the constructiveness of reviewer´s comments depends on paper quality	NA	NA	*I suppose there will be some submissions which are so poor, this will be difficult*
Extent of the comments	Statements on the consideration to what extent reviewer`s comments could be addressed	NA	NA	*I think that reviewers should also consider to what extent their comments can be addressed. For example, if it's a paper on a survey, it's not helpful for a reviewer to say that more people should be surveyed*
General comments	General statements	NA	NA	*The worst reviews are the ones where the reviewer just rambles on and does not provide something to respond to*
Importance of destructive comments	Statements on the importance of destructive comments	NA	NA	*Some bellicose reviews are pretty helpful*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*I would rank this as the most important*

Marker of quality	Statements on constructiveness as a marker of quality	NA	NA	*Constructiveness flags the reviewer's interest in improving the paper, so it is a marker of the likely value of their review*
Negativity of the comments	Statements on the total negativity of reviewer´s comments	NA	NA	*The reviewers' comments are important, however at times, without any reasons the comments are totally negative*
Not mandatory requirement	Statements on constructive comments as not a mandatory requirement	NA	NA	*They can be, but it's not mandatory. Some manuscripts shouldn't be published*
Not reviewers' responsibility	Statements on how reviewers should not rewrite the paper but be respectful	NA	NA	*It is not the reviewers' job to rewrite the paper or mentor the authors. However comments should always be respectful*
Related to author's experience	Statements on how constructive comments are related to the experience of authors	NA	NA	*It depends on the status of the author. A beginner in a field needs encouragement and support. An older expert who is talking rubbish deserves more direct language*
Related to recommendation	Statements on how constructive comments are also useful if the manuscript is rejected	NA	NA	*Important, even if the recommendation is to reject: the authors will probably submit elsewhere, the comments can be useful for them in order to improve the paper*
Related to the readership's interest	Statements on how constructive comments are related to the interest of readership	NA	NA	*Unless manuscript really not of interest to readership, then I would not expect a reviewer to*

				spend a lot of time essentially helping the authors
Related to other items	Statements on the link of the item with others	NA	NA	*Hands in hands with being courteous*
Sometimes difficult to be constructive	Statements on how comments are sometimes difficult to present in a constructive way	NA	NA	*I am not sure this reflects quality - valid concerns over methodology, results etc. are sometimes difficult to present in a constructive way. Clearly being constructive is preferable though*
Subjective term	Statements on the subjective interpretation of the term "constructiveness"	NA	NA	*I think this is a subjective term*
Usefulness for both authors and editors	Statements on how constructive comments are useful for both editors and authors	Useful for authors	NA	*Directly linked to helping the author improve the manuscript.*
		Useful for editors	NA	*In case the review aims to support the editor to offer a revision, constructiveness of the review is more relevant*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*I want respectful and helpful. Sometimes that is different than "constructive."*

Detail/Thoroughness (n= 62)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Accommodating reviewer's comments	Statements on how authors accommodate reviewer´s comments	NA	NA	*On the other hand, authors often spend a lot of time with accommodating reviewer comments that were maybe not that relevant to start with. So there is a limit to how detailed and thorough is still helpful and the authors should have the right to reject some of the requests*
Dependency on the paper quality	Statements on how detailed comments depends on the quality of the paper	Detailed when paper is inadequate	NA	*Sometimes, where a paper is clearly inadequate producing a detailed report is necessary*
Detailed but not useful review	Statements on how detailed comments are not always useful	NA	NA	*They can be detailed but not useful--for example, when they concentrate on grammar and spelling*
For improving or rejecting the manuscript	Statements on the importance of detailed review to improve or reject a manuscript	NA	NA	*Sometimes the length of the comments is greater than the length of the manuscript. Peer reviewers should provide positive suggestions how the paper can be improved or rejected. The forma of the comments should be the same as the format of the responses, e.g. what I propose and why*

Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*Most important!*
Inconsistency in length	Statements on how peer review report are inconsistent in length	NA	NA	*The most disconcerting thing about reviews is inconsistency - when one is five pages long, the other five lines*
Not always necessary	Statements on how detailed comments are not always necessary	NA	NA	*Ideally but not always necessarily*
Providing a justification	Statements on the importance of providing a justification in the comments	NA	NA	*This is a pet peeve of mine. Some reviewers say things like "it has been demonstrated that this method of analysis is flawed" without providing a reference, for instance*
Related to other items	Statements on the link of the item with others	NA	NA	*I would have thought clarity was a more important criteria then being detailed but agree about thoroughness*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers	NA	NA	*Reviewers' time is valuable*
Straight to the critical points	Statements on the importance of succinct comments	Detecting fatal flaws	NA	*Focusing on one major flaw is more important than recitating all the typos*
		Excessive details	NA	*But they can be too detailed leading to a report that is too long overwhelming the author with too many requested revisions*

		Expectation from the authors	NA	*the most important is that the reviewer clearly indicate what he/she is expecting from the authors*
Wording of the item	Statements on how to improve the wording of the item	NA	NA	*"detailed, thorough and clear" (or unambiguous)*

Objectivity (n= 62)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Citing own work	Statements on reviewers citing their own work in a peer review report	NA	NA	*Please can reviewers not cite their own work. This usually means they have approached the paper with bias*
Declaration of COI	Statements on the importance of reviewer's conflict of interest declaration	More important than be objective	NA	*Declarations of competing interest and bias are more important than the claim to be objective*
Dependency on the study type	Statements on how objective comments are related to the type of study	Related to study's quality	NA	*This is a fundamental principle, that the comments should be disinterested (i.e. not driven by the reviewer's self-interest) as this increases the chance of the comments relating to the paper's quality*
		Related to the novelty of the study	NA	*Moreover, reviewer's comments are strongly influenced by the reputation of the author and the novelty of the idea. The less known the author and the more novel the idea, the reviewer tends to be less objective*
Editor's objectivity rather than reviewer's objectivity	Statements on the importance of objectivity from editors rather than peer reviewers	NA	NA	*Editors are supposed to be objective, so reviewers can be opinionated if they wish*

Following a specific rubric	Statements on the importance to follow a specific rubric to guide comments by peer reviewers	NA	NA	*Reviewer should follow a specific rubric to guide comments and make revision manageable by author*
General comments	General statements	NA	NA	*This is one of the most critical elements of good peer review in my opinion but also one of the rarest things to find*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*This is a fundamental principle, that the comments should be disinterested (i.e. not driven by the reviewer's self-interest)*
Impossibility to be total objective	Statements on the impossibility for reviewers to be totally objective	Comments are subjective by definition	NA	*All reviews are subjective!*
		Desirable to express own opinion	Awareness of own experience	*A better expectation is that reviewers come to the role aware of their own backgrounds, culture, experiences, research and views on the topic will affect their assessment of the research*
		Reminding reviewers to be objective	NA	*As far as possible - reminding reviewers to be as objective as possible would be a good start*
		Subjective comments are helpful for both editors and authors	NA	*I think there are subjective comments that are still valuable to the authors and editors. For example, if the reviewer finds a section of the manuscript to be unclear, this is there subjective*

				opinion but can still help the authors re-assess and potentially improve that portion of the manuscript
Justification of the comments	Statements on the importance of substantiating the comments by peer reviewers	NA	NA	*It is important that the reviewer substantiates comments, and that the authors are able to respond in case of revising their manuscript, to the reviewer's comments*
Recruiting additional reviewers	Statements on recruitment of additional peer reviewers when the reviews are not objective	NA	NA	*When peer reviewers recommend citing own papers or clearly favour one treatment over the others, editors should recruit additional peer reviewers*
Related to authors' reputation	Statements on how peer reviewers are influenced by author's reputation	NA	NA	*Moreover, reviewer's comments are strongly influenced by the reputation of the author and the novelty of the idea. The less known the author and the more novel the idea, the reviewer tends to be less objective*
Related to other item	Statements on the link of the item with others	NA	NA	*Goes along with a courteous tone*
Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Every reviewer will always have their own perspective based on their expertise*
Wording of the item	Statements on how to improve the wording of the item	Unclear item	Difficult to define	*It would be very difficult to define this*

Fairness (n= 55)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Biases are unavoidable	Statements on how reviewer's biases are unavoidable	NA	NA	*Reviewer biases are a reality, but one should (where possible) recognize them and phrase criticism in that light*
Editor's responsibility	Statements on the editor's responsibility to be fair	NA	NA	*Fairness is the editor's responsibility to judge*
Fair depends on author's characteristics	Statements on how the reviewer´s comments are influenced by author's characteristics	NA	NA	*It is well known that reviewers comments are not fair in terms of the location, ethnicity and gender of the authors*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	NA	NA	*Fairness is extremely important*
Importance to back up opinions	Statements on the importance of backing up opinions by peer reviewers	NA	NA	*Crucial that a reviewer backs up their opinion with evidence from the paper/published literature*
Justification based on the paper quality	Statements on how justification of reviewers is based on the quality of the paper	NA	NA	*As per above, the reviewers comments must be justified based on the quality of the article rather than on their personal views*
Recognizing COI	Statements on recognition of reviewer´s conflict of interest	NA	NA	*More specifically, Editors should identify if the referee*

				has potential conflict of interest especially if he/she can have a conflict of interest working on the same field or topics. We all know such
Related to other items	Statements on the link of the item with others	NA	NA	*Objective, evidence-based, fair etc. are highly correlated*
Reviewers' perspectives	Statements on the importance of having peer reviewers with different perspective	NA	NA	*We often seek reviewers with different perspectives, so the entire editorial review team is constructed to be fair*
Subjective item	Statements on the subjective interpretation of the term "fairness"	NA	NA	*Fair, of course, is subjective*
Wording of the item	Statements on how to improve the wording of the item	Simplistic way to assess quality	NA	*Donald Trump sees the world in terms of fair or unfair. I think this is too simple a view*
		Unclear item	Difficult to define	*How do you define 'fair'?*
			Difficult to measure	*How would you measure "fairness" of a review?*

Support by evidence (n= 69)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Context dependency	Statements on how the importance of item depends on the context	NA	NA	*Completely depends on the context; sometimes common sense can suffice but other times evidence-based critiques are necessary to show authors and editors why something needs changing*
Dependency on the type of study	Statements on how comments supports by evidence depends on the type of study	NA	NA	*Depends on the type of study*
Editor's responsibility	Statements on the editor's responsibility to determine if reviewer´s comments are relevant or not	NA	NA	*It is the editor's role to determine whether they are relevant or not*
Especially for supporting criticism	Statements on the importance of supporting criticism using evidence	NA	NA	*Comments, especially criticisms, should be supported by citations wherever possible; subjective criticism ("I prefer such and such a method ...") is not constructive*
Helpful when there are disagreements	Statements on the particular importance of the item especially when there are disagreements	NA	NA	*Particularly if authors objections are rooted in disagreements with their own opinions or collaborators opinions*
Importance of perceptions, opinions and experiences	Statements on the importance of perceptions, opinions and	Especially for too innovative manuscript	NA	*The exceptions are when the manuscript is too innovative or*

	experience of a reviewer in assessing a paper			*groundbreaking. In these case, knowledge and expertise to identify a possible major breakthrough is of utmost importance*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Very important. As an author and a frequent reviewer I have seen reviewer comments which are anything from completely wrong to simple statements of opinion without any evidence-base. But in my experience, what I would regard as quality reviews, up to date and accurate, are exceptionally rare*
		Not important item	NA	*Do you mean that they offer citations for their comments? If so, that's probably not important*
More information needed by authors	Statements on the necessity to get more information by authors	NA	NA	*Sometimes comments may be based on a hunch -- and more information from authors may be needed*
Unnecessary to provide evidence to each comments	Statements on how it is unnecessary to provide evidence for each comment	NA	NA	*I don't think reviewers need to cite something for every point that they make*
Related to other items	Statements on the link of the item with others	NA	NA	*If you mean, is it an objective review, then it is important*

Reviewer's expertise	Statements on how the assessment of the item depends on reviewer´s expertise	NA	NA	*Based on the reviewer's knowledge and experience*
Selecting use of evidence	Statements on providing the use of no selective evidence	NA	NA	*Provided the use of evidence is not selective*
Difficulty of backing up all comments	Statements on the difficulty of backing up each comments	NA	NA	*Wild claims may need back up but it is unrealistic to expect reviewers to back up everything they say. A quicker system for raising and addressing queries would be a better response*
Wording of the item	Statements on how to improve the wording of the item	Difficult to measure	NA	*But I'm not sure how you judge this*
		Providing citations	NA	*Do you mean that they offer citations for their comments?*
		Unclear item	NA	*I don't really understand*

Knowledgeability (n= 57)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
CoI between reviewers and authors	Statements on possible conflicts of interest between peer reviewers and authors	NA	NA	*Although this is very important it can create a conflict of interest as the authors and reviewers may be involved in the same field of research and this could result in a degree of bias for or against the research described in the manuscript*
General comments	General statements	NA	NA	*Reviewers should be able to commit time and effort to the process and be held accountable to the commitment.*
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Obviously this is a key requirement*
Related to other items	Statements on the link of the item with others	NA	NA	*Peer reviewers should have understanding of research methodology as well*
Responsibility	Statements on editor or author´s responsibility to evaluate the item	Author's responsibility	NA	*Failures in this can be about whether the authors have communicated their work clearly*
		Editor's responsibility	NA	*They have to try to understand it - but sometimes they do not. This is where the editor must cast a critical eye to ensure the*

				reviewer has grasped the essence of the article.
Review as guide for editors	Statements on how a review is a guidance for editors		NA	*Extremely important. Nothing more annoying to an author than realising the reviewer has not fully read the paper. Also crucial if the review is to provide fair guidance for editors*
Reviewer as disadvantaged position	Statements on how the peer reviews is often in a disadvantaged position		NA	*The reviewer is often at a disadvantage as he/she is given limited information on which to make a decision on whether to accept or reject the offer to review*
Reviewer as unpaid extra job	Statements on the voluntary job of reviewers		NA	*Yes, but see issue above about late night reviewing*
Reviewers as readers proxy	Statements on how the reviewer acts as a proxy for the reader		NA	*The peer reviewer acts as a proxy for the reader, so a basic understanding of the manuscript's content is important*
Reviewers' expertise	Statements on different reviewer´s expertise	Assessment reviewers' expertise	NA	*This is very important. I've long thought that one of the review criteria should be self-ratings of the reviewer's expertise in the substantive and methodological aspects of the article*

		Declaration of competence by reviewers	NA	*Reviewers should declare their competence in the subject of the manuscript*
		More reviewers	NA	*Some reviewers know about methods and some about content. It would be ideal to always have both, but that is often not the case*
		Understanding also research methodology	NA	*Peer reviewers should have understanding of research methodology as well.*
Wording of the item	Statements on how to improve the wording of the item	Confusing item	NA	*I found this question confusing. Are you asking if the reviewer is competent to evaluate the content of the manuscript?*
		Difficult to assess	NA	*Not sure how you would know if the reviewer knows and understands correctly the content of the manuscript*
		Disagreement with the wording	NA	*"Knows" and "understands" are distinct concepts and should not be combined here*

Timeliness (n= 49)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Better quality rather than on time	Statements on how a high quality review is more important than an on-time review	NA	NA	*But it is better to wait a while and have a high-quality review than to receive a quick, superficial and/or unfair review.*
Depends on the delay	Statements on how the importance of the item depends on the type of delay	NA	NA	*A few additional days of delay is not a major issue, while months of delay are*
Dependency on the type of journal	Statements on how biomedical journals differently evaluate the item	NA	NA	*Less important for pre-prints or F1000Research*
Difficult for editors	Statements on how long delay can cause difficulties to the editor	NA	NA	*Difficult for the editor if the delay is too long (or, worse, need to find another reviewer)*
Feasible and flexible deadlines	Statements on the importance to provide reasonable deadlines	Tendency to give short deadlines	NA	*Important, but there seems to be a trend among the editors to get reviews done in shorter amounts of time. Reminders are very helpful, but also some flexibility*
General comments	General statements	NA	NA	*The peer review process needs to be helpful for getting quality research into the public domain in a timely manner. It should not be a road block*
Golden rule	Statements on how to be on time is the golden rule	NA	NA	*It's the Golden Rule - it's just polite to be on time! Do as you would be done by etc.*

Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*Very important for us as we try to provide a rapid response to the outcome of papers*
Journal´s reputation rather than good science	Statements on how the journal´s reputation is more important than good science	NA	NA	*Far too many editors now are asking for reviews to be complete in too little time. This is being done for the good of the journal's reputation, not for the good of science*
More time does not mean more quality	Statements on how giving more time does not means having more quality	NA	NA	*Increasing time for revision doesn't add anything to the quality*
Nor related to the quality of PR process	Statements on how the item is not related to the quality of the entire peer review process	NA	NA	*Again not sure this contributes directly to the quality of the peer review process but is important in terms of ensuring that publication timetables can be adhered to*
Orthogonal factors to review quality	Statements on orthogonal factors to review quality	NA	NA	*Availability, timeliness, reliability (reviewing when says will review) are orthogonal to review quality.*
Related to reviewer´s professionalism	Statements on how to be on time is related to reviewer´s professionalism	NA	NA	*This is more on the professionalism of the reviewer rather than the quality of the review*
Reviewer as extra unpaid job	Statements on the voluntary job of reviewers	Demanding work	NA	*There are heavy demands on our time. I find to carry out a quality review on a paper which needs*

				expert reviewing to achieve the two aims laid out initially, that it takes a minimum of half a day, and often a day to do a good job
		Difficult to find reviewers	NA	*We are all busy! biggest issue is finding someone to say yes in first place*
		Little delay	NA	*On time is pretty important but a little delay really is not a problem. I'm speaking as a journal editor... I always have plenty to do so a week delay is probably fine and will not cause authors too much pain*
		Reasonable time for the reviewer	NA	*But are the deadlines reasonable?*
Scope of the peer review	Statements on the scope of the peer review process	NA	NA	*Of course, with the caveat that peer review is voluntary and usually being fitted in around other work activities*
Time given by the journal	Statements on how journals give different deadlines	NA	NA	*Depends on how much time the journal gives. 2 weeks is not enough!*
Wording of the item	Statements on how to improve the wording of the item	Unclear item	NA	*Important to define what "on time" means. For example, is one day late a problem? Or a week late if the authors tell you?*

Tone (n= 40)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Academia as though environment	Statements on the aggressiveness and competition in academia	NA	NA	*There is enough competition and aggressiveness in academia without having to receive reviews which are rude or condescending.*
Dependency on the author´s work	Statements on how the tone used by reviewers depends on the author´s work	NA	NA	*Depends on the author, and how much rubbish is being put forward*
Dependency on the paper quality	Statements on how the tone depends on paper quality	NA	NA	*And how much rubbish is being put forward*
Editor's responsibility	Statements on the editor´s responsibility to tone down the peer review reports	Removal comments by editors	NA	*Editor can tone down or edit out obnoxious comments, but it is better not to have to do this*
Golden rule	Statements on how to use a courteous tone is the golden rule	NA	NA	*Golden Rule again. Hiding rudeness behind anonymity is odious*
Hiding behind anonymity	Statements on how peer reviewers use anonymity to hide rudeness	NA	NA	*All too often, peer reviewers use the anonymity of the review process as an excuse to be rude and scathing in a way they would be unlikely adopt in person*
Impact of a rude review	Statements on how rude reviews can impact authors	NA	NA	*And if authors receive discourteous comments, this really does 'sour' the process*

				and put people off, especially new researchers
Importance of the item	Statements on the importance of the item in assessing the quality of peer review report	Important item	NA	*This is important. Especially to keep his/her nerves when looking at the first revised version and notice that the revision is not answering the queries and comments! Should also provide authors the keys to improve the paper and answer politely to referees...*
Not always a necessary requirement	Statements on how courteous tone is not always a necessary requirement	NA	NA	*This is nice, but not totally necessary*
Related to cultural differences	Statements on how courteous tone is culturally bound	NA	NA	*I think this is important, but courtesy is culturally bound*
Related to other items	Statements on the link of the item with others	NA	NA	*This relates to constructiveness, above*
Requirement	Statements on the requirement to use a courteous tone in a peer review report	NA	NA	*But it's an ethical requirement, and helps improve everyone's experience*
Review quality is important than courteous tone	Statements on how an higher quality review is more important than a courteous review	NA	NA	*Would you rather be treated by skilled, but rude, surgeon, or by a courteous flop?*
Useful for authors	Statements on the usefulness of the item for the authors	NA	NA	*The reviewer's aim should be to give comments that make the next version of the*

				manuscript better (whether or not it is accepted for that specific journal)
Wording of the item	Statements on how to improve the wording of the item	Difficult to define	NA	*It is impossible to define 'courteous' so I doubt that this is operationalisable*

Codebook 3. Identification of new items to assess peer review report quality

Identification of new items to assess peer review report quality (n=152)				
Theme	**Definition**	**Code**	**Sub-code**	**Example**
Characteristics of reviewer's comments	Statements on the characteristics of the comments made by a peer reviewer	Clarity	Clarity of the expected changes	*Being clear about the changes they want to see (vs. vague comments about weaknesses - what would most strengthen)*
			Clarity of the language	*The peer review report should be comprehensive and written clearly. It should not be ambiguous.*
		Constructiveness	NA	*Constructive attitude should include orientation and suggestion to authors to improve the manuscript.*
		Detailed	NA	*Specific details for concerns and suggestions for improvement are the keys for me. I need substantive concerns.*
		Evidence-based	NA	*A good peer-review report also includes references*
		Fairness/Unbiased	NA	*The comments should not only be fair, but also unbiased.*
		Specificity	NA	*I think that the challenge is that the comments are context*

				specific. Reviews are helpful when they identify something that is a challenge or something that can be done better. These rely upon the context.
		Structure of the peer review report	Additional comment to the editor	*Provides additional comments to the editor that provides context to the reviewer's assessment.*
			Explicit recommendation	*The reviewer makes an explicit recommendation about what decision to make – i.e., "reject", "revise and re-review", "accept", etc.*
			Initial summary	*The reviewer should begin her report with a short synthesis of the study (goals and main results)*
			Length of comments	*Length, very short peer review reports make me suspicious that they have even read the paper in enough detail. E.g. I once received a peer review "report" that had a single sentence along the lines of: "Good methods and results".*
		Tone	NA	*Candid*
Related to peer review process	General statements on the peer review process	Anonymity _PR process type	NA	*Reviewers and authors should be anonymous on both sides*

		Dependency on the type of journal	NA	*Reviewers should understand the nature of the journal that the manuscript had been submitted to - and should be prepared to state if he / she thinks the paper is not appropriate or relevant to that journal's readership.*
		Disclosure of reviewer's COI	Editor's task	*Conflict of interest. We often can tell that the journal has sent our paper to a reviewer who will not be objective in their review, and sometimes even when we've asked the editor to not use a particular reviewer. Editors have an obligation to insure a fair review, and often they do not. In these instances, the outcome is a foregone conclusion*
			Requests motivated by reviewer's COI	*The reviewer does not make requests that seem to be motivated by a competitive attitude or a conflict of interest.*
			Reviewer's publication record	*Publication record of the reviewer*
		Editor's responsibility	Balanced and fair decision	*Editor' decisions should also be balanced and fair, especially when reviewing are*

				discordant... Editors should also read papers... .
			Filtering reviews	*Editors need to protect authors from poor reviewers.*
		High profit of scientific publishing industry	NA	*The scientific publishing industry makes very high profits, in fact it is the most profitable "legal business model" among all economic activities.*
		Peer reviewers' training	NA	*A good reviewer needs to be trained: should be important to organize courses*
		Poor quality of the second review	NA	*Completion of a second review after the first draft-this is often poorly done*
		Quality scale	NA	*It might be helpful to consider one of the research quality ratings scales that are used in quantitative reviews.*
		Review quality as usefulness to make an editor's decision	NA	*The quality of a peer-review report that an author receives is partially determined by what the editors contribute to it before sending it to the author and how quickly they use it to make a decision. Too many don't send enough feedback, especially when two or more reviewers disagree.*

				This leads to three, four, or more back-and-forth "reviews" where reviewers are trapped in a cycle of disagreement and the editors won't make any significant contributions or a decision to resolve the disagreement.
		Reviewer as unpaid extra job	NA	*Reviewers are scientists that perform a professional service for the scientific publishing industry that in the vast majority of the cases is not paid.*
		Reviewer's final choice	Difference between major and minor revisions	*Always a grey area between Major revisions and Minor revisions that foxes a reviewer*
			Explanation choice	*If the reviewer makes a recommendation, e.g. accept or reject, they must provide a reason why. A review that just says "accept", "good work", is not a valid peer review.*
		Reviewer's recognition	Professional evaluators by publishers	*Alternatively, publishers may turn to "professional evaluators", who they may find in consultancy firms (KPMG, McKinsey, etc.) and pay their fees.....*
			Rewards for reviewers	*Payment for the reviewers should be considered, and this*

				would also increase the quality of the evaluations
		Saturation of the system	Professional evaluators by publishers	*So far reviewers are working "pro bono", and with the tremendous growth in the number of journals and the overall increase of the scientific activity worldwide (think just of the soaring number of papers coming from China in recent years) the system is becoming saturated, and reviewers becoming fed up*
		Scope of review	NA	*The number of items listed is a factor. If too many and in too minute detail, the article could be suppressed by the sheer workload of trying to address the comments*
		Weighting reviewer's comments	NA	*It is important for the ae to weight reviewers comments - some are rubbish and can be disregarded*
Related to the study	Statements on different aspects of a study that should be commented in a peer review report	About references	Suggesting relevant references	*Including references not known to the author*
		Addressing study's aims	NA	*I think the 'does this study address its stated aims' issue that I raised in my earlier responses is very important*

		Adherence to ethical guidelines	NA	*Comment on the study's adherence to ethical guidelines*
		Appendices	NA	*Use of supplemental material/appendices when appropriate (e.g., sensitivity analyses)*
		Applicability of the study	NA	*And general applicability*
		Authors' contribution and acknowledgments	NA	*Clearly articulate the role of every team member, and their contribution to the study. For evidence syntheses, require librarian involvement and give them authorship, the same with statisticians. Everyone in the team, without whose knowledge the study would not be possible, sound, or complete, should be acknowledged.*
		Context of the study	NA	*Puts the study in appropriate context*
		Data availability and software	NA	*Referees check the data availability and if new software actually works*
		Data quality	NA	*Quality of the data is most important*
		Ensuring disclosure of COI	NA	*Conflict of interests could be included*

		Ensuring inclusion of data sharing statements	NA	*Reviewers should ensure data sharing statements are included*
		Ensuring language quality	NA	*Comment on readability*
		Ethics	NA	*Ethical considerations of research*
		Importance of methods	NA	*Perhaps reviewing upto methods and evaluating the study that way is worth more consideration.*
		Literature is adequately reviewed	Most recent research	*Reviewer rating of whether The authors discuss the most recent relevant research on the topic*
		Originality	NA	*The added value of the study to what is already known.*
		Potential impact	NA	*Potential impact of study*
		Presentation (tables and figures)	NA	*And appropriateness of accompanying visual aids (graphs, tables e.t.c.).*
		Publication study's protocol and deviation from it	NA	*Whether a protocol was lodged in publication or on an independent site e.g. OSF and whether it matches the paper and if not, if reporting of deviations is transparent.*

		Relevance	NA	*relevance*
		Replicability/Reproducibility	NA	*Whether the study can be replicated on current methods whether limitations are acknowledged (this was covered actually I think)*
		Study conclusions	NA	*And finally if the conclusion answers the research question.*
		Study introduction	NA	*If the in introduction leads to the research question*
		Study limitations	NA	*Whether limitations are acknowledged*
		Study theoretical framework	NA	*Logic of the theoretical framework*
		Study weaknesses	NA	*Reviewer comments on the limitations of the study*
		Suggestions for future studies	NA	*Suggestions for future studies*
Reviewer's expertise		Considering reviewer's expertise	NA	*I have experienced vastly different qualities of reviews, so I think that the reviewer's expertise in the area of the article needs to be considered.*
		Focus on the points suggested by editor	NA	*Suggestions from editor re: which points/themes to focus on*

		Knowing dimensions not assessed by reviewers	NA	*What the reviewer feels they cannot comment on (e.g. is outside their expertise)*
		Rating or commenting on own level of expertise	NA	*The reviewer should state those aspects of the study for which they have limited knowledge*
		Reviewer's type	NA	*Items need to be tailored for whether the reviewer is a stats (methodology) reviewer or clinical expert*

NA= not available

Supplementary file 6. New items suggested by survey participants

New items	Example
1. Adherence to ethical guidelines	*"Comment on the study's adherence to ethical guidelines"*
2. Author's contribution and acknowledgements	*"Clearly articulate the role of every team member, and their contribution to the study. For evidence syntheses, require librarian involvement and give them authorship, the same with statisticians. Everyone in the team, without whose knowledge the study would not be possible, sound, or complete, should be acknowledged."*
3. Data availability	*"Referees check the data availability and if new software actually works"*
4. Disclosure of COI	*"Conflict of interests could be included"*
5. Data sharing statements	*"Reviewers should ensure data sharing statements are included"*
6. Study protocol	*"Whether a protocol was lodged in publication or on an independent site e.g., OSF and whether it matches the paper and if not, if reporting of deviations is transparent."*
7. Addressing study aims	*"I think the 'does this study address its stated aims' issue that I raised in my earlier responses is very important"*
8. Study introduction	*"If the in introduction leads to the research question"*
9. Study limitations	*"Whether limitations are acknowledged"*
10. Study conclusion	*"And finally if the conclusion answers the research question."*
11. Theoretical framework	*"Logic of the theoretical framework"*
12. Relevant literature	*"Reviewer rating of whether The authors discuss the most recent relevant research on the topic"*
13. Reproducibility	*"Whether the study can be replicated on current methods"*

Supplementary file 7. Explanations of the items included in the ARCADIA tool

Domain 1: Importance of the study

Item 1.a Contribution

A study can contribute to scientific knowledge in many ways: it can be a novel or confirmatory study with little or great impact on society and/or the research community. The contribution of a study is therefore not only associated to its novelty. Studies also need to be replicated in order to verify the validity of their results. The peer reviewer should discuss the importance of the study's research question.

Item 1.b Relevant literature

The peer reviewer should check if the authors reviewed the relevant research related to the study's topic in order to situate the study within the context of the existing literature.

Domain 2: Robustness of the study methods

Item 2.a Study methods

The peer reviewer should evaluate the soundness of the study methods, such as the selection of the study design, assessment of the risk of bias, etc., to understand whether the methods were appropriate to the study's aims, as well as if they were properly used and reported.

Item 2.b Statistical methods

Data can be analysed in many ways, but the only appropriate statistical models are those that fit well with the study design and the characteristics of the variables. The peer reviewer with expertise in statistics should assess whether or not the study followed a suitable statistical procedure, as well as if they were correctly conducted and reported.

Domain 3: Interpretation and discussion of the study results

Item 3.a Study conclusions

The reviewer should verify if the study conclusions answer the research question(s) and correctly summarize the study results.

Item 3.b Study limitations

The reviewer should check if the weaknesses of the study are correctly identified and discussed in order to interpret the validity of the research.

Item 3.c Applicability and generalizability

The reviewer should comment on the applicability and generalizability of the study results. Applicability and generalizability are two underlying concepts of external validity [1]. The first concerns how "the results from a sample can be extended to the population from which the sample was drawn", while the second how "the inferences drawn from study participants can be used in the care of patients drawn from any populations" [1].

Domain 4: Reporting and transparency of the manuscript

Item 4.a Study protocol

Public access to study protocols is important to increase transparency and reduce waste of biomedical research. In the case of previous publication and/or inclusion as an additional file of a study protocol, the reviewer should verify that the major deviations from it are reported in the manuscript.

Item 4.b Reporting

The reviewer should comment if the reporting of the study is clear, complete and transparent enough for facilitating its reproducibility by verifying the adherence of the manuscript to the corresponding reporting guideline. The Enhancing the Quality and Transparency of Health Research (EQUATOR) Network provides a toolkit to be used during the peer review process for selecting the appropriate reporting guideline [2].

Item 4.c Presentation and organization

The reviewer should discuss the quality of the written language used in the manuscript, as well as of how the study results are presented (tables, figures, etc.).

Item 4.d Data availability

When applicable, the reviewer should ensure that the data and materials (e.g., dataset, software codes), supported the results reported in the manuscript, are available.

Domain 5: Characteristics of the reviewer's comments

Item 5.a Clarity

A peer review report should be clear (meaning that readers can easily understand its content), succinct and well organized (following the manuscript sections and, when it is necessary, providing line and page numbers) in order to be understood correctly by editors and authors.

Item 5.b Constructiveness

A peer review report should contain constructive and polite comments that allow the authors to improve the quality of their work and editors to take a decision.

Item 5.c Objectivity

Comments provided in a peer review report should be as objective as possible and, if considered appropriate, include references to support the reviewer's statements.

References

1. Murad MH, Katabi A, Benkhadra R, Montori VM. External validity, generalisability, applicability and directness: a brief primer. BMJ Evid-Based Med. 2018;23(1):17–9.
2. EQUATOR Network. Welcome to our toolkit for peer reviewing health research! [Internet]. Available from: https://www.equator-network.org/toolkits/peer-reviewing-research/