



How to evaluate reviewers – the international orthopedics reviewers score (INOR-RS)

Andreas F. Mavrogenis¹ · Jing Sun² · Andrew Quaille³ · Marius M. Scarlat⁴

Published online: 5 July 2019
© SICOT aisbl 2019

Peer-review and reviewers

Peer-review (refereeing) is the process of subjecting an author's work to the scrutiny of others who are experts in the same field. It is a collaborative process that allows manuscripts submitted to a journal, conference or a book to be evaluated and commented upon by independent, usually anonymous experts within the same field of research [1, 2]. It provides feedback for authors to help to improve their work, and further, it helps the editor to decide whether the work should be accepted, revised, or rejected.

Peer-reviewers (referees) are experts qualified to perform reasonably unbiased evaluation in a topic. They are ordinary scientists who provide their pro-bono scientific work for peer-review within their busy schedule; they should provide valuable comments to the editor, and they are expected to alert the editor to any problems they identify [1, 2]. The quality of the reviewers, therefore, determines the quality of the peer-review. Traditionally, reviewers have been anonymous, but there is currently a significant amount of open peer-review publications, where the comments are visible to readers, generally with the identities of the reviewers disclosed as well [3, 4], (<https://publons.com/benefits/researchers>).

Although peer-review is the current best method to evaluate a scientific work, it is considered "...a flawed process, full of easily identified defects with little evidence that it works..." [5, 6]. Not all peer-reviewers are equally skilled in the peer

review process. Moreover, evaluation of peer-review and reviewers is difficult incorporating a lot of biases related to authors and reviewers' gender, nationality, native language, personal and institutional preferences [5, 7–12]. Additional challenges are the limited formal training and assessment methods for peer-review and reviewers compared to other research topics [13]. Yet, proper acknowledgement of reviewers and reward for the most dedicated is necessary.

The fact is that every single individual that accepts a peer-review invitation is fueled by great intentions. It is like a surgeon who is sure about the outcome of a procedure in a defined case. And actually this is not true; no one can predict with 100% accuracy that a patient will be better after surgery. However, the majority of patients will be improved. The outcome of a case is individual; the outcome of a procedure is depending on many variables. There are procedures with a high predictability of success, such as the total hip arthroplasty or high tibial osteotomy, while other procedures have variable success rates; some procedures are debatable in terms of outcome or patient benefit or hospital economics. Outcome studies are unpredictable. So is the reviewers' analysis. By one side we have the specialized and qualified medical expert who can provide a reliable and professional opinion on some written paper. By another side we have a paper authored by motivated colleagues. Depending on the type of paper the reviewers' analysis can be more or less valuable. In fact the experience of some individuals - such as acknowledged board members or editors of professional publications - can be so important that the prevision of the outcome for a given study is high. Is this study true? Is it meaningful? Does it contribute in some extent to the advancement of the current science or is it just a confirmation of previous studies? Maybe it is a meaningful contradictory hypothesis and the confirmation of such a research would open new sources of interest and stir debate. Being a reviewer or an editor is a voluntary act. It could be indelicate to criticize someone who does a voluntary job for helping others, especially if the reviews are positive or if they offer solution to improve someone's production.

✉ Marius M. Scarlat
mscarlat@gmail.com

¹ First Department of Orthopaedics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece

² Editorial Office of Orthopaedic Surgery, Tianjin Hospital, Tianjin, China

³ Spineworks, Hampshire Clinic, Basingstoke, Hampshire, UK

⁴ Clinique St. Michel, Groupe ELSAN, Toulon, France

Some reviewers are too young and have limited experience in peer-review. They will eventually and hopefully improve in time. Peer-review was originally meant for conformity with medical standards because in the seventeenth century when medical periodicals were started; one paper was basically a story told by a physician who described specific conditions or unseen stuff. Everything was new and everything was to be described. With time and advance in science the peer-review process became technically demanding and specialized. Journal editing and Production became more technical and professional too.

Characteristics of reviewers

The variability of the reviewers is a major weakness of the peer-review process and generates the subsequent disparate quality of their reviews [14–19]. Studies showed that reviewers at younger age, up to 60 years old, could provide higher-quality peer-reviews than older reviewers [14–17]. Nylenna et al. reported that the gender of the reviewers may be a factor in obtaining high quality peer-reviews, and female reviewers may provide stricter assessment [17]. Other studies reported that women are underrepresented in first authorships, and articles with women in dominant author positions attract fewer citations than those with men in the same positions [11]. USA nationality reviewers and USA papers are favoured [9]; USA reviewers recommend acceptance of papers submitted by USA authors more often than non-USA reviewers do, and USA reviewers rank USA papers much more favourably than non-USA reviewers do. Interestingly, non-USA reviewers also rank USA papers slightly more favourably than non-USA papers [9]. Authors' suggested reviewers produce reviews of similar quality to editor-chosen reviewers [8, 12]; however, the former are more likely to recommend acceptance [7, 8, 12].

Some studies showed that the lower in the academic ranks reviewers are, the better the peer-reviews they provide [16, 18]. Other studies showed that reviewers who produce good-quality peer-reviews tend to have background of training in research methods, epidemiology or statistics [15, 16]. Characteristics positively related to high-quality peer-reviews include the experience of previous peer-reviewing [17], additional degree [16], working in university environment [19], and membership of an editorial board [19]. Characteristics negatively related to high-quality peer-reviews are administrative positions [16], serving on IRB [19], and private practice [16]. Due to the complex and variable factors relevant to the quality of peer-review, a standard, effective and comprehensive evaluation instrument of the review should be established.

Evaluation scores of reviewers

There have been few attempts to evaluate reviewers for the quality of their peer-reviews [4], (<https://publons.com/benefits/researchers>), [20]. Publishers' evaluations relate to the readability, downloads and marketability of their journals; this is fair, but most likely evaluates items that have nothing to do with the quality of the research. Van Rooyen et al. [4] described the Review Quality Instrument (RQI) for reviewers. The RQI is an easy to calculate, seven item instrument that evaluates the importance of the research question, originality of the manuscript, strengths and weaknesses of the methods, provides comments about writing and presentation, constructive criticism, evidence to support comments and comments on interpretation of the results. The authors concluded that reviewers performed best on aspects that help the authors to improve the quality of their manuscripts, and less well on aspects that help the editors select papers (originality of the research), which is obvious because more reviewers are experienced as authors and not as editors [4]. An open peer commentary or open peer-review and a post-publication peer-review have also been suggested [4]; reviewers would record authors' complaints and authors would formally reply to all reviewers' comments and their reply will be published in its entirety for all readers of the journal to see and appraise. In these situations, a non-anonymous peer-review is recommended, on the basis that if reviewers have to sign their peer-reviews they may put more effort and produce better analysis [4].

The effort to properly credit the reviewers is an increasing concern for several journals. This concern has resulted in the recent creation of the Publons (publons.com), a website that lists the completed peer-reviews of their members, after crosschecking the database of the journal members (<https://publons.com/benefits/researchers>). Publons can be used to confirm the work done by reviewers worldwide and to find proper reviewers for each scientific topic. When a peer-review is performed for a journal listed in Publons' website, the reviewer is asked whether he wants to get recognition for his review on Publons, and whether he wants to track the review contributions for any journal. Additionally, Publons has instituted global peer-review awards to honour the critical role reviewers play in ensuring the quality and integrity of published research (<https://publons.com/benefits/researchers>).

Messias et al. described the Reviewer Index (RI) in a mathematical formula [20]. The RI is a ratio that considers the substantial importance of the reliability of the reviewers as measured by the number of peer-reviews completed, the time spent from peer-review acceptance to submission, and the editor's score (in cubic) [20]. We concur with the variables included in the RI but we believe that the factors related to the peer-review itself should also be included in the index and not be assigned in the editors' score in cubic.

The international orthopedics reviewers score (INOR-RS)

As editors, we feel responsible to our readers and authors to assure that the peer-review process will be done in a fair and timely fashion and that the final decision will be based on the scientific validity of the information in the manuscript. To evaluate who are the best reviewers and to determine if there are any characteristics that would tend to predict the quality of their peer-reviews, we routinely review the peer-reviews performed in our journal by the reviewers included in the journal's reviewers' panel. We evaluate the time taken to reply to an invitation, the numbers of accepted assignments and returned reviews, and the scientific quality of reviews. Based upon these observations, we created an index to measure the performance of the reviewers. This index, named the *International Orthopaedics Reviewers Score – INOR-RS* considers reviewers' and reviewing variables calculated using the outputs of the "Editorial Manager" workflow platform and scored by the assigned editor. The INOR-RS can be used specifically for Surgery Journals or in general for any scientific publication. More specifically, the INOR-RS is a five item score; each letter stands for a variable related to the peer-review(er), and scored from 0 to 4 points by the assigned editor (Table 1).

IN stands for Instructional and Informative peer-review. It is the type of peer-review that is most helpful to the editor, and it shows that the reviewer carefully read the manuscript, throughout its sections. An instructive and informative review should (a) indicate the novelty and significance of the work and if/how it adds to the current literature, (b) identify major flaws in the study

hypothesis, materials and methods, techniques, approaches and statistics, (c) denote statements requiring clarification, and (d) give details for areas to be improved and list changes to be made. If the above are met, the maximum score (3 points) is given; failure to identify major flaws and/or to communicate to authors criticism to improve their manuscript indicates a poor peer-review, and a lower score is given.

O stands for Organized and Objective peer-review. An organized peer-review should (a) be in formal and clear writing, well-structured, extend at least one half to two pages, (b) number the points and refer to pages, paragraphs of sections or lines in the manuscript, (c) review all sections of the manuscript including the citations and illustrations, and (d) include fair and impartial comments without any bias (for example on nationality, gender, opposite theory or conflicts of interest) or criticism for reasons unrelated to the quality of the work. If the above are met, the maximum score (3 points) is given. A non-organized review that fails to review the manuscript sections and a non-objective review indicates a poor review, and a lower score is given.

R stands for Responsible and Reliable reviewer. A responsible and reliable reviewer (a) accepts promptly (no later than 2 days) an invitation, (b) does not turn down any manuscript he is asked to review, (c) returns his peer-review within 10 days of the invitation, (d) provides confidential comments to inform the editor if any fraud is detected, and (e) discloses potential conflicts of interest, and declines the invitation if there are any with the authors. Comments to the editor should be unbiased and fair, as though the authors might read them. Very occasionally, a reviewer may decline an

Table 1 The International Orthopaedics Reviewers Score (INOR-RS)

INOR-RS	Criteria/Observations of the Peer-Review(er)	Score
IN	Instructional and Informative Indicates the novelty and significance of the work Identifies major flaws in the study hypothesis, materials and methods, techniques, approaches and statistics Denotes statements requiring clarification Gives details for areas to be improved and changes to be made	0-3 points
O	Organized and Objective Includes formal and clear writing and structure, extends one half to two pages Numbers the points and refers to pages/paragraph/lines in the manuscript Reviews all sections of the manuscript including references and illustrations Includes fair and impartial comments	0-3 points
R	Responsible and Reliable Accepts within 2 days an invitation Does not turn down any invitation Returns within 10 days of the request Provides confidential comments to the editor (fraud detection) Discloses conflicts of interest	0-4 points
Total		0-10 points

invitation; in this case, the reason should be communicated to the assigned editor, and ideally an alternate reviewer should be recommended. If the above are met, the maximum score is given (4 points); if not, a lower score is given.

The sum of the total points (0–10 points) indicates the INOR-RS for the respective reviewer for the respective manuscript. This score is inserted by the assigned editor at the review rate section of the “Editorial Manager” workflow platform, and it accompanies the reviewer at the reviewers’ details section of the reviewer’s panel of the “Editorial Manager” for the respective reviewed paper. The score may change at the next peer-review performed by the reviewer for another manuscript, and then the mean score is inserted at the reviewer’s details section of the “Editorial Manager” workflow platform.

Characteristics of the INOR-RS

A score to be useful should be easy and reproducible. It should include variables related to quality of the peer-review; these include constructive criticism, organization and synopsis of the review, and responsibility and reliability of the reviewer. The available reviewers’ scoring systems focus mainly on the manuscript, whether the reviewer puts comments or not, and if his/her comments are appropriate for the different aspects of the manuscript, such as the study design, methodology and structure. In contrast to complex formulas, the INOR-RS is a four item score that we consider easy to calculate and use. It addresses the quality of the peer-review by evaluating the structure, writing, instructions and constructive criticism, and the practice of the reviewers by evaluating their efficacy and attitude. These factors are important in the peer-review process to improve the quality of a submitted paper. Authors invest a great deal of time and energy into their research and have strong incentives to have their work treated fairly and published in time. With the increasing volume of submitted articles and their growing sophistication, editors should emphasize on time keeping in the editorial process. For this reason, we consider the responsibility of the reviewers to be very important.

Conclusion

The quality of peer-review depends on the quality of the reviewers. The INOR RS is a simple measurement of the quality of reviewers by ensuring that the evaluation performed is informative, constructive, organized,

objective, reliable and timely. We are indebted to the reviewers who provide their qualified and responsible collaboration to the scholar peer-review process, and we acknowledge constantly their involvement in the publishing process. Working with voluntary colleagues demands competence and expertise in evaluation and management. Provide something positive in return and interfering in a positive way with feedback and support. Being reviewer is sometimes frustrating and it is an editor’s responsibility to choose wisely the papers for peer-review in a stimulating way. Proper communication between authors, editors and reviewers determines the eventual success of a publication.

References

1. Brand RA (2012 Sep) Reviewing for clinical orthopaedics and related research. *Clin Orthop Relat Res* 470(9):2622–2625
2. Zazyva A, Kon E, Mauffrey C, Mavrogenis AF, Scarlat MM (2017 Jan) Reviews, reviewers and reviewing. *Int Orthop* 41(1): 1–2
3. Van Rooyen S, Godlee F, Evans S, Smith R, Black N (1999 Oct) Effect of blinding and unmasking on the quality of peer review. *J Gen Intern Med* 14(10):622–624
4. van Rooyen S, Godlee F, Evans S, Black N, Smith R (1999 Jan 2) Effect of open peer review on quality of reviews and on reviewers’ recommendations: a randomised trial. *BMJ*. 318(7175):23–27
5. Smith R (2006) Peer review: a flawed process at the heart of science and journals. *J R Soc Med* 99:178–182
6. Peters DP, Ceci SJ (1980) A manuscript masquerade. How well does the review process work? *The new York Academy of Sciences*. <https://doi.org/10.1002/j.2326-1951.1980.tb01346.x>
7. Schroter S, Tite L, Hutchings A, Black N (2006 Jan 18) Differences in review quality and recommendations for publication between peer reviewers suggested by authors or by editors. *JAMA* 295(3): 314–317
8. Wager E, Parkin EC, Tamber PS (2006 May 30) Are reviewers suggested by authors as good as those chosen by editors? Results of a rater-blinded, retrospective study. *BMC Med* 4:13
9. Link AM (1998 Jul 15) US and non-US submissions: an analysis of reviewer bias. *JAMA* 280(3):246–247
10. Knobloch-Westerwick S, Glynn CJ (2013) The Matilda effect in science communication: an experiment on gender bias in publication quality perceptions and collaboration interest. *Sci Commun* 35: 603–625
11. Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR (2013 Dec 12) Bibliometrics: global gender disparities in science. *Nature*. 504(7479):211–213
12. Kowalczyk MK, Dudbridge F, Nanda S, Harriman SL, Patel J, Moylan EC (2015 Sep 29) Retrospective analysis of the quality of reports by author-suggested and non-author-suggested reviewers in journals operating on open or single-blind peer review models. *BMJ Open* 5(9):e008707. <https://doi.org/10.1136/bmjopen-2015-008707>
13. Kovanis M, Porcher R, Ravaud P, Trinquart L (2016) The global burden of journal peer review in the biomedical literature: strong imbalance in the collective enterprise. *PLoS One* 11(11):e0166387

14. Kliwer MA, Freed KS, DeLong DM et al (2005) Reviewing the reviewers: comparison of review quality and reviewer characteristics at the American journal of Roentgenology. *AJR Am J Roentgenol* 184(6):1731
15. Nick B, Susan VR, Fiona G (1998) What makes a good reviewer and a good review for a general medical journal? *JAMA* 280(3):231
16. Evans AT, Mcnutt RA, Fletcher SW, Fletcher RH (1993) The characteristics of peer reviewers who produce good-quality reviews. *J Gen Intern Med* 8(8):422–428
17. Nylenna M, Riis P, Karlsson Y (1994) Multiple blinded reviews of the same two manuscripts. Effects of referee characteristics and publication language. *JAMA* 272(2): 149
18. Stossel TP (1985) Reviewer status and review quality. Experience of the journal of clinical investigation. *N Engl J Med* 312(10):658–659
19. Callahan ML, Wears RL, Waeckerle JF (1998) Effect of attendance at a training session on peer reviewer quality and performance. *Ann Emerg Med* 32(3):318–322
20. Messias AMV, Lira RPC, Furtado JMF, Paula JS, Rocha EM. How to evaluate and acknowledge a scientific journal peer reviewer: a proposed index to measure the performance of reviewers. *Arq Bras Oftalmol.* 2017 Nov-Dec;80(6):V. <https://doi.org/10.5935/0004-2749.20170084>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.