

Case Report ■

Developing and Evaluating Criteria to Help Reviewers of Biomedical Informatics Manuscripts

ELSKE AMMENWERTH, PhD, ASTRID C. WOLFF, PhD, PETRA KNAUP, PhD, HANNO ULMER, PhD, STEFAN SKONETZKI, MSc, JAN H. VAN BEMMEL, PhD, ALEXA T. MCCRAY, PhD, REINHOLD HAUX, PhD, CASIMIR KULIKOWSKI, PhD

Abstract Peer-reviewed publication of scientific research results represents the most important means of their communication. The authors have annually reviewed a large heterogeneous set of papers to produce the International Medical Informatics Association (IMIA) *Yearbook of Medical Informatics*. To support an objective and high-quality review process, the authors attempted to provide reviewers with a set of refined quality criteria, comprised of 80 general criteria and an additional 60 criteria for specific types of manuscripts. Authors conducted a randomized controlled trial, with 18 reviewers, to evaluate application of the refined criteria on review outcomes. Whereas the trial found that reviewers applying the criteria graded papers more strictly (lower overall scores), and that junior reviewers appreciated the availability of the criteria, there was no overall change in the interrater variability in reviewing the manuscripts. The authors describe their experience as a “case report” and provide a reference to the refined quality review criteria without claiming that the criteria represent a validated instrument for quantitative quality measurement.

■ *J Am Med Inform Assoc.* 2003;10:512–514. DOI 10.1197/jamia.M1062.

Research is defined as carrying out an investigation into a subject or problem.¹ Communicating research results in recognized, peer-reviewed scientific journals is essential both to scientific progress and to individual professional advancement.² However, as Seglen states: “Evaluating scientific quality is a notoriously difficult problem which has no standard

solution.”³ The authors’ motivation to develop and evaluate quality criteria for scientific papers arose during their work in editing the *Yearbook of Medical Informatics* of the International Medical Informatics Association (IMIA).^{4,5} The *Yearbook* appears annually and presents approximately 50 significant papers that have been published during the previous year. Its aim is to give a broad overview of the latest significant research activities in the field of health and medical informatics.

Affiliations of the authors: Research Group Assessment of Health Information Systems, University for Health Informatics and Technology Tyrol (UMIT), Innsbruck, Austria (EA); Department of Medical Informatics, University of Heidelberg, Germany (ACW, PK, SS); Institute for Biostatistics and Documentation, University of Innsbruck, Austria (HU); Erasmus University, Rotterdam, The Netherlands (JHV); National Library of Medicine, Bethesda, Maryland (ATM); Institute for Health Information Systems, University for Health Informatics and Technology Tyrol, Innsbruck (UMIT), Austria (RH); Department of Computer Science, Rutgers, The State University of New Jersey, New Brunswick, New Jersey (CK).

The authors are Editors and Managing Editors of the IMIA *Yearbook* 2001. The authors thank Martina Hutter, Steven Huesing, Thomas Kleinöder, and the Schattauer Publishing Company for their support in publishing the IMIA *Yearbook*. They also want to thank the fellow Managing Editors A. Bohne, K. Ganser, C. Maier, A. Michel, V. Mludek, and R. Singer for their discussions and comments on this paper as well as the 16 reviewers (S. Abel, B. Baumgarten, A. Bess, B. Brigl, T. Bürkle, E. Finkeissen, S. Garde, E. Lang, F. Leiner, F. Phillip, J. Pilz, U. Prokosch, M. Schwabedissen, R. Weber, T. Wendt, T. Wetter) for their participation in the study. Thanks also to Frieda Kaiser for her support and the anonymous reviewers for their fruitful comments on an earlier version of this paper.

Correspondence and reprints: Asst.-Prof. Dr. Elske Ammenwerth, Research Group Assessment of Health Information Systems, University for Health Informatics and Technology Tyrol (UMIT), Innrain 98, 6020 Innsbruck, Austria; e-mail: <elske.ammenwerth@umit.at>.

Received for publication: 11/21/01; accepted for publication: 05/06/03.

During the selection process, approximately 10,000 medical informatics papers published annually and listed in MEDLINE are reviewed and filtered to retain about 50 papers for publication in the *Yearbook*. Eight managing editors, each assigned to different subfields, first preselect papers for review; those (approximately) 150 papers are reviewed by two external international experts, by the two editors of the *Yearbook*, and by the responsible managing editor. A purely quantitative review scale is used for the final review. An analysis of the external reviews of 118 papers preselected for the IMIA *Yearbook* 2001⁶ indicated a wide range of variability in “expert” scoring, with approximately one-third of the papers showing a difference in quantitative scores of 20% or more among the two external reviewers.

The editors and managing editors attempted to refine the main review criteria with the goal of decreasing rater variability and improving quality of reviews. The authors of the current report conducted a review of the relevant literature; developed new, refined review criteria; and conducted a small randomized controlled trial using the new criteria to assess variability and reviewers’ satisfaction with the new criteria. The trial indicated that the new criteria were not useful, in an absolute sense, to create better agreement among reviewers. This is not surprising in that reviewers with different expertise and different experiences will view “relevant” aspects of a paper from unique perspectives.

The goal of this article is to report the authors' attempt to define quality manuscript review criteria and to share what was learned about their application during reviews.

Development of Refined Quality Criteria for Paper Reviews

The authors first reviewed the available literature on review criteria for scientific articles. For example, Gunn⁷ discussed problems with the quality of electronically published clinical guidelines, such as low quality and irrelevance. Elliott et al.⁹ presented guidelines for reviewing qualitative research, and Jefferson et al.¹⁰ assessed whether *BMJ* guidelines for reviewing economics submissions influenced the quality of submitted and published manuscripts in this area. Also, a standard was developed for measuring quality of publications regarding randomized controlled clinical trials (the Consolidated Standards of Reporting Trials, or CONSORT system).¹¹ CONSORT implemented a checklist containing 21 items, covering manuscripts' methods, results, and discussion sections.¹¹ The German Research Association, among others, has published general guidelines for good scientific practice¹²; their recommendation 12 covers (co-)authorship, completeness of presentation of research results, and correct citation of previous work of other researchers.

Many scientific medical informatics journals provide information regarding their own quality criteria via instructions for authors and guidelines for reviewers (e.g., *BMJ*¹³ and *JAMIA*¹⁴). Review criteria of medical journals (such as *BMJ*) are not always directly relevant to biomedical informatics publications.

The authors' literature review showed that no previously published quality checklist met the objective of providing a comprehensive list of refined quality criteria useful for reviewing all scientific papers in medical informatics.

The authors chose a top-down approach to refining the previously used IMIA Yearbook main review criteria (with sections for significance, quality of scientific content, originality and innovativeness, coverage of related literature, and organization and clarity of presentation). Analysis of available literature^{11,13-16} and the authors' own experiences as reviewers provided the basis for the refined criteria, which then were discussed and revised by IMIA Yearbook editors and managing editors in an iterative process.¹⁷

The revised review criteria developed by this methodology are presented on the Web pages of the IMIA Yearbook at <<http://www.yearbook.uni-hd.de>>. The revised criteria included five quality categories with 15 subgroups, totaling approximately 80 general questions, with approximately 60 additional questions for specific subtypes of articles. Table 1 (available as an online data supplement at www.jamia.org) highlights some of the differences between the review criteria of *BMJ*, *JAMIA*, and CONSORT and compares them with the authors' refined quality criteria.

Evaluation of Revised Quality Criteria and Lessons Learned

After developing and elaborating the revised quality criteria, the authors conducted a randomized trial, comparing the new review criteria with "standard" previous review methods (see Appendix A in an online data supplement at

www.jamia.org for methods and results of that study). The randomized study failed to show an effect of the revised criteria on interrater concordance. Nevertheless, the authors found the criteria to be helpful in the following manner.

1. The small-scale evaluation of the refined quality criteria showed that the reviewers' absolute quality ratings fell significantly (lower scores of merit) on papers they reviewed while applying the new criteria. Reviewers commented that the refined quality criteria helped to increase their awareness of all quality criteria, so that they more easily identified weaknesses in the reviewed papers, justifying lower ratings. Stricter ratings may be a desired effect, reflecting improved review quality—especially for less experienced reviewers. More experienced reviewers may already have most of the criteria in mind.
2. An observed increased time required to apply the refined quality criteria was not surprising. In the study, all reviewers had to grade each of the 140 individual review criteria for every paper reviewed, so that the study could be certain that the refined criteria had been applied. During "normal" reviews, it would be expected that refined criteria would serve as useful information for novice reviewers and as infrequently used but available reference material for more experienced reviewers. The authors believe that the increased time required to use the refined criteria may not reduce their usefulness as long as the criteria are available as a reference and not as an absolute requirement during review (a subject for future study).
3. Different reviewers of equal stature and ability will always judge the manuscript differently due to their varied scientific backgrounds, their variable familiarity with similar projects, and their variegated knowledge of the authors and the authors' prior work. It is therefore not surprising that the authors' randomized study found no reduction in variation between the reviewers after using the refined quality criteria.
4. The refined quality criteria do not constitute an instrument that can be used to obtain an objective quantitative assessment from reviewers and cannot lead to a "definitive" quality score. Rather, the criteria provide a list of important considerations that can help to support thorough reviews of scientific papers (even though the results will be different due to different backgrounds of the reviewers).
5. The reviewer (or author) of a paper can use the refined criteria as a reference in case of questions, for aspects to consider when judging originality or significance of a paper. Refined quality criteria may help novice reviewers to think more thoroughly about reviews (thus perhaps leading to better reviewers). The criteria may help in assessing how and why to arrive at certain decisions.
6. The list of refined quality criteria is not designed to be exhaustive and will be continually revised annually by the IMIA Yearbook editorial staff.

The quality of reporting in a paper does not automatically reflect the quality of the underlying research project it describes. But a good paper makes it easier to assess the quality of a research project. Publication quality is an important aspect of research quality.¹⁵

The authors hope that the refined review criteria will be helpful for authors of scientific papers in medical informatics

and for reviewers and editors to come to a balanced and more explicit assessment of the quality of medical informatics papers. The journal *Methods of Information in Medicine* has already adopted a draft of the refined quality criteria for its review guidelines.¹⁸

References ■

1. Pons Collins Dictionary of the English Language. Glasgow: HarperCollins, 1998.
2. de Vries J. Peer review: the holy cow of science. In: Fredriksson E (ed). *A Century of Science Publishing* Amsterdam: IOS Press, 2001, pp 231–44.
3. Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ*. 1997;314:498–502.
4. van Bommel J, McCray A. *Yearbook of Medical Informatics* (annual edition). Stuttgart: Schattauer, 1992–2000.
5. Haux R, Kulikowski C. *Yearbook of Medical Informatics* (annual edition). Stuttgart: Schattauer, since 2001.
6. Ammenwerth E, Knaup P, Maier C, et al. Digital libraries and recent medical informatics research—Findings from the IMIA yearbook of medical informatics 2001. *Methods Inf Med*. 2001;40:163–7.
7. Gunn IP. Evidence-based practice, research, peer review, and publication. *CRNA* (clinical forum for nurse anesthetists). 1998;9:177–82.
8. Boyer C, Selby M, Scherrer J-R, Appel R. The health on the net code of conduct for medical and health websites. *Comput Biol Med*. 1998;28:603–10.
9. Elliott R, Fischer CT, Rennie DL. Evolving guidelines for publication of qualitative research studies in psychology and related fields. *Br J Clin Psychol*. 1999;38(pt 3): 215–29.
10. Jefferson T, Smith R, Yee Y, et al. Evaluating the BMJ guidelines for economic submissions: prospective audit of economic submissions to *BMJ* and *The Lancet*. *JAMA*. 1998; 280:275–7.
11. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *JAMA*. 1997;277:927–34.
12. DFG. German Research Association: Proposals for safeguarding good scientific practice. Weinheim: Wiley-VCH, 1998.
13. BMJ. Checklist for editors and peer reviewers. <<http://bmj.com/advice/>>. Accessed October 2002.
14. Journal of the American Medical Informatics Association. Instructions for authors. <<http://www.jamia.org/misc/ifora.shtml>> Accessed October 2002.
15. Köbberling J. The quality of German medical journals [in German]. *Dtsch Med Wschr*. 2000;125:1106–8.
16. Paice E. How to write a peer review. *Hosp Med*. 2001;62: 172–5.
17. Kulikowski C, Ammenwerth E, Bohne A, et al. Medical imaging informatics and medical informatics: opportunities and constraints—findings from the IMIA Yearbook of Medical Informatics 2002. *Methods Inf Med*. 2002;41:183–9.
18. *Methods of Information in Medicine*. Instructions for authors. Accessed October 2002. <<http://www.schattauer.de/zs/startz.asp?load=/zs/methods/richtl.asp>>.