

# Evidence-Based Peer Review? Highlights of the Fifth International Congress

### Ana Marusic

Sixteen years after the pioneering work of Drummond Rennie and Annette Flanagin in peer-review research and the first congress in Chicago in 1989, the best in research on peer review came back to Chicago, where the Fifth International Congress on Peer Review and Biomedical Publication was held on 16-18 September 2005. During the 16 years, peer-review research flourished in the spirit of the Olympic Games—*citius, altius, fortius*. Seasoned CSE members will remember that the research presented at the First Congress was published in the book<sup>1</sup> by CSE (then the Council of Biology Editors), in addition to the theme issue of *JAMA*.<sup>2</sup> After 23 articles in 1989, the congresses in Chicago in 1993, Prague in 1997, and Barcelona in 2001 gathered increasing numbers of researchers and journal editors, as well as the number of articles published on peer review, mostly in the theme issues of *JAMA*.<sup>2-5</sup> The Fifth Congress—with 42 oral presentations during 3 days, 58 posters in two sessions selected from more than 200 submissions, and 470 participants from 38 countries—showed that there is more research than ever.

Most of the evidence presented at the congress illustrated the failures of peer-review and editorial processes. The journals' current focus on trial registration as a requirement for publication drew attention to research into the ties between the pharmaceutical industry and publications. The session on publication bias and funding and sponsorship showed that those ties were strong and influenced the publication of sponsored articles. Federico Soldani and colleagues were not at the congress to present their study of articles on drug treatment of bipolar disorder, but the numbers

from their abstract led to the disturbing conclusion “that even in highly respected journals, most reports on bipolar disorder are of limited methodological quality and that source of funding may have an important influence on psychopharmacology research results, conclusions, and dissemination”.

More bad news followed with the other presentations at the session. Veronica Yank and her colleagues showed that meta-analyses with disclosed financial ties to the pharmaceutical industry, although they had similar results and quality, had a higher proportion of positive conclusions in favor of the study drugs than meta-analyses without such financial ties. Anders W Jorgensen and Peter C Gotzsche, of the Nordic Cochrane Centre, studied pairs of meta-analyses of the same two drugs sponsored by either a nonprofit institution or industry and showed that all reviews supported by the industry recommended the use of the experimental drug, whereas Cochrane Reviews, reviews with unknown support, and those with nonprofit support or no support had more cautious conclusions. Moreover, published reports of randomized controlled trials may differ greatly from the original trial protocols, as shown by the study of Lucy Chappell and her colleagues in the UK journal *BJOG: An International Journal of Obstetrics and Gynaecology*.

The research presented at the Congress also showed that most of the editorial efforts to increase the quality of reporting trials are futile. In the session on this topic, An-wen Chan and his colleagues in a UK-Canada collaborative group showed that reports of crossover trials often omit important methodologic issues in design, analysis, and presentation. Catherine Walter and her colleagues, from the UK, showed that trials in operative surgery are similar in quality to those in other medical fields, but fewer than half of them reported adequate

methodology. A systematic review of the effects of adopting CONSORT guidelines for reporting randomized controlled trials on the quality of reports, performed by Amy Plint and her colleagues, in Canada, UK, and the United States, showed that journal adoption of CONSORT results in better description of the method of randomization and reporting of more checklist items but not in reporting other important items, such as allocation concealment and participant flow. Finally, the introduction of Standards for Reporting Studies of Diagnostic Accuracy (STARD) improved the quality of reported articles, but only to well below an optimal level, according to data presented by Nynke Smidt, of the Netherlands.

Of course, a meeting on peer review cannot pass without impact factors, the famous and equally infamous child of Eugene Garfield. His plenary speech in Chicago again (and again!) emphasized the true meaning of impact factors, as well as their misinterpretation and misuses, and the session on impact factors showed that they still matter for both big and small journals. Mabel Chew and her colleagues, of the *Medical Journal of Australia*, followed the impact factors of seven prestigious general medical journals and showed that they increased steadily over the 11 years of the follow-up. However, some achieved that by increased citations to their articles (increasing the numerator in the impact-factor equation), and the others by decreasing the number of citable articles (decreasing the denominator in the equation). Yuan-I Min and her colleagues, in the United States, confirmed that the journal is the site of confirmatory publication bias related to the impact factor: trials with statistically significant results are more likely to be published in high-impact journals. Reports from small journals showed that they increased their visibility by providing their content online and by

---

ANA MARUSIC is editor-in-chief of the Croatian Medical Journal.

## Peer Review continued

getting indexed in bibliographic databases; the examples presented were the *Journal de Pediatria* in Brazil, by Danilo Blank and his colleagues, and the *Journal of Postgraduate Medicine* in India, by Dev Kumar Sahu and his colleagues.

The sessions on research misconduct reminded us that although we may have good policies and procedures for handling misconduct, it may be hard to undo the damage in the published literature. The study presented by Rhonda Dailey showed that although articles flawed by misconduct have identifiable corrections or retractions in PubMed, few citing articles refer to them. The research of Benjamin Druss and his colleagues illustrated the problem of publication retractions because of unintentional errors in biomedical publications, which were substantially more common than retractions in the published literature because of misconduct.

Finally, research on the peer-review process again revealed how it is heavily burdened by bias. Two independent groups—Sara Schroter and colleagues, of *BMJ*, and Elizabeth Wager and colleagues, of *BioMed Central*—showed that reviewers suggested by authors, although not different in the quality of their reviews, tended to give more favorable recommendations than reviewers selected by editors. We also learned that it pays to cite the works of potential reviewers: the study of Matthias Egger and colleagues, of the *International Journal of Epidemiology*, showed that the more reviewers' work is cited in a manuscript, the more likely they are to give favorable recommendations.

So does anything work in peer review and biomedical publishing? Not much. Blinding reviewers to authors' identities and affiliations in the selection of abstracts for the annual meeting of the American Heart Association may increase the number of non-US and non-English-speaking countries, as shown by the research of Joseph Ross and his colleagues. And introduction of continuing medical education credits may be an incentive for peer reviewers, at least to provide a timely review, as reported by Mary Beth Schaeffer

and her colleagues, of *Annals of Internal Medicine*.

Although it became clear very early in the history of peer-review research that peer review and editorial work should not be studied as interventions in clinical medicine but rather as psychosocial and behavioral factors in publishing, the congress in Chicago presented little research on these aspects of peer review. There were two qualitative studies: one on peer reviewers and the other on editorial decision-making. Michael Callahan and John Tercier, of *Annals of Emergency Medicine*, showed that reviewers' attitudes and value systems differed from those of editors. Kay Dickersin and Catherine Mansell, of Brown University, examined the processes during editorial meetings in *JAMA* and reported that the most frequent discussion themes were scientific merit, journalistic goals, and writing—more than the statistical significance of study results.

A study by John Gilstad and Thomas Finucane, of the US National Naval Medical Center, explored the rhetoric of different scientific articles reporting clinical trials of the same intervention. They found that the rhetoric of the articles varied widely although the experimental findings were consistent, and they concluded that such rhetoric may influence clinical interpretations of the studies.

My research group presented the results of our analysis of psychologic aspects of authorship, studying contribution disclosure forms as self-reports of behavior. We showed that authors of scientific articles had problems with all cognitive aspects of survey methodology when filling out contribution disclosure forms: they may not understand the questions and have views of authorship different from those formally prescribed in the biomedical community, they may have difficulty in recalling relevant behavior, they have problems in inferring and estimating the behavior in question or in mapping the answer to the response format, and they may edit answers for social desirability. Another illustration of the subjectivity of the publication process was the study of

Penelope Green, of the Harvard School of Public Health, who showed in an elegant study that the same sets of original data analyzed by 12 statisticians yielded statistically (!) different interpretations.

After the Fifth Congress, you will not have to wait for the *JAMA* theme issue on peer review. This year, three journals—*JAMA*, *BMJ*, and the *Medical Journal of Australia*—will publish a selection of articles from the oral presentations. For those who will not publish their research in the three journals, perhaps it is time to renew the practice of the First Congress in 1989 and publish their research with *CSE*—in *Science Editor*! Judging by the interest shown at the *CSE* booth at the Congress in Chicago, especially regarding the invitation to publish editorial research in *Science Editor*, perhaps we shall not have to wait 3 years for a single theme issue on peer review in general medical journals. *Science Editor* is a great place to publish such research, not only from biomedicine but especially from other scientific fields with different peer-review and editorial practices. Only through interdisciplinary research across scientific fields can we start understanding how and why we behave like authors, reviewers, or editors. 🌐

## References

1. Peer review in scientific publishing. Reston, VA: Council of Biology Editors, 1991.
2. Guarding the guardians: research on editorial peer review: selected proceedings from the First International Congress on Peer Review in Biomedical Publication. *JAMA* 1990;263(theme issue):1317-441.
3. The Second International Congress on Peer Review in Biomedical Publication. *JAMA* 1994;272(theme issue):91-173.
4. The Third International Congress on Peer Review in Biomedical Publication. *JAMA* 1997;280(theme issue):213-302.
5. The Fourth International Congress on Peer Review in Biomedical Publication. *JAMA* 2002;287(theme issue):2749-871.