

6. Smith R. Opening up *BMJ* peer review. *BMJ*. 1999; 318:4-5.
7. Rennie D. Anonymity of reviewers. *Cardiovasc Res*. 1994;28:1142-1143.
8. McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. *JAMA*. 1990;263:1371-1376.
9. Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA*. 1998;280:237-240.
10. van Rooyen S, Godlee F, Evans S, Smith R, Black N. Effect of blinding and unmasking on the quality of peer review: a randomized controlled trial. *JAMA*. 1998;280:234-237.
11. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D. Does masking author identity improve peer review quality? a randomized controlled trial. *JAMA*. 1998;280:240-242.
12. van Rooyen S, Godlee F, Evans S, Black N, Smith R. Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised controlled trial. *BMJ*. 1999;318:23-27.
13. Walsh E, Rooney M, Appleby L, Wilkinson G. Open peer review: a randomised controlled trial. *Br J Psychiatry*. 2000;176:47-51.
14. van Rooyen S, Black N, Godlee F. Development of the review quality instrument (RQI) for assessing peer reviews of manuscripts. *J Clin Epidemiol*. 1999; 52:625-629.
15. Odlyzko AM. Tragic loss or good riddance? the impending demise of traditional scholarly journals. Available at: <http://www.research.att.com/~amo/doc/tragicloss.txt>. Accessed May 13, 2002.
16. Harnad S. Implementing peer review on the net: scientific quality control in scholarly electronic journals. In: Peek R, Newby G. *Scholarly Publishing: the Electronic Frontier*. Cambridge, Mass: MIT Press; 1995.
17. Ginsparg P. Creating a global knowledge network. Available at: <http://www.biomedcentral.com/1417-8219/1/9>. Accessed May 13, 2002.
18. Gura T. Peer review unmasked. *Nature*. 2002; 416:258-260.
19. Godlee F, Jefferson T. *Peer Review in Health Sciences*. London, England: *BMJ* Publishing Group; 1999.
20. Altman DG. Poor-quality medical research: what can journals do? *JAMA*. 2002;287:2765-2767.

Poor-Quality Medical Research What Can Journals Do?

Douglas G. Altman, DSc

THERE IS CONSIDERABLE EVIDENCE that many published reports of randomized controlled trials (RCTs) are poor or even wrong, despite their clear importance.¹ The results of several reviews of published trials are briefly summarized in TABLE 1. Poor methodology and reporting are widespread.

Similar problems afflict other study types. A review of 308 phase 2 trials in cancer (295 of which were single-arm studies) found that 250 (81%) did not report an identifiable statistical design. Further, positive findings were reported in 48% of designed studies but 70% of studies with no reported design ($P=.003$).³ Of 40 molecular genetics articles published in leading general medical journals, 15 (38%) failed to meet at least 2 of 7 methodological standards. The authors wrote: "Without suitable attention to fundamental methodological standards, the expected benefits of molecular genetic testing may not be achieved."⁴

In recent years, systematic reviews have become common. In these, all reliable evidence relating to a clinical question is sought, systematically appraised, and, if suitable, combined statistically in a meta-analysis.⁵ A key component is an assessment of the methodological qual-

The aim of medical research is to advance scientific knowledge and hence—directly or indirectly—lead to improvements in the treatment and prevention of disease. Each research project should continue systematically from previous research and feed into future research. Each project should contribute beneficially to a slowly evolving body of research. A study should not mislead; otherwise it could adversely affect clinical practice and future research. In 1994 I observed that research papers commonly contain methodological errors, report results selectively, and draw unjustified conclusions. Here I revisit the topic and suggest how journal editors can help.

JAMA. 2002;287:2765-2767

www.jama.com

ity of the individual (primary) studies.⁶ Reviewers often conclude that the available evidence is of poor scientific quality,^{7,8} sometimes leading to heated debate about interpretation.⁹

General reviews also find much to be concerned about. Serious statistical errors were found in 40% of 164 articles published in a psychiatry journal¹⁰ and in 19% of 145 articles published in an obstetrics and gynecology journal.¹¹ I suspect that many basic errors have become less common, but statistics has become more complex, and there is evidence of frequent misapplication of newer advanced techniques.¹²

Also, when interpreting a study, readers need to know how it relates to existing knowledge. Many authors interpret their findings narrowly, failing to

either identify previous studies or place their findings in the context of those previous studies.¹³

Why Are There So Many Errors in Medical Articles?

Errors in published research articles indicate poor research that has survived the peer-review process. But the problems arise earlier, so a more important question is, Why are submitted articles poor?

Much research is done without the benefit of anyone with adequate training in quantitative methods.¹⁴ Many in-

Author Affiliation: Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, England.
Corresponding Author and Reprints: Douglas G. Altman, DSc, Cancer Research UK/NHS Centre for Statistics in Medicine, Institute of Health Sciences, Old Road, Headington, Oxford OX3 7LF, England (e-mail: doug.altman@cancer.org.uk).

Table 1. Summary of Empirical Evidence of Prevalence of Methodological Problems in Published Reports of Randomized Trials*

Deficiency	Evidence
Failing to specify eligibility criteria	25% of 364 reports in surgery journals
Not reporting an adequate method for generating random numbers	68% of 206 reports in obstetrics and gynecology journals; 52% of 80 reports in general medical journals
Not reporting the mechanism used to allocate interventions	89% of 196 reports in rheumatoid arthritis journals; 48% of 206 reports in obstetrics and gynecology journals; 44% of 80 reports in general medical journals
Failing to state whether blinding was used	51% of 506 reports in cystic fibrosis journals; 33% of 196 reports in rheumatoid arthritis journals; 38% of 68 reports in dermatology journals
Incorrect analysis of multiple observations	63% of 196 reports in rheumatoid arthritis journals
Inadequate information on harmful consequences of interventions	61% of 192 reports in 7 medical areas
Incorrect method of comparison of subgroups	58% of 50 reports in general journals

*Data from Altman et al.²**Table 2.** Time Limit on Submitting Letters Commenting on Published Articles

Journal	Time Limit, wk	Word Limit
<i>Annals of Internal Medicine</i>	6	300
<i>BMJ</i>	4	400
<i>CMAJ</i>	8	250
<i>JAMA</i>	4	400
<i>Lancet</i>	8	500
<i>New England Journal of Medicine</i>	4	250

investigators are not professional researchers; they are primarily clinicians. “. . . [I]f they had any training in research methods it was usually a single course in statistics in the first or second year of their degree, before they really appreciated how important rigorous research methods are in order to do good science.”¹⁵ Also, training in statistics often focuses on data analysis, an emphasis reinforced by several statistics textbooks, often by nonstatisticians, in which design issues are not addressed.¹⁶ Yet study design is a crucial element of education in research methods and appropriately forms a key aspect of training in critical appraisal generally and evidence-based medicine in particular.^{17,18}

A contributory reason is inadequate review by research ethics committees (institutional review boards). Such review should detect studies with important flaws in design but clearly often fails to do so. Unfortunately, committees tend to use a narrow interpretation of ethics that downplays scien-

tific quality, despite the clear ethical implications of allowing research that is not scientifically valid.¹⁹

A further issue is the copying of incorrect or inappropriate methods. Once incorrect procedures become common, it can be hard to stop them from spreading through the medical literature like a genetic mutation. Many editors have wrestled with the problem of authors objecting to a reviewer's criticism on the grounds that the same methods have appeared in previous articles, quite possibly by the same authors in the same journal. Examples of incorrect practices that persist despite published warnings include using the correlation to compare 2 methods of measurement,²⁰ using significance tests to compare baseline characteristics in randomized trials,²¹ conducting multiple tests of data recorded at multiple times,²² and ignoring the clustering in the design and analysis of cluster randomized trials.²³

Peer review can and should weed out serious methodological errors. However, expert methodological input is in short supply. Only a third of high-impact journals reported statistical review of all published manuscripts.²⁴ The vast majority of research is published in low-impact journals where peer review is undoubtedly less thorough.

Postpublication Peer Review

Many readers seem to assume that articles published in peer-reviewed jour-

nals are scientifically sound, despite much evidence to the contrary. It is important, therefore, that misleading work be identified after publication.

As Gehlbach²⁵ noted, “[t]he ultimate interpretation and decision about the value of an article rests with the reader.” Recent draft recommendations from the World Association of Medical Editors say that “[e]ditors should promote self-correction in science and participate in efforts to improve the practice of scientific investigation by . . . publishing corrections, retractions, and letters critical of articles published in their own journal.”²⁶ Although journals do publish correspondence, there are weaknesses in the way they do so. Most obviously, editors select which letters to publish.

Editors should give special attention to letters making criticisms of methodology. They should do one of the following: satisfy themselves (perhaps by having the letter peer reviewed) that the criticisms are unfounded or unimportant, agree to publish the letter and invite the authors to respond, or invite a response from the authors and then decide whether to publish. Letters should not be rejected because of previously published correspondence (making different points) or lack of space.

Time limitation on correspondence denies readers the opportunity to draw attention to methodological deficiencies. TABLE 2 shows the current rules of 6 general medical journals. In effect, there is a statute of limitations by which authors of articles in these journals are immune to disclosure of methodological weaknesses once some arbitrary (short) period has elapsed, which cannot be right.

None of these journals suggests that there are exceptions, but from personal experience, at least 3 of them have occasionally published letters received beyond the stated time limit. The *BMJ* recently published a letter pointing out errors in an article published 6 years earlier. In it, Bland commented: “Potentially incorrect conclusions, based on faulty analysis, should not be allowed to remain in the literature to be cited uncritically by others.”²⁷ A time limit discourages poten-

tial postpublication peer review; potential correspondents will surely be deterred by the unambiguous cutoff. Journals with such a policy should reconsider.

A few journals (eg, *BMJ* and the *CMAJ*) have rapid publication of correspondence on their Web pages. All (or most) letters are published, and there is no apparent time limit. Nor is there the same limit on length as in print journals (Table 2). Electronic letters are linked to the original publication and are relatively easily accessed. It is remarkable and disappointing that as yet so few journals have such a capability. Restricting the facility to current subscribers, as currently done by *Neurology* and *Pediatrics*, is inadequate. A weakness yet to be resolved is the absence of pressure on authors to respond to criticisms.²⁸ For such journals there is uncertainty about which version is definitive. Although the *BMJ* considers *bmj.com* to be the definitive version,²⁹ only the letters that appear in the print journal are indexed on MEDLINE.

What Journal Editors Can Do

Authors and editors should have the same goals: the advancement of scientific understanding and improvement in the treatment and prevention of disease. Poor research is the fault of authors, not journals. Poor research methods, unnecessary research, redundant or duplicate publication, thinly sliced study results, selective reporting, and scientific fraud, as well as a general tendency to inflate the importance of the results, should all be resisted vigorously. All could be less likely if research were not a career necessity for physicians.

Rather than abandon peer review, as some have suggested, journals should work to strengthen it. In particular, methodological review should be implemented much more widely. It will never be possible to eliminate misleading studies, but our imperfect peer-review system is a safeguard without which the quality of published research would be lower.

Journals can also help improve the literature by requiring the full and transparent reporting of research. Guidelines have been developed for RCTs,³⁰ systematic reviews and meta-analyses of RCTs³¹ and observational studies,³² and studies of diagnostic tests,³³ and other initiatives are under way. Editors should continue to be involved in the development of reporting recommendations and explicitly require authors to follow them.

Journals can enable and encourage the publication of research protocols.³⁴⁻³⁶ They can use their Web pages to publish extended versions of articles. They should also enable and encourage publication of the raw data used in medical research articles (eg, *Clinical Chemistry* and *Neurology*). If journals are willing to publish data, they should explicitly suggest this possibility to authors.

Acknowledgment: I thank Iain Chalmers, DSc, and a reviewer for helpful suggestions.

REFERENCES

- Altman DG. The scandal of poor medical research. *BMJ*. 1994;308:283-284.
- Altman DG, Schulz KF, Moher D, et al, for the CONSORT Group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134:663-694.
- Mariani L, Marubini E. Content and quality of currently published phase II cancer trials. *J Clin Oncol*. 2000;18:429-436.
- Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA*. 1999;281:1919-1926.
- Egger M, Davey Smith G, Altman DG, eds. *Systematic Reviews in Health Care: Meta-analysis in Context*. 2nd ed. London, England: BMJ Books; 2001.
- Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42-46.
- Hotopf M, Lewis G, Normand C. Putting trials on trial—the costs and consequences of small trials in depression: a systematic review of methodology. *J Epidemiol Comm Health*. 1997;51:354-358.
- Lawlor DA, Hopker SW. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *BMJ*. 2001;322:763-767.
- Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet*. 2001;358:1340-1342.
- McGuigan SM. The use of statistics in the *British Journal of Psychiatry*. *Br J Psychiatry*. 1995;167:683-688.
- Welch GE II, Gabbe SG. Review of statistics usage in the *American Journal of Obstetrics and Gynecology*. *Am J Obstet Gynecol*. 1996;175:1138-1141.
- Schwarzer G, Vach W, Schumacher M. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Stat Med*. 2000;19:541-561.
- Clarke M, Alderson P, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals. *JAMA*. 2002;287:2799-2801.
- Altman DG, Goodman SN, Schroter S. How statistical expertise is used in medical research. *JAMA*. 2002;287:2817-2820.
- Chanter DO. Maintaining the integrity of the scientific record: new policy is unlikely to give investigators more control over studies [letter]. *BMJ*. 2002;324:169.
- Bland JM, Altman DG. Caveat doctor: a grim tale of medical statistics textbooks. *BMJ*. 1987;295:979.
- Sackett DL, Straus S, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. 2nd ed. Edinburgh, Scotland: Churchill-Livingstone; 2000.
- Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Chicago, Ill: AMA Press; 2002.
- Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA*. 2000;283:2701-2711.
- Westgard JO, Hunt MR. Use and interpretation of common statistical tests in method comparison studies. *Clin Chem*. 1973;19:49-57.
- Rothman KJ. Epidemiologic methods in clinical trials. *Cancer*. 1977;39(4 suppl):1771-1775.
- Oldham PD. A note on the analysis of repeated measurements of the same subjects. *J Chronic Dis*. 1962;15:969-977.
- Donner A, Brown KS, Brasher P. A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979-1989. *Int J Epidemiol*. 1990;19:795-800.
- Goodman SN, Altman DG, George SL. Statistical reviewing policies of medical journals: caveat lector? *J Gen Intern Med*. 1998;13:753-756.
- Gehlbach SH. *Interpreting the Medical Literature: A Clinician's Guide*. 3rd ed. New York, NY: McGraw-Hill; 1993.
- Report of the World Association of Medical Editors (WAME): an agenda for the future. Available at: http://www.wame.org/bellagioreport_1.htm. Accessed January 20, 2002.
- Bland M. Fatigue and psychological distress: statistics are improbable. *BMJ*. 2000;320:515-516.
- Rennie D. Freedom and responsibility in medical publication: setting the balance right. *JAMA*. 1998;280:300-302.
- Smith R. The *BMJ*: moving on. *BMJ*. 2002;324:5-6.
- Moher D, Schulz KF, Altman D, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001;285:1987-1991.
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, for the QUOROM Group. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet*. 1999;354:1896-1900.
- Stroup DF, Berlin JA, Morton SC, et al, for the Meta-analysis of Observational Studies in Epidemiology (MOOSE) Group. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000;283:2008-2012.
- The STARD Group. The STARD initiative: towards complete and accurate reporting of studies on diagnostic accuracy. Available at: <http://www.consort-statement.org/stardstatement.htm>. Accessibility verified May 1, 2002.
- Horton R. Pardonable revisions and protocol reviews [commentary]. *Lancet*. 1997;349:6.
- Chalmers I, Altman DG. How can medical journals help prevent poor medical research? some opportunities presented by electronic publishing. *Lancet*. 1999;353:490-493.
- Godlee F. Publishing study protocols: making them visible will encourage registration, reporting and recruitment [editorial]. *BMC News Views*. 2001;2:4.