Association of Journal Quality Indicators With Methodological Quality of Clinical Research Articles

Kirby P. Lee, MA
Marieka Schotland, BA
Peter Bacchetti, PhD
Lisa A. Bero, PhD

Context The ability to identify scientific journals that publish high-quality research would help clinicians, scientists, and health-policy analysts to select the most up-to-date medical literature to review.

Methods To assess whether journal characteristics of (1) peer-review status, (2) citation rate, (3) impact factor, (4) circulation, (5) manuscript acceptance rate, (6) MEDLINE indexing, and (7) Brandon/Hill Library List indexing are predictors of methodological quality of research articles, we conducted a cross-sectional study of 243 original research articles involving human subjects published in general internal medical journals.

Results The mean (SD) quality score of the 243 articles was 1.37 (0.22). All journals reported a peer-review process and were indexed on MEDLINE. In models that controlled for article type (randomized controlled trial [RCT] or non-RCT), journal citation rate was the most statistically significant predictor (0.051 increase per doubling; 95% confidence interval [CI], 0.037-0.065; $P<.001$). In separate analyses by article type, acceptance rate was the strongest predictor for RCT quality (–0.113 per doubling; 95% CI, –0.148 to –0.078; $P<.001$), while journal citation rate was the most predictive factor for non-RCT quality (0.051 per doubling; 95% CI, 0.044-0.059; $P<.001$).

Conclusions High citation rates, impact factors, and circulation rates, and low manuscript acceptance rates and indexing on Brandon/Hill Library List appear to be predictive of higher methodological quality scores for journal articles.

JAMA. 2002;287:2805-2808 www.jama.com

IT IS DIFFICULT FOR CLINICIANS, scientists, and health policy analysts to keep up with the more than 2 million new research articles published each year in medical and scientific journals.[1] Furthermore, many published reports are of poor-to-average methodological quality,[2-6] and most scientific articles are never cited.[7,8]

One approach to facilitating identification of sound medical evidence is to identify high-quality journals that are likely to publish high-quality research. Peer-review and bibliometric methods (such as journal citation rates, impact factors, circulation, manuscript acceptance rates, and indexing on MEDLINE or Brandon/Hill Library List) may be useful in evaluating the quality of a journal.[9-14] However, these methods are controversial due to potential biases in citation, impact factor, and inherent limitations of the sources of information used to calculate them.[8,12,15-28] Currently, none of these bibliometric parameters have been validated as predictors of journal quality.

We determined whether journal characteristics of peer-review status, citation rate, impact factor, circulation, manuscript acceptance rate, and indexing on MEDLINE or the Brandon/Hill Library List are associated with the methodological quality of original research articles they publish. Studies have also suggested that the source of research funding is associated with article quality.[3,4,6,29-31] Therefore, we also estimated the effect of funding source on article quality score.

METHODS

Selection of Journals and Articles

Using a computer-generated list of random numbers, we randomly selected 30 journals from 107 categorized as general internal medical journals by the Institute for Scientific Information.[32] We excluded journals that were not in English or were unavailable through the University of California library system. Original research articles published in the journals were identified by searching MEDLINE and HealthSTAR from January 1, 1999, through December 31, 1999, using exact journal title, human subjects only, and publication type (journal article). We excluded reviews, historical articles, meta-analyses, case reports or case series, clinical conferences, comments, and consensus development conferences because they would require a different instrument for quality assessment. For each journal, we initially randomly sampled 3 randomized controlled trial (RCT) articles (or all, if <3 were published) and 3 other (non-RCT) ar-

Author Affiliations: Institute for Health Policy Studies, School of Medicine and Department of Clinical Pharmacy, School of Pharmacy (Mr Lee, Ms Schotland, and Dr Bero), and Department of Epidemiology and Biostatistics (Dr Bacchetti), University of California, San Francisco. Ms Schotland is now with the Department of Psychology, New York University, New York.
Corresponding Author and Reprints: Lisa A. Bero, PhD, Clinical Pharmacy and Health Policy, University of California, Suite 420, Box 0613, 3333 California St, San Francisco, CA 94143 (e-mail: bero@medicine.ucsf.edu).

ticles. We scored these as described below and examined the article-to-article variability within each article type (RCT or non-RCT) within each journal. We then randomly sampled up to 6 additional articles of a type from the journals with the greatest variability, for a total of 97 RCT and 146 non-RCT articles. This additional sampling from the most variable journals improved the amount of statistical information provided per article scored.

Journal Characteristics and Data Sources

We collected data on the following 7 journal characteristics for each of the journals: (1) peer-review status, defined as manuscript review by a journal's editor(s) and outside experts, was verified by examining the journal's published peer-review policy or contacting the journal's editorial office; (2) citation rate, defined as the average number of times current articles in a specific journal were cited during the year they were published, was obtained from the Institute for Scientific Information[32]; (3) impact factor, defined as the total number of citations during a given year that a journal receives to articles from the 2 previous years, divided by the total number of "source" articles published in the journal during that same 2-year period, was obtained from the Institute for Scientific Information[32]; (4) circulation, defined as the number of subscriptions for a journal publication, was obtained from *Ulrich's International Periodicals Directory*[33] or the journal's editorial office or publisher; (5) manuscript acceptance rate, defined as the percentage acceptance of original research articles in a given year, was obtained from the journal's editorial office or publisher; (6) whether each journal was indexed in MEDLINE in 1999; (7) and whether each journal was indexed on the Brandon/Hill Library List in 1999.[34] (The Brandon/Hill Library List is a selected list of books and journals that are recommended for the small medical library.[34])

Article Quality Assessment

Two reviewers independently assessed the quality of each article using an instrument previously tested for validity and reliability.[5] Our quality assessment instrument includes 22 items designed to measure the methodological quality of articles (defined as the minimization of systematic bias and the consistency of conclusions with results) with a wide range of study designs, regardless of article topic. We selected this instrument rather than an instrument that only assesses the quality of RCTs because our objective was to assess the validity of quality characteristics of journals based on the quality of RCTs and non-RCTs that the journal publishes. We chose the instrument because it compared favorably in terms of validity and reliability with instruments that assess the quality of RCTs only,[35] and it performed similarly to other well-accepted instruments for scoring the quality of trials included in meta-analyses.[36] We did not choose the component approach for assessing quality of research articles, since empirical evidence for the approach applies primarily to RCTs.[37,38]

Reviewers were trained to use the instrument and were given detailed written instructions. Because previous studies suggest that masking of the articles to the reviewers does not influence quality scores,[39,40] reviewers were not blinded to the identity of the articles or the hypothesis of the study, but they did not have access to the data on the journal quality characteristics.

Scores can range continuously from 0 (lowest quality) to 2 (highest quality).[5] The average of the scores of the 2 reviewers was used for the analyses unless the reviewers' scores differed by more than 1 SD. In this case, the article was discussed by both reviewers until consensus was achieved, and the consensus score was used in the analyses. Ten percent of methodological quality scores required adjudication. The interrater reliability of overall scores, which was measured by intraclass correlation,[41] was fair (r=0.45).

Source of Funding

For each article, funding sources were categorized as government, private nonprofit, industry, government plus private nonprofit, industry plus any other source, unable to be determined, or none disclosed.

Statistical Analyses

Quality scores were modeled in terms of journal characteristics and article characteristics, with a random journal effect included to account for within-journal correlation (Mixed Procedure, Version 8.2, SAS Institute Inc, Cary, NC). Because predictor variables had skewed distributions, we used log transformations to prevent large values from being too influential. Because some journals had much more variability in their quality scores than others, models allowed for different variances for each journal.

Logistic regression with a random journal effect was used to model article type and journal characteristics as predictors of whether funding source was disclosed.

RESULTS

Article Quality Scores

The mean (SD) methodological quality score of all articles was 1.37 (0.22) (range, 0.62-1.88) on a scale of 0 to 2. All 30 journals had non-RCT articles and 28 had RCT articles. Non-RCT articles had an estimated average quality score of 1.31, and RCT articles were estimated to average 0.13 points higher (95% confidence interval [CI], 0.10-0.17; $P<.001$).

Predictors of Journal Quality

All journals reported a peer-review process and were indexed on MEDLINE, so these variables could not be analyzed. There were significant associations between quality scores and higher citation rates ($P<.001$; citation rate range, 0.03-6.06), higher impact factors ($P<.001$; impact factor range, 0.22-28.66), higher circulation ($P=.001$; circulation range, 1080-3.7 million), lower manuscript acceptance rates ($P<.001$; manuscript acceptance rate range,

7.5%-72.0%), and indexing on Brandon/Hill Library List ($P<.001$; 33.3% indexed and 66.7% not indexed) (TABLE 1). Residuals and fitted random journal effects from these models appeared to be approximately normally distributed, and there were no extreme outliers.

In models that controlled for article type (RCT or non-RCT), citation rate was the predictor with the smallest *P* value (0.051 increase per doubling, 95% CI, 0.037-0.065; $P<.001$). No other predictors substantially improved this model.

In separate analyses by article type, acceptance rate was the strongest predictor for RCT quality (–0.113 per doubling; 95% CI, –0.148 to –0.078; $P<.001$), while journal citation rate was the most predictive factor for non-RCT quality (0.051 per doubling; 95% CI, 0.044-0.059; $P<.001$). This means that the estimated effect of acceptance rate on the quality of RCTs is that for every doubling of acceptance rate, article quality score decreases by 0.11 point. The estimated effect of citation rate on the quality of non-RCTs is that for every doubling in journal citation rate, article quality score increases by 0.05 point.

Funding Sources

For the entire sample, only 66% (160/243) of articles reported a source of study funding. These studies reported funding sources solely by government (34%; 54/160), private nonprofit (19%; 30/160), and industry (14%; 23/160). We were unable to determine the type of funding for 4 studies (2%). Thirty-one percent (49/160) reported multiple sources of funding, with 53% (26/49) being funded by government and private nonprofit followed by 37% (18/49) by government and industry, 6% (3/49) by government, industry, and private nonprofit, and 4% (2/49) by industry and private nonprofit. The authors of RCTs (77%; 75/97) were more likely to disclose a funding source than authors of non-RCTs (58%; 85/146) ($P=.05$ from random-effects logistic regression).

After controlling for article type (RCT or non-RCT), articles disclosing funding sources were estimated to score 0.022 higher than those without disclosure (95% CI, –0.029 to 0.074; $P=.39$). All of the journal characteristics had statistically significant associations with disclosure. Significance ranged from $P<.001$ for impact factor and citation rate to $P=.006$ for circulation.

TABLE 2 shows the estimated effects of different funding sources on article quality. For RCTs, private nonprofit funding was associated with the highest scores, while industry-funded studies received the lowest scores, particularly those with a mix of industry and other funding. For non-RCTs, the pattern was qualitatively similar, but effects were smaller and did not reach statistical significance.

COMMENT

Articles of higher methodological quality are published in journals whose articles are cited more frequently (higher citation rates and impact factors), read more widely (higher circulation, indexed on the Brandon/Hill Library List), and scrutinized more carefully by editors and outside peer-reviewers (lower manuscript acceptance rates). These 5 journal characteristics may be valid predictors of journal quality when evaluating journals within the same category, such as general internal medicine. Journal citation and manuscript acceptance rates were the best predictors of the quality of research articles published in the journals.

One limitation of our study is that we used a scale, rather than a component method, to assess the quality of research articles.[36,38] However, there is limited empirical data to support the selection of components to measure the quality of non-RCTs that are published by the journals in our study. The quality scores for the articles in this study are slightly higher than previ-

Table 1. Estimated Effects of Journal Characteristics on Article Quality

Journal Characteristic	No. of Journals	Coefficient (95% Confidence Interval)*	*P* Value
Citation rate	30	0.055 (0.037 to 0.073)	<.001
Impact factor	30	0.062 (0.043 to 0.081)	<.001
Circulation rate	29	0.034 (0.017 to 0.051)	<.001
Acceptance rate	25	–0.127 (–0.161 to –0.093)	<.001
Brandon/Hill Library List	30	0.189 (0.100 to 0.279)†	<.001

*The coefficients are the estimated effect on journal quality per doubling of the predictor.
†The coefficient is the estimated effect on journal quality of being listed vs not.

Table 2. Estimated Effects of Types of Funding on Article Quality*

Funding Source	Estimated Effect (95% CI)†	*P* Value
Randomized Controlled Trials (n = 75)‡		
Government only	1.000	
Private nonprofit only	0.042 (0.010 to 0.073)	.01
Government and private nonprofit	–0.048 (–0.178 to 0.082)	.46
Industry only	–0.063 (–0.153 to 0.026)	.16
Industry and any other	–0.183 (–0.218 to –0.148)	<.001
Non–Randomized Controlled Trials (n = 85)		
Government only	1.000	
Private nonprofit only	0.029 (–0.031 to 0.090)	.34
Government and private nonprofit	–0.057 (–0.141 to 0.028)	.18
Industry only	–0.031 (–0.156 to 0.093)	.18
Industry and any other	–0.085 (–0.311 to 0.141)	.45

*Only articles with at least 1 disclosed funding source are analyzed. CI indicates confidence interval.
†Estimated effect is based on the quality score.
‡Industry and any other is estimated to score 0.120 lower than industry only (95% CI, 0.032-0.207; $P = .009$) and 0.135 lower than government and private nonprofit (95% CI, 0.006-0.263; $P = .04$).

ous studies using the same instrument,[3,5,6] which may be the result of improvements in study design and/or the quality of reporting over the last few years. The interrater reliability score in this study is somewhat lower than previous reports using the same instrument,[3,5,6] possibly because of the difficulty in assessing the quality of non-RCT articles, which are more variable in their design.[35]

Our findings that about one third of research articles did not disclose any funding source and that those with disclosed funding sources were of higher quality than those without, suggests that journal editors should continue to encourage disclosure of sources of study funding, affiliations, and other potential financial conflicts of interest.[42,43] Our finding that studies with industry support have lower quality scores than government-funded studies, as well as other studies showing associations between industry-funded sources and outcomes, suggest that funding should be considered when assessing the usefulness of an article.[3,4,6,29-31]

Author Contributions: *Study concept and design:* Lee, Schotland, Bero.
Acquisition of data: Lee, Schotland.
Analysis and interpretation of data: Lee, Schotland, Bacchetti, Bero.
Drafting of the manuscript: Lee, Bacchetti, Bero.
Critical revision of the manuscript for important intellectual content: Lee, Schotland, Bacchetti, Bero.
Statistical expertise: Bacchetti.
Obtained funding: Bero.
Administrative, technical, or material support: Lee, Schotland, Bero.
Study supervision: Lee, Bero.
Funding/Support: This study was funded by the Soros Foundation.
Acknowledgment: We gratefully acknowledge Erin VanScoyoc, BS, for quality assessment, James Lightwood, PhD, for technical assistance, and Julia Kochi for library assistance.

REFERENCES

1. Arndt KA. Information excess in medicine: overview, relevance to dermatology, and strategies for coping. *Arch Dermatol.* 1992;128:1249-1256.
2. MacLehose R, Reeves B, Harvey I, Sheldon T, Russell I, Black A. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technol Assess.* 2000;4:1-154.
3. Barnes D, Bero L. Scientific quality of original research articles on environmental tobacco smoke. *Tob Control.* 1997;6:19-26.
4. Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. *JAMA.* 1998;279:1566-1570.
5. Cho M, Bero L. Instruments for assessing the quality of drug studies published in the medical literature. *JAMA.* 1994;272:101-104.
6. Cho M, Bero L. The quality of drug studies published in symposium proceedings. *JAMA.* 1994;272:101-104.
7. de Jong J, Schaper W. The international rank order of clinical cardiology. *Eur Heart J.* 1996;17:35-42.
8. Seglen P. The skewness of science. *J Am Soc Inform Sci.* 1992;43:628-638.
9. Bloom B, Retbi A, Dahan S, Jonsson E. Evaluation of randomized controlled trials on complementary and alternative medicine. *Int J Technol Assess Health Care.* 2000;16:13-21.
10. Birken CS, Parkin PC. In which journals will pediatricians find the best evidence for clinical practice? *Pediatrics.* 1999;103:941-947.
11. Rennie D. The present state of medical journals. *Lancet.* 1998;SII:18-22.
12. Opthof T. Sense and nonsense about the impact factor. *Cardiovasc Res.* 1997;33:1-7.
13. Schoonbaert D, Roelants G. Citation analysis for measuring the value of scientific publications: quality assessment tool or comedy of errors? *Trop Med Int Health.* 1996;1:739-752.
14. Bruer JT. Methodological rigor and citation frequency in patient compliance literature. *Am J Public Health.* 1982;72:1119-1123.
15. Gehanno J, Thirion B. How to select publications on occupational health: the usefulness of medline and the impact factor. *Occup Environ Med.* 2000;57:706-709.
16. Pittler M, Abbot N, Harkness E, Ernst E. Location bias in controlled clinical trials of complementary/alternative therapies. *J Clin Epidemiol.* 2000;53:485-489.
17. Joyce J, Rabe-Hesketh S, Wessely S. Reviewing the reviews: the example of chronic fatique syndrome. *JAMA.* 1998;280:264-266.
18. Gallagher EJ, Barnaby DP. Evidence of methodologic bias in the derivation of the Science Citation Index impact factor. *Ann Emerg Med.* 1998;31:83-86.
19. Seglen P. Citations and journal impact factors: questionable indicators of research quality. *Allergy.* 1997;52:1050-1056.
20. Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ.* 1997;314:498-502.
21. Moed H, Van Leeuwen T, Reeduck J. A critical analysis of the journal impact factors of *Angewandte Chemie* and *The Journal of the American Chemical Society*: inaccuracies in published impact factors based on overall citations only. *Scientometrics.* 1996;37:105-116.
22. Larsson K. The dissemination of false data through inadequate citation. *J Intern Med.* 1995;238:445-450.
23. Seglen P. Causal relationship between article citedness and journal impact. *J Am Soc Inform Sci.* 1994;45:1-11.
24. Evans JT, Nadjari HI, Burchell SA. Quotational and reference accuracy in surgical journals: a continuing peer review problem. *JAMA.* 1990;263:1353-1354.
25. Campbell F. National bias: a comparison of citation practices by health professionals. *Bull Med Libr Assoc.* 1990;78:376-382.
26. MacRoberts M, MacRoberts B. Problems of citation analysis: a critical review. *J Am Soc Inform Sci.* 1989;40:342-349.
27. Gotzsche PC. Reference bias in reports of drug trials. *BMJ.* 1987;295:654-656.
28. Moravcsik M, Murugesan P. Some results on the function and quality of citations. *Soc Stud Sci.* 1975;5:86-92.
29. Djulbegovic B, Lacevic M, Cantor A, et al. The uncertainty principle and industry-sponsored research. *Lancet.* 2000;356:635-638.
30. Bero L, Rennie D. Influences on the quality of published drug studies. *Int J Technol Assess Health Care.* 1996;12:209-237.
31. Davidson R. Source of funding and outcome of clinical trials. *J Gen Intern Med.* 1986;1:155-158.
32. Institute for Scientific Information. *Science Citation Index: Journal Citation Reports*. Philadelphia, Pa: Institute for Scientific Information; 1998.
33. *Ulrich's International Periodicals Directory*. 38th ed. New Providence, NJ: Bowker; 2000.
34. Hill D. Brandon/Hill selected list of books and journals for the small medical library. *Bull Med Libr Assoc.* 1999;87:145-169.
35. Moher D, Jadad A, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials.* 1995;16:62-73.
36. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA.* 1999;282:1054-1060.
37. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA.* 1999;282:1061-1066.
38. Berlin J, Rennie D. Measuring the quality of trials: the quality of quality scales. *JAMA.* 1999;282:1083-1085.
39. Clark H, Wells G, Huet C, et al. Assessing the quality of randomized trials: reliability of the Jadad scale. *Control Clin Trials.* 1999;20:448-452.
40. Berlin J, Miles C, Crigliano M, et al. Does blinding of readers affect the results of meta-analyses? results of a randomized trial. *Online J Curr Clin Trials.* 2001. Available at: http://90rt.ucsd/newjour/o/ms901600.html.
41. Haggard E. *Intraclass Correlation and the Analysis of Variance*. New York, NY: Dryden Press; 1958.
42. Davidoff F, DeAngelis CD, Drazen JM, et al. Sponsorship, authorship, and accountability. *JAMA.* 2001;286:1232-1234.
43. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med.* 1997;126:36-47.