

review in clinical neuroscience. *Brain*. 2000;123:1964-1969.

15. Caswell A. Letters to the editor 1991. *Med J Aust*. 1992;157:63-64.

16. Bingham C. Peer review on the Internet: a better class of conversation. *Lancet*. 1998;351(suppl 1):10-14.

17. Bingham CM, Higgins G, Coleman R, Van der Weyden MB. The *Medical Journal of Australia* Internet peer-review study. *Lancet*. 1998;352:441-445.

18. McConnell J, Horton R. *Lancet* electronic research archive in international health and eprint server. *Lancet*. 1999;354:2-3.

19. Smith R. Electronic publishing in science. *BMJ*. 2001;322:627-629.

20. Horton R. North and South: bridging the information gap. *Lancet*. 2000;355:2231-2236.

21. Horrobin DF. The philosophical basis of peer review and the suppression of innovation. *JAMA*. 1990; 263:1293-1295.

22. Grant JM. Timely publication of critical letters. *Lancet*. 1999;353:1190.

23. Popper K. *Realism and the Aim of Science*. London, England: Routledge; 1983.

24. Mill JS. *On Liberty*. London, England: Penguin; 1985.

25. Edwards R, White M, Gray J, Fischbacher C. Use of a journal club and letter-writing exercise to teach critical appraisal to medical undergraduates. *Med Educ*. 2001;35:691-694.

Journal Prestige, Publication Bias, and Other Characteristics Associated With Citation of Published Studies in Peer-Reviewed Journals

Michael Callaham, MD

Robert L. Wears, MD, MS

Ellen Weber, MD

ALTHOUGH PUBLICATION IS A crucial portion of the scientific process, an equally important part is the subsequent use and citation of these published articles by other researchers and authors. We studied a cohort of all research submitted to a scientific meeting and subsequently published to determine how these studies were cited by other authors and determine what characteristics (including positive results) were associated with more frequent citation.

METHODS

We previously reported the methods of the first phase of this study.¹ To summarize, all abstracts of scientific studies submitted to the Society for Academic Emergency Medicine (SAEM) meeting in 1991 were examined. Each submitted abstract was categorized in a blinded fashion according to research design, number of subjects, and other characteristics (TABLE 1 and

Context Citation by other authors is important in the dissemination of published science, but factors predicting it are little studied.

Methods To identify characteristics of published research predicting citation in other journals, we searched the Science Citations Index database for a standardized 3.5 years for all citations of published articles originally submitted to a 1991 emergency medicine specialty meeting. Analysis was conducted by classification and regression trees, a non-parametric modeling technique of regression trees, to determine the impact of previously determined characteristics of the full articles on the outcome measures. We calculated the the number of times an article was cited each year and calculated the mean impact factor (citations per manuscript per year) in other citing journals.

Results Of the 493 submitted manuscripts, 204 published articles met entry criteria. The mean citations per year was 2.04 (95% confidence interval, 1.6-2.4; range, 0-20.9) in 440 different journals. Nineteen articles (9.3%) were never cited. The ability to predict the citations per year was weak (pseudo $R^2=0.14$). The strongest predictor of citations per year was the impact factor of the original publishing journal. The presence of a control group, the subjective newsworthiness score, and sample size predicted citation frequency (24.3%, 26.0%, and 26.5% as strongly, respectively). The ability to predict mean impact factor of the citing journals was even weaker (pseudo $R^2=0.09$). The impact factor of the publishing journal was the strongest predictor, followed by the newsworthiness score (89.9% as strongly) and a subjective quality score (61.5%). Positive outcome bias was not evident for either outcome measure.

Conclusion In this cohort of published research, commonly used measures of study methodology and design did not predict the frequency of citations or the importance of citing journals. Positive outcome bias was not evident. The impact factor of the original publishing journal was more important than any other variable, suggesting that the journal in which a study is published may be as important as traditional measures of study quality in ensuring dissemination.

JAMA. 2002;287:2847-2850

www.jama.com

Author Affiliations: Division of Emergency Medicine, University of California, San Francisco (Drs Callaham and Weber), Department of Emergency Medicine, University of Florida at Jacksonville, Jacksonville (Dr Wears).

Corresponding Author and Reprints: Michael Callaham, MD, Box 0208, University of California, San Francisco, CA 94143-0208 (e-mail: mlc@medicine.ucsf.edu).

Table 1. Predictors of Citations per Year*

Variable	Relative Contribution†
Impact factor of original publishing journal	100.0
No. of subjects (quartiles)	26.5
Newsworthiness score (delphi)	26.0
Control group	24.3
Quality score (delphi)	15.8
Accepted for presentation at meeting	5.5
Explicit hypothesis	4.7
Retrospective vs prospective	2.7
Type of subjects	2.1
Blinded	0.7
Randomized	0.0
Positive results	0.0

*Pseudo- $r^2 = 0.14$. Delphi refers to the author delphi ranking score described in "Methods" and in Callahan et al.¹

†The strongest variable is rated 100, with all others measured against it.

Table 2. Predictors of the Mean Impact Factor of Citing Journals*

Variable	Relative Contribution†
Impact factor of original publishing journal	100.0
Newsworthiness score (delphi)	89.9
Quality score (delphi)	61.5
No. of subjects (quartiles)	7.5
Explicit hypothesis	6.6
Accepted for presentation at meeting	1.4
Type of subjects	0.0
Retrospective vs prospective	0.0
Control group	0.0
Blinded	0.0
Randomized	0.0
Positive results	0.0

*Pseudo- $r^2 = .09$. Delphi refers to the author delphi ranking score described in "Methods" and in Callahan et al.¹

†The strongest variable is rated 100, with all others measured against it.

TABLE 2), and rated subjectively for scientific quality and newsworthiness using a modified delphi method. Searches of MEDLINE, EMBASE, and Cochrane databases were conducted for 4 years after the meeting to identify publication in any journal listed in the National Library of Medicine (NLM). If needed, further information was obtained from authors in writing.² The characteristics of submitted abstracts that predicted full publication were reported previously.¹

In March 2000, the cohort was further examined by searching the Science Citation Index (SCI) database for articles of every study submitted to this meeting that had been published in full (<http://www.webofscience.com>).

For each published article, all citations of that article from publication to the time of the search were identified. Citations were minimal until 2 years after publication, so results were analyzed beginning then and for the next 3.5 years.

Two outcome measures were determined. The number of citations per year during the study period was calculated for each article. This can be considered the article's "impact factor," analogous to the traditional impact factor for journals (the annual number of citations per published article in a journal).³ The second outcome measure was the mean citing journal impact factor (CJIF) of each published article, calculated by averaging the impact factors of each journal citing that publication. Together, these 2 outcomes provide estimates of both the quantity and quality of citations of a manuscript.

Summary statistics were calculated using Systat 9.02 statistical software (SPSS, Evanston, Ill). The relationship of the primary outcome measures (citations per year and CJIF) to characteristics of the original abstracts was assessed using regression trees (CART, or Classification and Regression Trees).⁴ Exploratory analyses of the relationships of interest in this study suggested that simpler multiple regression models would not fit the data as well as regression trees, which are particularly suited to complex interactions among predictors.

The analysis was conducted using CART 4.0 (Salford Systems, San Diego, Calif). The candidate predictor variables were impact factor of the publishing journal, number of subjects (in quartiles, 0-33, 34-101, 102-425, and >425), subjects (human, animal, or other), binary variables representing presence or absence of an explicit hypothesis, control group, blinding, acceptance for presentation at the research meeting, method (prospective or retrospective), and positive results (as previously defined¹), and the subjective "newsworthiness" and quality scores, derived from the full article, not the abstract.

The model was allowed to continue splitting until a node contained 20 cases or fewer (about 10% of the sample) and the resultant trees were tested by 10-fold cross-validation analysis to eliminate over-fit trees and to identify the best model. The cross-validation relative error was used as a pseudo- R_2 measure (analogous to but not truly R_2), a global measure of the final model's explanatory power on a new data set (a more detailed description of the methodology is available from the authors.)

This study was approved by the committee on human research of the University of California.

RESULTS

Four hundred ninety-three studies were submitted to the 1991 SAEM meeting from a total of 144 institutions and 103 medical schools, of which 179 (36%) were accepted for presentation at the meeting. Two hundred nineteen (44%) of the 493 studies submitted to this meeting were published in 44 peer-reviewed journals (37 in specialties other than emergency medicine) with impact factors ranging from 0.23 to 24.5.¹

Fifteen of the published articles appeared in journals that are not SCI citation sources and thus have no SCI impact factor. These were excluded from subsequent analysis, leaving 204 published articles.

Of the 204 included published articles, 19 (9.3%) had no citations during the study period; the remainder were cited a total of 1446 times by 440 different journals with impact factors ranging from 0.01 to 24.5, 434 of them from disciplines other than emergency medicine. Seventy-nine studies (39%) were cited 1 or 2 times only during the study period. The mean citations per year was 2.04 (95% confidence interval [CI], 1.6-2.4; range, 0-20.9) and the mean impact factor of the citing journals was 1.69 (95% CI, 1.50-1.87; range, 0.01-9.9). Citing journals included all the large general medical journals; 18% of citing journals had an impact factor greater than 3, compared with only 10% of all journals in the SCI.⁵

Univariate analysis is not reported for purposes of brevity. CART regression for citations per year yielded a pseudo- R^2 statistic of 0.14, suggesting relatively low explanatory power. The regression tree suggested that impact factor was the only variable of importance (Table 1); other characteristics of the studies either had no influence, or were almost completely subsumed by the impact factor of the publishing journal. After adjustment for the impact factor of the publishing journal, presence of a control group, the subjective “newsworthiness” score, and sample size were the next most important determinants of citation (Table 1).

The regression analysis for CJIF yielded a pseudo- R^2 statistic of 0.09, also suggesting little predictive ability. The impact factor of the publishing journal was the most important variable here also; however, in contrast to citations per year, the newsworthiness and quality scores also contributed substantially (Table 2).

Positive outcome bias was not evident in this sample in either univariate analysis or the regression model. Full manuscripts with negative outcomes had a mean 1.96 citations per year (95% CI, 1.2-2.7), whereas those with positive outcomes had 1.84 citations per year (95% CI, 1.2-2.5). Results for CJIF were similar.

COMMENT

The publication of research in peer reviewed journals is only an intermediate outcome, satisfying to authors but not necessarily useful to others. There is no way to measure how useful a published article is to clinicians, but we can measure its impact on other authors by how frequently they cite it in their publications. Such citations are a hallmark of academic achievement for authors and for journals, correlate highly with the opinion of peers as to a scientist's contribution to his/her field and are used by medical school deans for promotion reviews.^{6,7} Citations complete the chain of publication and that underpins the evolution of scientific knowledge.

Of the 204 publications studied, 185 were cited a total of 1446 times by 440 different journals. The average article was cited 2.04 times per year, and the mean citing journal impact factor was 1.84. This is roughly comparable to the citation rate of all material published in all journals in the SCI,⁸ and the proportion of citing journals with high impact factors in our sample was greater than that of the SCI as a whole.⁵

Only 19 (9.3%) of these studies were never cited during the study period (even by their own authors), compared with 22% for all the international medical literature,⁹ 11% of AIDS articles, and 15% of biology articles.¹⁰ These figures suggest that the study cohort was broadly representative of the biomedical literature.

We found that the impact factor of the original publishing journal, not the methodology or quality of the research, was the strongest predictor of citations per year. A pessimist would suggest that despite the era of accessibility due to electronic searching and retrieval, citation may be more strongly influenced by the reputation of the publishing journal than by the design merits of the study. Thus, a strong or seminal paper submitted to a minor journal might not receive the scientific recognition it deserves. Likewise, a weak article published in a major journal will probably receive more recognition than it deserves.

An optimist might interpret our results to mean that journals are perfectly efficient in publishing studies of a uniform quality, with the same citation value as the journal itself. However, the SCI reports marked variability in citation rates of individual articles in any given journal,⁵ and we found no relationship between study design (and other measures of quality), and the impact factor of the original publishing journal.¹

Once the impact factor of the publishing journal was accounted for, the subjective newsworthiness score (from a delphi panel rating), sample size and presence of a control group were the only major predictors of citations (and

whose respective contribution was only 26.0%, 26.5%, and 24.3% as strongly). This is disappointing since one would hope that characteristics of sound design would be the most important predictors. However, it may also indicate that quality checklists for assessing studies are not meaningful or accurate^{11,12} or that editors are able to identify importance and originality in research independent of the “plumbing” of study design.

It is encouraging that studies with positive results were not cited more frequently or by more prestigious journals. This is in contrast to the consistent bias toward acceptance and publication of studies with positive results previously reported for this study cohort and others.^{1,13-20} Although documented at earlier steps of the publication process (from submission to research meetings to publication of full articles), the final step, citation by subsequent authors, appears to be relatively free of this bias.

In this study, acceptance of research for presentation at the meeting had no predictive power for citation frequency or impact factor of citing journals. As with our previous finding that acceptance failed to predict publication of a manuscript or the impact factor of the publishing journal,¹ the current results suggest methods of screening research for meeting presentation are very fallible.

Our study has several limitations. It examined research (mostly clinical) from only 1 meeting although this cohort has previously been shown to share key characteristics with research from other specialty meetings¹ and was published in 44 journals covering a broad spectrum of specialties and impact factors. Eighty-four percent of publishing journals and 98% of the citing journals were from specialties other than that sponsoring the meeting. Citations per year is not an ideal tool for detecting positive outcome bias because a citation may not credit or praise a study but instead refute or criticize it. However, most citations in most articles do not refute the cited paper, and

authors, journals, and promotion committees consider the number of citations to constitute a measure of merit. Finally, the ability of the regression model to predict outcomes was poor, meaning that most of the variance was accounted for by unmeasured variables (such as, hopefully, the relevance of the study to the citing research).

Author Contributions: *Study concept and design:* Callaham, Wears.
Acquisition of data: Callaham, Wears, Weber.
Analysis and interpretation of data: Wears.
Drafting of the manuscript: Callaham.
Critical revision of the manuscript for important intellectual content: Callaham, Wears, Weber.
Statistical expertise: Wears.
Administrative, technical, or material support: Wears.
Study supervision: Callaham.

REFERENCES

- Callaham ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting [published erratum appears in *JAMA*. 1998;280:1232]. *JAMA*. 1998;280:254-257.
- Weber EJ, Callaham ML, Wears RL, Barton C, Young G. Unpublished research from a medical specialty meeting: why investigators fail to publish. *JAMA*. 1998;280:257-259.
- Garfield E. *SCI Journal Citation Reports: A Bibliometric Analysis of Science Journals in the ISI Database*. Philadelphia, Pa: Institute for Science Information Inc; 1996.
- Breiman L, Friedman J, Olshen R, Stone C. *Classification and Regression Trees*. New York, NY: Chapman & Hall; 1984.
- Garfield E. How can impact factors be improved? *BMJ*. 1996;313:411-413.
- Garfield E. From citation indexes to informetrics: is the tail now wagging the dog? *Libri*. 1998;48:67-80.
- Davies HD, Langley JM, Speert DP, for the Pediatric Investigators' Collaborative Network on Infections in Canada. Rating authors' contributions to collaborative research: the PICNIC study of university departments of pediatrics. *CMAJ*. 1996;155:877-882.
- Garfield E. Random thoughts on citationology: its theory and practice. *Scientometrics*. 1998;43:69-76.
- Schwartz C. The rise and fall of uncitedness. *College Res Libr*. 1997;48:19-29.
- Brown P. Has the AIDS research epidemic spread too far? *New Sci*. 1993;15:12-15.
- Berlin JA, Rennie D. Measuring the quality of trials: the quality of quality scales. *JAMA*. 1999;282:1083-1085.
- Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282:1054-1060.
- Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results: follow-up of applications submitted to two institutional review boards. *JAMA*. 1992;267:374-378.
- Dickersin K, Min YI. NIH clinical trials and publication bias. *Online J Curr Clin Trials* [serial online]. 1993;doc 50.
- Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337:867-872.
- Goldman L, Loscalzo A. Fate of cardiology research originally published in abstract form. *N Engl J Med*. 1980;303:255-259.
- Meranze J, Ellison N, Greenhow D. Publications resulting from anesthesia meeting abstracts. *Anesth Analg*. 1982;61:445-48.
- McCormick MC, Holmes JH. Publication of research presented at the pediatric meetings. *Am J Dis Child*. 1985;139:122-126.
- Scherer RW, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts: a meta-analysis. *JAMA*. 1994;272:158-162.
- Koren G, Graham K, Shear H, Einerson T. Bias against the null hypothesis: the reproductive hazards of cocaine. *Lancet*. 1989;2:1440-1444.

CME ANNOUNCEMENT

CME Hiatus: July Through December 2002

CME from *JAMA/Archives Journals* will be suspended between July and December 2002. Beginning in early 2003, we will offer a new *online* CME program that will provide many enhancements:

- Article-specific questions
- Hypertext links from questions to the relevant content
- Online CME questionnaire
- Printable CME certificates and ability to access total CME credits

We apologize for the interruption in CME and hope that you will enjoy the improved online features that will be available in early 2003.