Relative Citation Impact of Various Study Designs in the Health Sciences

Nikolaos A. Patsopoulos, MD
Apostolos A. Analatos
John P. A. Ioannidis, MD

Context The relative merits of various study designs and their placement in hierarchies of evidence are often discussed. However, there is limited knowledge about the relative citation impact of articles using various study designs.

Objective To determine whether the type of study design affects the rate of citation in subsequent articles.

Design and Setting We measured the citation impact of articles using various study designs—including meta-analyses, randomized controlled trials, cohort studies, case-control studies, case reports, nonsystematic reviews, and decision analysis or cost-effectiveness analysis—published in 1991 and in 2001 for a sample of 2646 articles.

Main Outcome Measure The citation count through the end of the second year after the year of publication and the total received citations.

Results Meta-analyses received more citations than any other study design both in 1991 ($P<.05$ for all comparisons) and in 2001 ($P<.001$ for all comparisons) and both in the first 2 years and in the longer term. More than 10 citations in the first 2 years were received by 32.4% of meta-analyses published in 1991 and 43.6% of meta-analyses published in 2001. Randomized controlled trials did not differ significantly from epidemiological studies and nonsystematic review articles in 1991 but clearly became the second-cited study design in 2001. Epidemiological studies, nonsystematic review articles, and decision and cost-effectiveness analyses had relatively similar impact; case reports received negligible citations. Meta-analyses were cited significantly more often than all other designs after adjusting for year of publication, high journal impact factor, and country of origin. When limited to studies addressing treatment effects, meta-analyses received more citations than randomized trials.

Conclusion Overall, the citation impact of various study designs is commensurate with most proposed hierarchies of evidence.

JAMA. 2005;293:2362-2366 www.jama.com

SEVERAL AUTHORS AND ORGANIzations have proposed hierarchies of evidence, based on the relative reliability of various types of study designs.[1-4] Although many people recognize that expert opinions and nonsystematic reviews provide the least reliable level of information,[5,6] such articles continue to have a massive influential presence.[7] Controlled studies assume higher places in hierarchies of evidence than uncontrolled studies, and randomized trials are considered the gold standard for clinical research.[1-4] However, randomized trials cannot be conducted for all questions of interest[8] and there is debate on whether they give different results than nonrandomized studies.[9-14] Finally, meta-analyses are becoming increasingly frequent in the literature. Meta-analyses are often placed at the highest level of evidence,[1-4] despite their critics.[15,16] No hierarchy of evidence is unanimously accepted.

An important issue is whether the impact of various studies is different and changing over time. Impact on clinical practice and decision making is difficult to measure comprehensively. However, one important measure of impact is the use of citations in the published literature. Citations have limitations,[17] but they provide an objective measurement of how often scientists use a specific published work. One may ask: What is the relative citation impact of published articles using various types of designs? Is this impact commensurate with the proposed hierarchies of evidence? Has it changed over time? We aimed to answer these questions using citation analysis.

METHODS

Identification and Eligibility of Relevant Studies

We compared the citation impact across various study designs and between studies published in 1991 and 2001. We searched the Institute for Scientific Information (ISI) Science Citation Index at the Web of Science Database (www.isinet.com) for meta-analyses, randomized controlled trials (RCTs), cohort studies, case-control studies, case reports, nonsystematic reviews, and decision analysis or cost-effectiveness analysis records published in 1991 and 2001. These types of publications cover the major, readily identifiable designs used in collecting

Author Affiliations: Clinical and Molecular Epidemiology Unit and Clinical Trials and Evidence Based Medicine Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece (Drs Patsopoulos and Ioannidis and Mr Analatos); Department of Immunology and Histocompatibility, Faculty of Medicine, University of Thessaly, Larissa, Greece (Dr Patsopoulos and Mr Analatos); Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Mass (Dr Ioannidis).
Corresponding Author: John P. A. Ioannidis, MD, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece (jioannid@cc.uoi.gr).

and synthesizing medical information. Secondarily, meta-analyses were also classified as meta-analyses including RCTs vs others. Both meta-analyses and RCTs were also classified according to their subject or purpose (treatment effect [therapy or prevention], prognosis, diagnosis, and etiology or association for meta-analyses; treatment effect and diagnosis for RCTs).

It is impractical to identify and analyze all the tens of thousands of publications fitting in these study designs. Often it is impossible to accurately classify the study design unless the whole article is carefully scrutinized. Sometimes even this may not suffice. Thus we used a strategy that aimed to yield an adequate number of relevant publications for each design with high specificity in characterizing design. The search strategies for each type of publication sought the appearance of the relevant study design terms in the article title (TI). Meta-analysis was searched with *TI=meta-analy** or *metaanaly**, randomized controlled trial with *TI=random* AND TI=trial*, decision analysis or cost-effectiveness analysis with *TI=decision analy* OR TI=cost effectiveness analy* OR TI=cost-effectiveness analy**, nonsystematic review with *TI=review NOT TI=systemat* NOT TI=meta-analy* NOT TI=overview NOT TI=case report**, case-control study with *TI=case control study*, cohort study with *TI=cohort study*, and case report with *TI=case report NOT TI=review NOT TI=overview*. When the search algorithm yielded an excessive number of articles, we screened systematically 1:5 or 1:10 batches of records, for study designs with 1200 to 3000 records and more than 3000 records retrieved in a year, respectively.

Two investigators (N.A.P. and A.A.A) independently screened both the title and abstract of identified articles. Articles were eligible if they represented applications of the type of study design under which they were identified. We excluded ISI records without abstract; letters; editorials; news and meeting abstracts; methodology-and-theory articles; and articles not on human subjects or material, not on health, or both. Discrepancies were discussed between investigators; a third investigator (J.P.A.I.) resolved disagreements.

Data Extraction

For each article eligible for citation analysis, we recorded total citations until December 10, 2004; citations received up to the end of the second year after publication (1991-1993 and 2001-2003, respectively); country of authors; and journal.

Analysis

The main analyses addressed citation counts for 1991-1993 and 2001-2003 (early citations). Most articles are rarely cited, if at all, during the same year in which they were published, but the citation count of the 2 subsequent years is representative (it forms the basis of estimating journal impact factors). Secondary analyses counted total citations until December 10, 2004 (long-term impact); this time frame unavoidably differed between the 1991 and 2001 publication cohorts.

Citation counts per publication type and year were summarized with medians and interquartile ranges (citation distributions are left-skewed).[18] Mann-Whitney U tests and Kruskal-Wallis analysis of variance compared 2 or several groups, respectively.

We also identified articles that received more than 10 citations in the first 2 years (approximately the top 10% most-cited ISI-indexed articles in Clinical Medicine).[19] Logistic regressions addressed the year and type of publication (dummy variables) as predictors of more than 10 citations in 2 years, adjusting also for country of authors and high journal impact factor.

Analyses were conducted using SPSS statistical software version 12.0 (SPSS Inc, Chicago, Ill). *P* values are 2-tailed. Statistical significance was considered at the .05 level.

RESULTS

We identified 17 813 articles (6052 from 1991, 11 761 from 2001) and screened 5769 of those for eligibility (1936 and 3833, respectively); 2646 articles (904 and 1742, respectively) were eligible for citation counting (TABLE 1).

Early Citations

Both in 1991 and in 2001, there was a significant difference in citation counts between various study designs ($P<.001$). Citation counts were statistically significantly higher in the 2001 publications compared with the 1991 publications for all designs ($P<.05$), except for cohort studies and decision and cost-effectiveness analyses (Table 1).

Both in 1991 and 2001, meta-analyses received the highest number of citations and RCTs were second (Table 1). For 1991, comparisons against other designs were always formally statistically significant, with the exception of decision and cost-effectiveness analysis publications ($P=.11$), and the difference against RCTs was also modest ($P=.04$); for 2001, *P* values were $<.001$ for meta-analyses compared with any other design. Differences in citation counts for other designs were more subtle, except for the case reports that always had negligible citation impact.

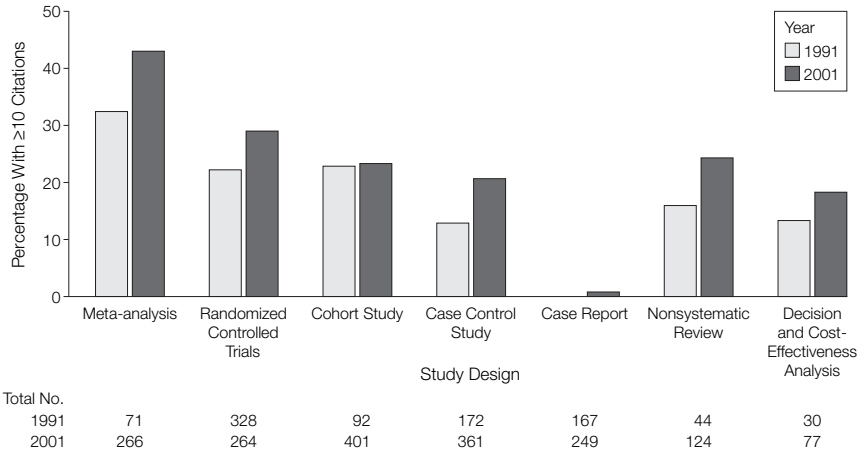
Twenty-three meta-analyses (32.4%) published in 1991 received more than 10 citations within 2 years. This rose to 43.6% (116/266) for meta-analyses published in 2001. Randomized controlled trials had the next highest impact (23.2% [76/328] and 29.5% [78/264], respectively). Other designs had percentages in the range of 10% to 25%, except for case reports for which less than 1% received more than 10 citations (FIGURE 1). In multivariable logistic regression, articles published in 2001 (odds ratio [OR], 1.56; 95% confidence interval [CI], 1.23-1.99), having US authors (OR, 1.69; 95% CI, 1.37-2.08), and published in journals with impact factors greater than 10 (OR, 12.8; 95% CI, 8.4-19.5) were more likely to be cited more than 10 times than articles published in 1991, without US authors, or in other journals; and meta-analyses were significantly more likely to be cited more than 10 times compared with all other designs (for

Table 1. Screened and Eligible Studies and Citation Counts Per Study Design and Year of Publication*

Study Design	Eligible (Screened) Articles		Articles, No. (%) Including US Authors		In Journals With High Impact Factor†		2-Year Citation Count, Median (Interquartile Range)		*P* Value‡
	1991	2001	1991	2001	1991	2001	1991	2001	
Meta-analysis	71 (171)	266 (608)	40 (56)	135 (51)	7 (10)	19 (7)	5 (2-15)	9 (4-20)	.004
Randomized controlled trials	328 (513)	264 (420)	127 (39)	85 (32)	35 (11)	39 (15)	4 (1-10)	6 (2-14)	.001
Cohort	92 (224)	401 (1023)	45 (49)	131 (33)	5 (5)	18 (4)	3 (1-9)	5 (2-10)	.14
Case-control	172 (302)	361 (659)	57 (33)	118 (33)	5 (3)	13 (4)	3 (1-7)	4 (2-9)	.02
Case report	167 (320)	249 (463)	44 (26)	57 (23)	0 (0)	1 (0)	0 (0-1)	1 (0-1)	<.001
Nonsystematic review	44 (310)	124 (480)	18 (41)	58 (47)	1 (2)	1 (1)	2 (0-7)	4 (2-10)	.005
Decision or cost-effectiveness analysis	30 (96)	77 (180)	22 (73)	39 (51)	3 (10)	1 (1)	4 (1-8)	4 (2-10)	.52

*The total number of retrieved articles in the case report search was 1605 in 1991 and 2343 in 2001 and the total number of retrieved articles in the nonsystematic review search was 3141 in 1991 and 4811 in 2001 while the total number of retrieved articles in the randomized controlled trial search was 2139 in 2001. For these study designs and years, systematically 1 in 5 or 1 in 10 batches of records were screened (see "Methods" section for more details). The κ coefficient between the 2 investigators for eligibility of the screened articles was 0.94. Final reasons for exclusions included: no abstract (n = 352), letter (n = 239), editorial (n = 159), news or meeting abstract (n = 1082), methods or theory article (n = 277), not a practical application of the study design being searched (n = 720), not on human health (n = 294).
†Journal impact factor exceeding 10 according to the latest ratings (Institute for Scientific Information, Journal Citation Reports 2003).
‡For the comparison of the 1991 vs 2001 publication cohort in terms of the 2-year citation count.

Figure 1. Percentage of Articles by Study Design With at Least 10 Citations in First 2 Years of Publication



RCTs, OR, 0.49; 95% CI, 0.36-0.68; for cohort studies, OR, 0.46; 95% CI, 0.34-0.63; for case-control studies, OR, 0.37; 95% CI, 0.27-0.52; for case reports, OR, 0.01; 95% CI, 0.00-0.04; for nonsystematic reviews, OR, 0.47; 95% CI, 0.31-0.73; and for decision or cost-effectiveness analysis articles, OR, 0.29; 95% CI, 0.16-0.51 vs meta-analyses).

Long-term Impact

Both in 1991 and in 2001, there was a statistically significant difference in citation count between various designs (*P*<.001, FIGURE 2). For 1991, meta-analyses were statistically significantly cited more times than all other designs (*P*<.05 for all comparisons). Conversely, RCTs had significantly more citations only from case reports (*P*<.001) and possibly decision or cost-effectiveness analysis articles (*P*=.05) but did not differ significantly in citation impact vs other designs. Case reports were statistically significantly less cited than anything else (*P*<.001 for all comparisons). For 2001, meta-analyses had greater impact than all other designs (*P*<.001 for all comparisons) and RCTs were cited significantly more times than all the remaining designs (*P*<.05 for all comparisons). Case reports had once again a very low impact (*P*<.001 for all comparisons). All other comparisons of designs were not statistically significant.

Subgroups

Citations of subgroups of meta-analyses and RCTs are shown in TABLE 2. There were no statistically significant differences in the citations received by meta-analyses including or not including RCTs, both in 1991 and 2001 and both for the 2 years and for the long term (*P*>.19 for all analyses). Similarly, citations did not differ significantly for meta-analyses of different purpose or subject (*P*>.58 for all analyses). Meta-analyses addressing treatment effects tended to receive more citations than RCTs of treatment effects in 1991 (*P*=.08 for 2-year citations, *P*=.10 for long-term citations), and the difference became more clear in 2001 (*P*=.001 and *P*=.001, respectively).

COMMENT

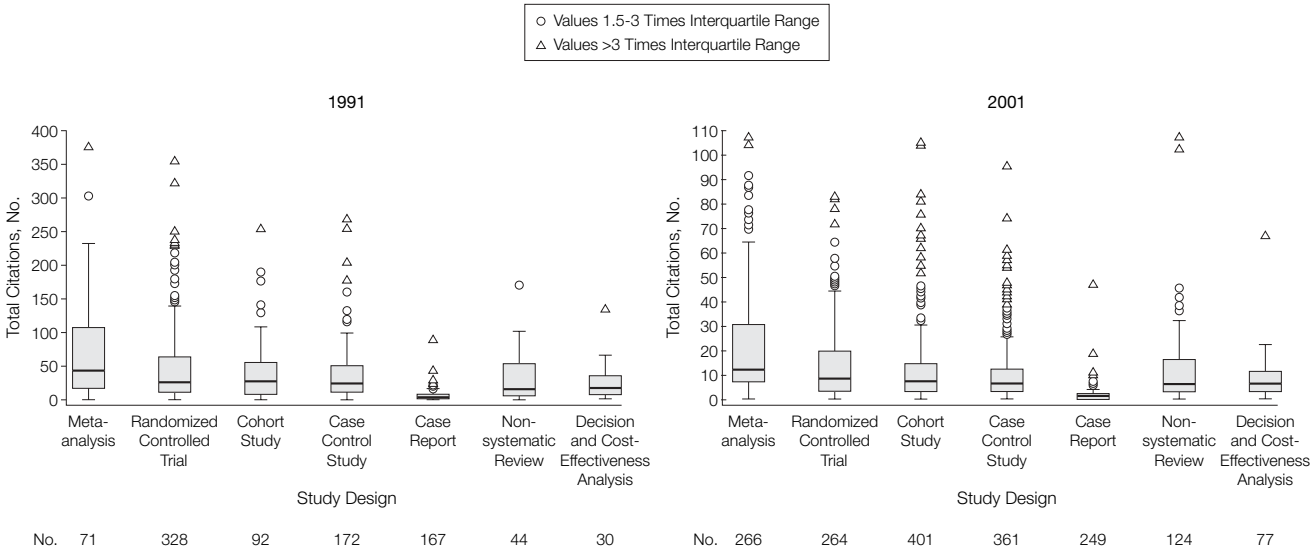
The citation impact of various study designs follows the order proposed by most current theoretical hierarchies of evidence.[1-3] On average meta-analyses currently receive more citations than any other type of study design. Meta-analyses have clearly surpassed in citation impact both decision or cost-effectiveness analysis articles and RCTs,

against which they had mostly modest differences, if any, in the early 1990s. Although RCTs have become the second most cited study design, decision or cost-effectiveness analysis has not followed this growth. Epidemiological studies are now lagging behind randomized research; however, this was not as evident for articles published in 1991. Nonsystematic reviews continue to have a citation impact similar to that of epidemiological studies. Finally, case reports have negligible impact.

The superiority in citation impact of meta-analyses and secondarily RCTs is consistent with the prominent role given to these designs by evidence-based medicine,[1-4] despite the criticisms leveled against both designs.[15,20] The further dissemination of hierarchies of evidence may further increase the citations for meta-analyses and RCTs. If the proposal that each study should start and end when a meta-analysis is adopted,[21] meta-analyses may become even more highly cited. Interestingly, high citations for meta-analyses extend to meta-analyses of nonrandomized research. Of course, we acknowledge that primary studies are required for quantitative syntheses ever to be performed.

The relative impact of epidemiological research has lost ground recently.

Figure 2. Box Plots for Total Citation Counts for Various Study Designs for Articles Published in 1991 and 2001



Box length and error bars represent the interquartile range and 1.5 times the interquartile range, respectively. Outliers (high values extending beyond 1.5 times and up to 3 times the interquartile range) are shown by circles and extremes (high values extending beyond 3 times the interquartile range) are shown by triangles. There were 25 articles with citation counts outside the depicted range (3 randomized controlled trials and 1 review in 1991; 7 meta-analyses, 10 randomized controlled trials, 3 cohort studies, 1 review in 2001). The thick lines in the boxes represent medians.

Table 2. Citations of Subgroups of Meta-analyses and Randomized Controlled Trials

	No. (%)		Median (Interquartile Range)			
			Two-Year Citations		Long-term Citations	
	1991	2001	1991	2001	1991	2001
Meta-analyses						
Per inclusion of randomized controlled trials						
Yes	39 (55)	146 (55)	5 (1-13)	9 (5-23)	34 (16-91)	12 (7-32)
No	32 (45)	120 (45)	6 (3-17)	8 (4-18)	62 (16-119)	12 (6-26)
Per subject or purpose*						
Treatment effect	46 (65)	164 (62)	6 (2-15)	9 (4-21)	40 (18-104)	12 (7-31)
Prognosis	4 (6)	18 (7)	10 (2-24)	13 (6-22)	115 (25-187)	17 (10-33)
Diagnosis	1 (1)	22 (8)	2	10 (4-22)	22	13 (4-30)
Etiology or association	18 (25)	62 (23)	3 (3-8)	8 (4-17)	40 (13-101)	14 (7-30)
Randomized Trials						
Treatment effect	323 (98)	261 (99)	4 (1-10)	6 (2-14)	26 (11-64)	9 (3-20)
Diagnosis†	5 (2)	3 (1)	1 (1-9)	3 (0-3)	13 (9-38)	4 (0-4)

*Two meta-analyses on the physiological range of blood pressure published in 1991 were not counted in any category.
†Clinical trials randomizing participants to 2 diagnostic methods and comparing accuracy.

Perhaps there is increasing uncertainty due to the refutation of several key cohort studies on important questions such as vitamins or hormone therapy.[11] Decision or cost-effectiveness analysis has also not managed to keep a high impact. Nevertheless, many important questions simply cannot be answered with randomized research.

Also many nonsystematic reviews continue to be published. In our study, we excluded nonquantitative reviews that seemed to use some systematic approaches. Empirical evaluations of orthopedic and general medical journals have shown that systematic reviews received double the number of citations compared with nonsystematic ones.[22,23] Efforts to enhance the accuracy and usefulness of all reviews are important because even nonsystematic expert reviews are still extensively read by practitioners.[24]

Some caveats should be discussed. First, higher citation rates in articles published in 2001 than in those published in 1991 probably reflect simply the increase of journals worldwide (especially journal articles listed by ISI). Second, we excluded several types of reports such as nonhuman studies and hybrid designs (eg, reports describing both cohort and case-control studies). However, we wanted to focus sharply on the key study designs. Third, we did not exclude self-citations. Fourth, we used very strict screening criteria to ensure high specificity for characterizing study designs. Most studies probably still do not mention their design in their title. It is unknown whether among studies having the same design, those that state it in the title would get more citations or less. Nevertheless, even if such differences exist, they probably would not affect selectively some study designs over others.

Finally, a citation does not guarantee the respect of the citing investigators. Occasionally a study may be cited only to be criticized or dismissed. Nevertheless, citation still means that the study is active in the scientific debate. Moreover, we should acknowledge that citation impact does not necessarily translate into clinical or scientific impact, but this is extremely difficult to measure and could vary on a case-by-case basis. Allowing for these caveats, our evaluation provides empirical evidence on the relative impact of various study designs in the literature.

Author Contributions: Dr Ioannidis had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Ioannidis, Patsopoulos.
Acquisition of data: Ioannidis, Patsopoulos, Analatos.
Analysis and interpretation of data: Ioannidis, Patsopoulos, Analatos.
Drafting of the manuscript: Ioannidis, Patsopoulos.
Critical revision of the manuscript for important intellectual content: Analatos.
Statistical analysis: Ioannidis, Patsopoulos, Analatos.
Study supervision: Ioannidis.
Financial Disclosures: None reported.
Acknowledgment: We thank Tom Trikalinos, MD, for help in retrieving articles and Anastasios E. Germenis, MD, for encouragement.

REFERENCES

1. West S, King V, Carey TS, et al. *Systems to Rate the Strength of Scientific Evidence.* Rockville, Md: Agency for Healthcare Research and Quality; 2002: 64-88. AHRQ publication 02-E016.
2. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ*. 2001;323:334-336.
3. Phillips B, Ball C, Sackett D, et al. Levels of evidence and grades of recommendations. Oxford, England: Oxford Centre for Evidence-Based Medicine. Available at: http://www.cebm.net/levels_of_evidence.asp. Accessed December 10, 2004.
4. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328:1490.
5. Mulrow CD. The medical review article: state of the science. *Ann Intern Med*. 1987;106:485-488.
6. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA*. 1992;268:240-248.
7. McAlister FA, Clark HD, van Walraven C, et al. The medical review article revisited: has the science improved? *Ann Intern Med*. 1999;131:947-951.
8. Stampfer M. Observational epidemiology is the preferred means of evaluating effects of behavioral and lifestyle modification. *Control Clin Trials*. 1997;18:494-499.
9. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342:1878-1886.
10. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342:1887-1892.
11. Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational vs randomised trial evidence? *Lancet*. 2004;363: 1724-1727.
12. Vandenbroucke JP. When are observational studies as credible as randomised trials? *Lancet*. 2004;363: 1728-1731.
13. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA*. 2001;286:821-830.
14. Ioannidis JP, Haidich AB, Lau J. Any casualties in the clash of randomised and observational evidence? *BMJ*. 2001;322:879-880.
15. Feinstein AR. Meta-analysis: statistical alchemy for the 21st century. *J Clin Epidemiol*. 1995;48:71-79.
16. Shapiro S. Meta-analysis/shmeta-analysis. *Am J Epidemiol*. 1994;140:771-778.
17. Garfield E. How can impact factors be improved? *BMJ*. 1996;313:411-413.
18. Weale AR, Bailey M, Lear PA. The level of noncitation of articles within a journal as a measure of quality: a comparison to the impact factor. *BMC Med Res Methodol*. 2004;4:14.
19. Essential Science Indicators ISI [available on subscription]. Stamford, Conn: Thompson Corp; November 2004. Available at: http://www.isinet.com/products/evaltools/esi. Accessed December 10, 2004.
20. Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. *JAMA*. 1998;279: 319-320.
21. Clarke M, Alderson P, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals. *JAMA*. 2002;287:2799-2801.
22. Bhandari M, Montori VM, Devereaux PJ, Wilczynski NL, Morgan D, Haynes RB; The Hedges Team. Doubling the impact: publication of systematic review articles in orthopaedic journals. *J Bone Joint Surg Am*. 2004;86-A:1012-1016.
23. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Hedges Team. Systematic reviews: a cross-sectional study of location and citation counts. *BMC Med*. 2003;1:2.
24. Loke YK, Derry S. Does anybody read "evidence-based" articles? *BMC Med Res Methodol*. 2003;3:14.

