

Negative Results and Impact Factor

A Lesson From Neonatology

Yoav Littner, MD; Francis B. Mimouni, MD, FAAP, FACN;
Shaul Dollberg, MD, FACN; Dror Mandel, MD

Objective: To test the hypothesis that articles with negative results are more likely than articles with positive results to be published in journals with lower impact factor.

Design and Setting: We selected all of the randomized, placebo-controlled trials conducted during the neonatal period between October 1, 1998, and October 1, 2003. Trials were classified as having positive results or negative results (significant or no significant difference, respectively). Only studies dealing with primary outcomes (efficacy) were included.

Main Outcome Measures: The impact factor of each journal was determined, and the sample size for each study was noted.

Results: There were 233 articles that fulfilled the inclusion criteria. There was a significant difference between the 2 groups in terms of impact factor ($P=.03$) but not sample size ($P=.30$). Impact factor correlated with both sample size and the type of study results (positive results vs negative results; $P<.05$).

Conclusion: Articles with negative results are more likely than articles with positive results to be published in journals with lower impact factor.

Arch Pediatr Adolesc Med. 2005;159:1036-1037

A RECOGNIZED WEAKNESS OF meta-analyses is the so-called "publication bias," ie, the greater likelihood of a study with "positive" results (PR; significant difference between treatment groups) to be published.¹ The impact factor (IF) has been developed as a means of assessing the quality of a journal (ie, the impact it has on the scientific community), is mostly based on the quotation rate of the articles published in it, and may vary from year to year. The higher it is, the more prestigious the journal.² We tested the hypothesis that articles with negative results (NR) are more likely than articles with PR to be published in journals with lower IF.

METHODS

We selected all of the randomized, placebo-controlled trials conducted during the neonatal period (birth to age 1 month) that were published between October 1, 1998, and October 1, 2003, using PubMed to search *placebo*, limiting to *newborn*, *randomized controlled trials*, and *human*. We classified the studies as having PR or NR (significant or no significant dif-

ference, respectively). Only the primary outcomes (efficacy), not secondary outcomes (such as adverse effects), were included. To ensure consistency, only 1 of us (Y.L.) reviewed the articles.

The IF of each journal was determined,³ and the sample size for each study was noted as a potential confounder. We gave an arbitrary classification of 0 for an IF when a journal was not included in the citation index.³

Minitab version 13.1 software (Minitab Inc, State College, Pa) was used for statistical analyses. Since the distribution of IF was not normal and the variance of IF in studies with NR was significantly different ($P<.05$) from that in studies with PR, we used nonparametric tests (Kruskal-Wallis tests) to study the difference in IF and in sample size between groups. Regression analysis was used to study the independent effect of sample size and type of study (PR vs NR) on the ranking of IF. A P value of less than .05 was considered significant.

RESULTS

We identified 286 articles by the search, and 233 of them fulfilled the inclusion criteria. The primary outcome was identified in all of the articles. There was a significant difference between the 2 groups

Author Affiliations:

Department of Neonatology, Lis Maternity Hospital, Tel Aviv Sourasky Medical Center, and the Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

in terms of IF ($P=.03$), but not in terms of sample size ($P=.30$) (**Table**). In ranked multiple regression analysis using the IF rank as the dependent variable, both sample size and the type of study results (findings of PR vs NR) were significant ($P<.05$).

COMMENT

We found that in our field of expertise (newborn medicine), placebo-controlled, prospective clinical trials with NR (ie, no significant difference between control and study groups) are more likely to appear in peer-reviewed medical journals of lower IF than trials with PR (ie, treatment superior to placebo).

There are several possible explanations for this finding. One is related to publication bias. Indeed, a study with NR is less likely to be published than a study with PR.⁴ Although the IF of a journal is only 1 way of measuring its quality (and is a controversial one⁵), the competition for acceptance in high-IF journals is fierce, and it is possible that reviewers and editors are biased toward articles with PR. In a recent study⁶ of articles submitted to *JAMA*, this was not found to be the case, as there was no statistically significant difference in the publication rates of submitted articles with positive vs negative results. This study theoretically indicates that the editorial board of *JAMA* is not likely to be influenced by the PR or NR of a study, but this may not be true for the editorial boards of other journals. The issue of publication bias is greater than just the problem of combining results in meta-analyses. It also directly impacts clinical care, as the published literature will be biased⁷ and the studies with NR are likely to be delayed in their publication.⁸ Callaham et al⁹ and Hartling et al¹⁰ found that, in the field of emergency medicine, positive-outcome bias was present at each of the submission acceptance and publication phases. Furthermore, in the study by Callaham and colleagues, the presentation of an abstract at a scientific meeting and the publication of the complete article were not strongly related to study design or study quality. It is also possible that studies with PR are truly better than those with NR because of better selection of hypotheses, study design, funding, more accurate methods, and so on. Alternatively, it is possible that authors of studies with NR lower their expectations at the time of submission and systematically submit their articles to journals with lower IF (ie, submission bias).¹¹

Interestingly, the sample sizes in articles with NR were, on average, 50% greater than those with PR. This difference was not significant in univariate analysis, but it became significant in multiple regression analysis when the type of the study (PR vs NR) was included. This finding might be related to the fact that when no significant difference is found between groups, authors attempt to increase sample size to gain statistical power, reduce type

Table. Impact Factor and Sample Size in Studies With Positive and Negative Results

	Mean \pm SD	Median	Range	P Value
Impact factor				
Studies with positive results	4.567 \pm 5.878	2.778	0-31.736	.03
Studies with negative results	4.242 \pm 7.153	2.126	0-31.736	
Sample size				
Studies with positive results	421 \pm 1525	80	8-11 619	.30
Studies with negative results	600 \pm 2207	90	12-15 832	

2 error, and increase the likelihood of having their articles accepted for publication. On the other hand, it is also possible that statistical design (eg, sequential analyses) or interim analyses lead to stopping trials with PR early. Alternatively, acceptance bias may favor the publication of trials with small sample sizes only if they have PR. Nevertheless, in the multiple regression analysis, the impact of PR vs NR on the IF of the journal in which the study was published remained significant, even after taking the sample size into account.

Accepted for Publication: April 7, 2005.

Correspondence: Shaul Dollberg, MD, Department of Neonatology, Lis Maternity Hospital, 6, Weitzman St, Tel Aviv, Israel 64239 (dollberg@tasmc.health.gov.il).

REFERENCES

- Simes RJ. Confronting publication bias: a cohort design for meta-analysis. *Stat Med*. 1987;6:11-29.
- Taubes G. Measure for measure in science. *Science*. 1993;260:884-886.
- Thomsom Corp. Science citation index. Available at: <http://isi4.isiknowledge.com/portal.cgi>. Accessed October 1, 2003.
- Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337:867-872.
- Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ*. 1997;314:498-502.
- Olson CM, Rennie D, Cook D, et al. Publication bias in editorial decision making. *JAMA*. 2002;287:2825-2828.
- Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA*. 1990;263:1385-1389.
- Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ*. 1997;315:640-645.
- 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. *JAMA*. 1998;280:254-257.
- Hartling L, Craig WR, Russell K, Stevens K, Klassen TP. Factors influencing the publication of randomized controlled trials in child health research. *Arch Pediatr Adolesc Med*. 2004;158:983-987.
- Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results: follow-up of applications submitted to 2 institutional review boards. *JAMA*. 1992;267:374-378.