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
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. Although the IF of a journal is only 1 way of measuring its quality (and is a controversial one), 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 study 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 biased and the studies with NR are likely to be delayed in their publication. Callaham et al and Hartling et al 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).

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 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

5. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ. 1997;314:498-502.

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

<table>
<thead>
<tr>
<th>Impact factor</th>
<th>Mean ± SD</th>
<th>Median</th>
<th>Range</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies with positive results</td>
<td>4.567 ± 5.878</td>
<td>2.778</td>
<td>0-31.736</td>
<td>.03</td>
</tr>
<tr>
<td>Studies with negative results</td>
<td>4.242 ± 7.153</td>
<td>2.126</td>
<td>0-31.736</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies with positive results</td>
<td>421 ± 1525</td>
<td>80</td>
<td>8-11 619</td>
<td>.30</td>
</tr>
<tr>
<td>Studies with negative results</td>
<td>600 ± 2207</td>
<td>90</td>
<td>12-15 832</td>
<td></td>
</tr>
</tbody>
</table>