

The proper use of citation data in journal management

Catherine M. Ketcham

Department of Pathology, Immunology and Laboratory Medicine, University of Florida College of Medicine, Gainesville, FL, USA

Received: 2008.10.03, **Accepted:** 2008.10.27, **Published online first:** 2008.12.01

Abstract

Journal impact factors (IF) are often maligned in editorials found in scientific publications, yet citation data can be used appropriately in journal management. The editors of Laboratory Investigation have found that weekly tracking of citation data for this and other highly ranked pathology journals provides valuable feedback on editorial performance and enables us to predict accurate IFs at least six months in advance. Once the IFs are released, it is useful to quantify the contributions of specific article categories, such as reviews and research articles, to the official IFs. In an ongoing attempt to understand the relationship between article downloads and eventual citations, we also analyze the citation rate of papers that had previously been the most frequently accessed on our web site. Finally, as a measure of editorial judgment, the papers that contributed no citations to the journal's IF are examined as are the papers that were rejected by Laboratory Investigation (Lab. Invest.) but subsequently published elsewhere. Thus the editors of Lab. Invest. use citation data in several ways to measure our progress in elevating the quality of the journal and understand the citation dynamics of papers we publish, while remaining true to the journal's fundamental operating premise: Publish high-quality original work relating to the mechanisms of disease.

Key words: impact factor, Web of Science, editorial policy, pathology.

Abbreviations: IF – impact factor, JCR – Journal Citation Report®, PIF – Pathobiology in Focus, WoS – Web of Science®.

Corresponding author: Catherine M. Ketcham, Ph.D., Department of Pathology, Immunology and Laboratory Medicine, University of Florida College of Medicine, P.O. Box 100275, 1600 SW Archer Road, Gainesville, FL 32610-0275, USA, tel.: +1 352-3922511, fax: +1 352-3926249, e-mail: ketcham@pathology.ufl.edu

INTRODUCTION

Many editors and publishers eagerly await the release of the Thomson Reuters journal impact factors (IFs) in June of each year. However, an inevitable rash of editorials materializes every spring to denounce IFs and their pernicious influence on the scientific community. It is far more popular to point out the shortcomings and misuse of citation data than to offer insight as to how it can be used constructively. The purpose of this article is to outline the ways that familiarity with the Web of Science® (WoS) and the Journal Citation Reports® (JCR) can enhance the journal editorial process. The editors of Laboratory Investigation (Lab. Invest.) have determined that it is possible to track and predict the IFs for several pathology journals at least six months, and potentially a year, before they are issued. We are also able to determine the contributions of various article types to a journal's IF, learn the fate of rejected articles that were subsequently published else-

where, and investigate the relationship between internet traffic and eventual citations. These endeavors provide positive feedback and constructive criticism for the editors and aid in the continued improvement of the journal.

PREDICTION OF IF

Perhaps the most useful feature of the WoS is that it allows tracking of journals' IFs in real time. IFs are issued only once a year and reflect papers that were published 1½–2½ years earlier and handled by an editorial group up to 4 years previously. Thus, monitoring the data in a more proximate time frame gives editors more continuous feedback about the results of their editorial decisions.

The process for collecting the data and predicting IF has been described previously [5]. Briefly, the method involves tracking the weekly IF as it develops using cita-

Table 1. Prediction of journal IFs from Thomson WoS® data

Journal	Predicted 2006 IF*	Official 2006 IF	Predicted 2006 IF as % of official IF	Predicted 2007 IF*	Official 2007 IF	Predicted 2007 IF as % of official IF
Am. J. Pathol.	5.665	5.917	95.7	5.448	5.487	99.3
J. Pathol.	5.612	5.759	97.4	5.421	5.423	99.9
Lab. Invest.	4.396	4.453	98.7	4.444	4.479	99.2

Abbreviations: Am. J. Pathol. – American Journal of Pathology, J. Pathol. – Journal of Pathology, Lab. Invest. – Laboratory Investigation.

*The 2006 data were previously published in ref. [5] and are reproduced with permission. The 2007 data are unpublished but were originally presented in an annual report issued to the United States and Canadian Academy of Pathology, February 2008.

tion data on the WoS and mapping the rate of IF accumulation to historical IF accrual over previous years to calculate the predicted IF. Table 1 shows the predictions made according to this formula and actual values of IFs for three top pathology journals in 2006 and 2007. The predicted IF values range from 95.7–99.9% of the actual IFs for 2006 and 2007 [7], which debunks the myth that the raw data from the WoS are inconsistent with the official IFs posted on the JCR [1, 8].

One of the drawbacks of the prediction method is that it requires a historical record of citation data from previous years. Figure 1 shows the weekly accumulated IFs for Lab. Invest. over a 3-year timeframe, 2005–2007, as a percentage of official IF. Note that there are differences in the accumulation and/or posting rates of citation data on the WoS between years; 2007 shows a much slower accumulation rate than the previous two years. It is important to note that the rates of accumulation of IF differ between journals; as each of the pathology journals monitored has a unique set of chronologic data (not

shown), the data for accumulation of the Lab. Invest. IFs cannot necessarily be used to extrapolate predicted IF for other journals.

Annual variation in the rate of data posting, which is apparent in Fig. 1, accounts for some of the inaccuracy in the prediction formula. In the future we will employ a “sliding” calendar that is based on the publication dates of the citing publications, not the dates of their posting in the WoS, to obtain a more chronologically accurate prediction formula.

ANALYSIS OF EACH OF THE TWO YEARS THAT CONTRIBUTE TO IF

Somewhat arbitrarily, the IF involves two years of citation data. The 2007 IF for any journal is a calculated ratio of the number of 2007 citations of articles published in 2005 and 2006 to the number of eligible articles published in the same two years. It is a simple matter to

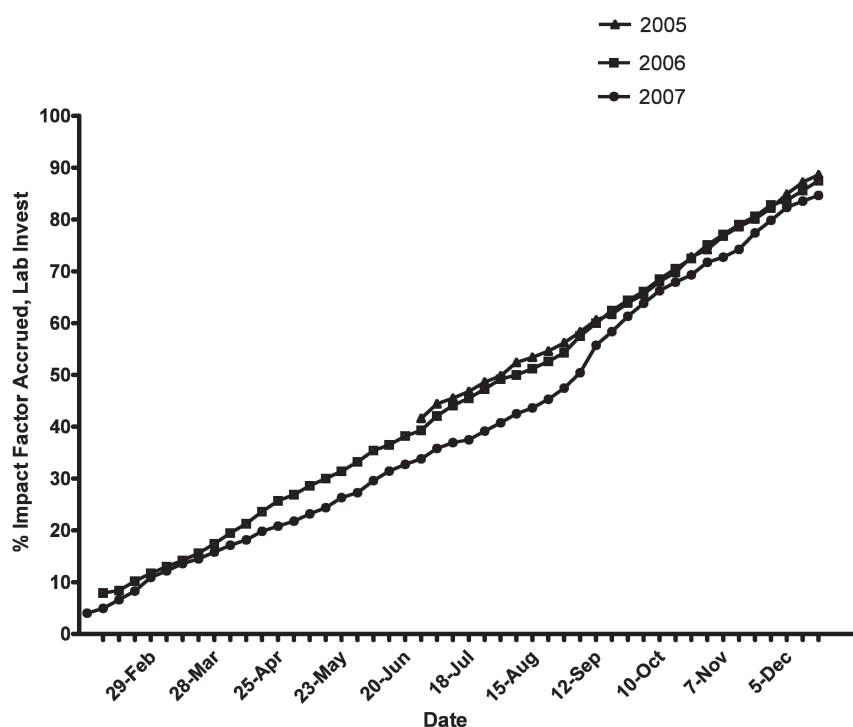


Fig. 1. Accrual rate of IF citations for Lab. Invest., 2005–2007. The citations for Lab. Invest. papers were determined weekly from WoS for the period of January 2005–December 2008. The developing IFs are presented as a percentage of the official IFs from the JCR.

Table 2. IFs analyzed by annual contribution

	Lab. Invest.	Am. J. Pathol.	J. Pathol.
2007 IF	4.479	5.487	5.423
2006 papers only	3.848	4.589	4.828
2005 papers only	5.031	6.531	6.080
2006 IF	4.453	5.917	5.759
2005 papers only	3.977	4.853	5.195
2004 papers only	4.836	6.732	6.301
2005 IF	3.859	5.796	6.213
2004 papers only	3.528	4.964	4.727
2003 papers only	4.149	6.580	7.437
2004 IF	3.702	6.441	5.333
2003 papers only	2.928	5.201	5.184
2002 papers only	4.475	7.678	5.483
2003 IF	4.418	6.946	4.933
2002 papers only	3.541	5.905	4.266
2001 papers only	5.363	7.946	5.565
2002 IF	4.000	6.750	4.563
2001 papers only	3.823	5.437	3.850
2000 papers only	4.354	8.160	5.283
2001 IF	3.934	7.103	5.064
2000 papers only	3.000	5.674	3.783
1999 papers only	4.982	8.587	6.211

The results for 2001–2006 were originally published in a different format in ref. [6]. Modified and reproduced with permission.

take the citation data shown in the JCR and calculate “one-year impact factors” as well, for example the citations of 2005 papers in 2007 and of 2006 papers in 2007. Why do this? Breaking down the IF to annual contributions can provide useful information about the upward or downward trajectory of a journal.

Table 2 shows that Lab. Invest. endured a historically low IF in 2004 (3.702), reflecting the papers published in 2002 and 2003. This table reveals that the low 2004 IF is mainly the result of the poor performance of the 2003 papers, because the 2003 papers’ impact in 2004 is the lowest “later year” number on the table for this journal (2.928, while the range for the rest of the years is 3.000–3.977). Additionally, the 2003 Lab. Invest. papers continued to perform poorly in 2005; their impact was the lowest for all reported “earlier years” for the journal (4.149, range of other years: 4.354–5.363). However, the editors noted that their 2004 papers performed well in 2005 (3.528, range for 2001–2005: 2.928–3.823) and therefore predicted correctly that the IF would start to recover in 2006.

THE EFFECT OF REVIEW ARTICLES ON IF

It is widely believed that publishing review articles is a good way for a journal to increase visibility and IF. The appeal of publishing review articles is apparent by

examining the JCR: the IFs of review journals are inordinately high. In 2007, of the top 50 journals (ranked by IF), almost half (22 titles) were review journals. However, it does not necessarily follow that publishing reviews in “regular” journals that publish original research will significantly increase the IFs of these journals, despite the fact that 1/7 of articles published in the biomedical literature are reviews [6]. Editors who put forth the effort to recruit and publish review articles might find it informative to determine the contribution of their review articles to IF.

An examination of review articles that were published in pathology journals in 2005 revealed that only 21% of the reviews had been cited more than 10 times since publication [6]. Accordingly, the IFs of 12 pathology journals were compared with or without review articles for the period 2000–2006 [6] and results for three of the journals are shown in Table 3. Inclusion of review articles resulted in a large increase for only one of the 12 journals examined, Journal of Pathology (J. Pathol.), by 0.610 ± 0.153 units. The increase for J. Pathol. was significantly greater than that for American Journal of Pathology (Am. J. Pathol.; 0.109 ± 0.086) and Lab. Invest. (0.147 ± 0.088). Notably, J. Pathol. publishes almost four times more review articles than Am J. Pathol. or Lab. Invest. Nonetheless, most of the IF for all three journals comes from the citation of original articles.

These results indicate that the IF of a research-oriented journal will not substantially increase unless it consistently publishes a relatively large proportion of highly cited review articles. Therefore, publication of high-quality original research articles is, and should be, the primary mechanism for elevation of IF for this category of journals.

Table 3. IF, proportion of IF due to review articles, and number of review articles for Am. J. Pathol., J. Pathol., and Lab. Invest., 2000–2006

Journal	Mean IF	Mean IF contribution from review articles	Mean # review articles/year
Am. J. Pathol.	6.489	0.109	4.5
J Pathol	4.797	0.610	17.8
Lab. Invest.	4.179	0.147	4.7

These results were originally published in different format ref. [6]; modified and reproduced with permission. The methodology for the calculation of IF without review articles is also outlined in detail in ref. [6].

THE IMPACT OF OTHER ARTICLE TYPES

It is also possible to determine the impact of any type on article on a journal’s IF. In the last half of 2006, Lab. Invest. unveiled the series “Pathobiology in Focus” (PIF), which presents novel and/or controversial hypotheses of disease mechanisms and summarizes complex

Table 4. Citations of 2006 Lab. Invest. papers

Month of publication, 2006	Article type	Times cited*
August	Pathobiology in focus (1)	4.0
	Review (1)	4.0
	All research articles (7)	4.3
	Technical report (1)	1.0
	Mean, all papers in issue	3.9
September	Pathobiology in focus (1)	13
	All research articles (8)	3.9
	Technical report (1)	14
	Mean, all papers in issue	5.8
October	Pathobiology in focus (1)	5.0
	All research articles (9)	4.1
	Mean, all papers in issue	4.2
November	Pathobiology in focus (1)	6.0
	All research articles (9)	3.3
	Mean, all papers in issue	3.6
December	Pathobiology in focus (1)	22
	All research articles (8)	5.5
	Mean, all papers	7.3

* Times cited (for individual papers) or mean times cited (for multiple papers in category) from publication through August 27, 2008.

research concepts for practicing and investigative pathologists [9]. Now that two years have passed since the first PIFs, it is of interest to determine whether or not the earlier articles in the series have been cited more frequently than other Lab. Invest. article types. Accordingly, Table 4 shows the citation rate (or average citation rate, as appropriate) for each article type in the August through December 2006 issues of Lab. Invest. These data reveal that each month from August through December 2006, the PIF article has been cited as well as, or better than, the average article in the issue. The table also shows that the one review article published during that timeframe (August 2006) had the same citation rate as the average article for that issue. One Technical Report (September 2006) has been relatively highly cited, but the other (August 2006) has not. These results are further support for the Lab. Invest. editors' conclusions that the best way to advance the quality of the journal is to publish high-quality work that is within the scope of the journal (the mechanisms of disease) and of interest to the members of the society that owns and publishes the journal (The United States and Canadian Academy of Pathology).

PAPERS WITH NO CITATIONS

Editors can use the WoS to find and examine papers that have received no citations during any timeframe of interest. If a paper has no citations (not even by the authors), one could argue that the article may represent a scientific "dead end", and the study of such papers

Table 5. Ranking of Am. J. Pathol., J. Pathol., and Lab. Invest. by 2007 IF and "zero citation" papers from 2006 and 2007

Journal	2007 IF	% papers with no citations in 2006*	% papers with no citations in 2007*
Am. J. Pathol.	5.487	1.9	16.6
J. Pathol.	5.423	2.1	18.0
Lab. Invest.	4.479	4.4	18.3

*Reviews and articles only, as of September 13, 2008.

could help editors calibrate their future decision-making processes. A paper with no citations during its first three publication years, i.e. the period of time in which citations to the paper would contribute to IF or Immediacy Index (same-year citation index), clearly has a measurable detrimental effect on the metrics for the journal.

When the rate of un-cited articles was used as a quantitative measure of journal quality in over 30,000 articles in more than 200 journals, a significant negative correlation between the percentage of never-cited articles and journal IF was found [10]. In practical terms, analysis of the "zero citation" papers is quicker and easier than predicting the IF when editors have an interest in ranking competing journals. Table 5 compares the 2007 IFs of three pathology journals with the percentages of non-cited papers in 2006 and 2007 (each of the journals had no "zero citation" papers for 2005). The rank order of the journals by IF and by non-cited papers from either year is the same. The predicted IFs for 2008 also follow this pattern (data not shown).

THE FATE OF REJECTED PAPERS

The editors of Lab. Invest. pondered the fate of papers rejected by the journal: what proportion was eventually published and did these papers appear in higher or lower IF journals? Through PubMed and the WoS, the editors found that 81% of the papers rejected from January through June 2005 were eventually published elsewhere; of those published, 24% were published in higher IF journals, 72% were published in lower IF journals, and 4% were published in journals with no IF.

Remembering that the IF of a journal is not an indication of the quality of any one article in that journal, the question arises: how well cited are these particular "recycled" papers compared with other papers appearing at the same time in the same journal or in corresponding issues of Lab. Invest.? More importantly, were the papers that were subjected to full review and then rejected subsequently modified according to the reviewers' and editors' comments before publication elsewhere? To the best of our knowledge, these issues have never been addressed comprehensively in the literature,

Table 6. Most frequently downloaded pdf articles from Lab. Invest. Website, January–September 2005

Title	Article type	Rank by pdf downloads	Rank by citations*
Downregulation of major histocompatibility complex antigens in invading glioma cells: stealth invasion of the brain	research article	1	41/129
Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes	technical report	2	1/129
The role of shear stress in the pathogenesis of atherosclerosis	review	3	2/129
Local overexpression of HB-EGF exacerbates remodeling following myocardial infarction by activating noncardiomyocytes	research article	4	90/129
Rapid detection of <i>VHL</i> exon deletions using real-time quantitative PCR	research article	5	19/129

*From publication through December 2007. All of the papers were published in 2005 and there were a total of 129 papers published in Lab. Invest. in 2005.

but are of great interest to scientists who donate their time for peer review. Such analyses are the topic of a manuscript in progress.

INTERNET VIEWS/DOWNLOADS AND CITATIONS

Information on internet views and downloads of a journal's papers is available shortly after publication and provides valuable feedback to the editorial team. However, editors cannot help but wonder whether the articles that receive a lot of attention online will be highly cited in the future. There is a murky relationship between journal articles that are popular on the web and their eventual citation rate. Reports on this topic present conflicting results [2, 3]; nevertheless, models have been developed to predict citations from downloads [4]. The editors of Lab. Invest. have found that, at least for this journal, all that can be said with confidence is that there is some overlap between the most accessed papers on the web and the most highly cited papers.

The five Lab. Invest. papers with the highest pdf downloads for January–September 2005 were all published in 2005 and are shown in Table 6. They were also ranked by the number of citations they received from publication through December 2007. This range of dates was chosen because it covers the citations used in the calculations of both the Immediacy Index and the IF. Of the five most downloaded papers in that timeframe, three are original research articles, one is an authoritative review article, and one is a broadly applicable technical report. The technical report is the most highly cited Lab. Invest. paper of 2005, and the review article falls into second place. However, the research articles vary widely in their citation ranks: 19/129, 41/129, and 90/129. These results affirm one of the fundamental philosophies of the editors of Lab. Invest.: It is not always possible to predict the behavior of the reading public; therefore, the editors should not select papers for publication based on our predictions of their ultimate popularity, but rather on our assessment of their scientific value.

In conclusion, the examples discussed above are not an exhaustive list of the uses for the WoS and the JCR; there is much more for editors to garner. We encourage other editors to venture into these web sites in order to enrich their own editorial experiences.

Acknowledgment: Special thanks are due to Lab. Invest. Editor-in-Chief James M. Crawford, M.D., Ph.D., for five years of excellent mentorship. The other members of the Lab. Invest. editorial team, 2004–2008, are also appreciated for their encouragement and support. They are: Senior Associate Editors Jerrold R. Turner, M.D., Ph.D. and Anthony T. Yachnis, M.D.; and Associate Editors Raul C. Braylan, M.D., Laurence Morel, Ph.D., and Naohiro Terada, M.D., Ph.D.

Any opinions presented herein are entirely those of the author and do not represent an official position of Lab. Invest., the United States and Canadian Academy of Pathology, or Nature Publishing Group. The author is not associated with and has not received payment from Thomson Reuters, publisher of the Web of Science® and Journal Citation Reports®.

REFERENCES

1. Brumback R. A. (2008): Worshiping false idols: The impact factor dilemma. *J. Child Neurol.*, **23**, 365–367.
2. Chu H. and Krichel T. (2007): Downloads vs. Citations: Relationships, Contributing Factors and Beyond. The 11th International Conference on Scientometrics and Informetrics, Madrid, Spain.
3. Coats A. J. (2005): Top of the charts: download versus citations in the International Journal of Cardiology. *Int. J. Cardiol.*, **105**, 123–125.
4. Jahandideh S., Abdolmaleki P. and Asadabadi E. B. (2007): Prediction of future citations of a research paper from number of its internet downloads. *Med. Hypotheses*, **69**, 458–459.
5. Ketcham C. M. (2006): Predicting impact factor one year in advance. *Lab. Invest.*, **86**, 632.
6. Ketcham C. M. and Crawford J. M. (2007): The impact of review articles. *Lab. Invest.*, **87**, 1174–1185.
7. Ketcham C. M. and Crawford J. M. (2008): Can impact factor data be trusted? *Lab. Invest.*, **88**, 340–341.
8. Rossner M., Van Epps H. and Hill E. (2007): Show me the data. *J. Exp. Med.*, **204**, 3052–3053.

9. Turner J. R., Maitra A., Natkunam Y., Rubin B. P., Rubin M. A. and Teitell M. A. (2006): Bringing pathobiology into focus. *Lab. Invest.*, **86**, 632.
10. Weale A. R., Bailey M. and Lear P. A. (2004): The level of non-citation of articles within a journal as a measure of quality: a comparison to the impact factor. *BMC Med. Res. Methodol.*, **4**, 14.