The History and Meaning of the Journal Impact Factor

Eugene Garfield, PhD

I FIRST MENTIONED THE IDEA OF AN IMPACT FACTOR IN Science in 1955.1 With support from the National Institutes of Health, the experimental Genetics Citation Index was published, and that led to the 1961 publication of the Science Citation Index.2 Irving H. Sher and I created the journal impact factor to help select additional source journals. To do this we simply re-sorted the author citation index into the journal citation index. From this simple exercise, we learned that initially a core group of large and highly cited journals needed to be covered in the new Science Citation Index (SCI). Consider that, in 2004, the Journal of Biological Chemistry published 6500 articles, whereas articles from the Proceedings of the National Academy of Sciences were cited more than 300,000 times that year. Smaller journals might not be selected if we rely solely on publication count,3 so we created the journal impact factor (JIF).

The Table provides a selective list of journals ranked by impact factor for 2004. The Table also includes the total number of articles published in 2004, the total number of articles published in 2002 plus 2003 (the JIF denominator), the citations to everything published in 2002 plus 2003 (the JIF numerator), and the total citations in 2004 for all articles ever published in a given journal. Sorting by impact factor allows for the inclusion of many small (in terms of total number of articles published) but influential journals. Obviously, sorting by total citations or other provided data would result in a different ranking.

The term “impact factor” has gradually evolved to describe both journal and author impact. Journal impact factors generally involve relatively large populations of articles and citations. Individual authors generally produce smaller numbers of articles, although some have published a phenomenal number. For example, transplant surgeon Tom Starzl has coauthored more than 2000 articles, while Carl Djerassi, inventor of the modern oral contraceptive, has published more than 1300.

Even before the Journal Citation Reports (JCR) appeared, we sampled the 1969 SCI to create the first published ranking by impact factor.4 Today, the JCR includes every journal citation in more than 5000 journals—about 15 million citations from 1 million source items per year. The precision of impact factors is questionable, but reporting to 3 decimal places reduces the number of journals with the identical impact rank. However, it matters very little whether, for example, the impact of JAMA is quoted as 24.8 rather than 24.831.

A journal’s impact factor is based on 2 elements: the numerator, which is the number of citations in the current year to items published in the previous 2 years, and the denominator, which is the number of substantive articles and reviews published in the same 2 years. The impact factor could just as easily be based on the previous year’s articles alone, which would give even greater weight to rapidly changing fields. An impact factor could also take into account longer periods of citations and sources, but then the measure would be less current.

Scientometrics and Journalology

Citation analysis has blossomed over the past 4 decades. The field now has its own International Society of Scientometrics and Informetrics.5 Stephen Lock, former editor of BMJ, aptly named the application of bibliometrics to journals evaluation “journalology.”6

All citation studies should be adjusted to account for variables such as specialty, citation density, and half-life.7 The citation density is the average number of references cited per source article and is significantly lower for mathematics journals than for molecular biology journals. The half-life (ie, number of retrospective years required to find 50% of the cited references) is longer for physiology journals than that for physics journals. For some fields, the JCR’s 2-year period for calculation of impact factors may or may not provide as complete a picture as would a 5- or 10-year period. Nevertheless, when journals are studied by category, the rankings based on 1-, 7-, or 15-year impact factors do not differ significantly.8,9 When journals are studied across fields, the ranking for physiology journals improves significantly as the number of years increases, but the rankings within the category do not significantly change. Similarly, Hansen and Henriksen10 reported “good agreement between the journal impact factor and the cumulative citation frequency of papers on clinical physiology and nuclear medicine.”

There are exceptions to these generalities. Critics of the JIF will cite all sorts of anecdotal citation behavior that do not represent average practice. Referencing errors abound,
but most are variants that do not affect journal impact, since only variants in cited journal abbreviations matter in calculating impact. These are all unified prior to issuing the JCR each year.

The impact factors reported by the JCR tacitly imply that all editorial items in BMJ, JAMA, Lancet, New England Journal of Medicine, etc., can be neatly categorized, but such journals publish large numbers of items that are not substantive in regards to citations. Correspondence, letters, commentaries, perspectives, news stories, obituaries, editorials, interviews, and tributes are not included in the JCR’s denominator. However, they may be cited, especially in the current year. For that reason, they do not significantly affect impact calculations. Nevertheless, since the numerator includes later citations to these ephemera, some distortion will result, although only a small group of leading medical journals are affected.

The assignment of publication codes is based on human judgment. A news story might be perceived as a substantive article, and a significant letter might not be. Furthermore, no effort is made to differentiate clinical vs laboratory studies or, for that matter, practice-based vs research-based articles. All these potential variables provide grist for the critical mill of citation aficionados.

**Size vs Citation Density**

There is a widespread belief that the size of the scientific community that a journal serves significantly affects impact factor. This assumption overlooks the fact that while more authors produce more citations, these must be shared by a larger number of cited articles. Most articles are not well-cited, but some articles may have unusual cross-disciplinary impact. It is well known that there is a skewed distribution of citations in most fields. The so-called 80/20 phenomenon applies, in that 20% of articles may account for 80% of the citations. The key determinants of impact factor are not the number of authors or articles in the field but, rather, the citation density and the age of the literature cited. The size of a field, however, will increase the number of “super-cited” papers. And while a few classic methodology papers exceed a high threshold of citation, thousands of other methodology and review papers do not. Publishing mediocre review papers will not necessarily boost a journal’s impact. Some examples of super-citation classics include the Lowry method, cited 300,000 times, and the Southern Blot technique, by E. M. Southern, cited 30,000 times. Since the roughly 60 papers cited more than 10,000 times are decades old, they do not affect the calculation of the current impact factor. Indeed, of 38 million items cited from 1900-2005, only 0.5% were cited more than 200 times. Half were not cited at all, and about one quarter were not substantive articles but rather the editorial ephemera mentioned earlier.

The skewness of citations is well known and repeated as a mantra by critics of the impact factor. If manuscript refereeing or processing is delayed, references to articles that are no longer within the JCR’s 2-year impact window will not be counted. Alternatively, the appearance of articles

**Table. Selected Biomedical Journals Ranked by Impact Factor**

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>2004 Impact Factor</th>
<th>2004 Articles</th>
<th>2002 + 2003 Articles</th>
<th>Citations in 2004 To 2002 + 2003 Articles Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual Review of Immunology</td>
<td>52.4</td>
<td>30</td>
<td>51</td>
<td>2674</td>
</tr>
<tr>
<td>New England Journal of Medicine</td>
<td>38.6</td>
<td>316</td>
<td>744</td>
<td>28,696</td>
</tr>
<tr>
<td>Nature Reviews: Cancer</td>
<td>36.6</td>
<td>79</td>
<td>149</td>
<td>54,477</td>
</tr>
<tr>
<td>Physiological Reviews</td>
<td>33.9</td>
<td>35</td>
<td>61</td>
<td>2069</td>
</tr>
<tr>
<td>Nature Reviews: Immunology</td>
<td>32.7</td>
<td>80</td>
<td>151</td>
<td>49,357</td>
</tr>
<tr>
<td>Nature</td>
<td>32.2</td>
<td>878</td>
<td>1478</td>
<td>56,255</td>
</tr>
<tr>
<td>Science</td>
<td>31.9</td>
<td>845</td>
<td>1736</td>
<td>55,297</td>
</tr>
<tr>
<td>Nature Medicine</td>
<td>31.2</td>
<td>168</td>
<td>318</td>
<td>99,29</td>
</tr>
<tr>
<td>Cell</td>
<td>28.4</td>
<td>288</td>
<td>627</td>
<td>17,800</td>
</tr>
<tr>
<td>Nature Immunology</td>
<td>27.6</td>
<td>130</td>
<td>273</td>
<td>7,531</td>
</tr>
<tr>
<td>JAMA</td>
<td>24.8</td>
<td>351</td>
<td>751</td>
<td>18,648</td>
</tr>
<tr>
<td>Nature Genetics</td>
<td>24.7</td>
<td>191</td>
<td>420</td>
<td>10,372</td>
</tr>
<tr>
<td>Annual Review of Neuroscience</td>
<td>23.1</td>
<td>26</td>
<td>42</td>
<td>972</td>
</tr>
<tr>
<td>Pharmacological Reviews</td>
<td>22.8</td>
<td>19</td>
<td>49</td>
<td>1119</td>
</tr>
<tr>
<td>Lancet</td>
<td>21.7</td>
<td>415</td>
<td>1020</td>
<td>22,147</td>
</tr>
<tr>
<td>Annals of Internal Medicine</td>
<td>13.1</td>
<td>189</td>
<td>396</td>
<td>5193</td>
</tr>
<tr>
<td>Annual Review of Medicine</td>
<td>11.2</td>
<td>29</td>
<td>65</td>
<td>728</td>
</tr>
<tr>
<td>Archives of Internal Medicine</td>
<td>7.5</td>
<td>282</td>
<td>967</td>
<td>42,657</td>
</tr>
<tr>
<td>BMJ</td>
<td>7.0</td>
<td>623</td>
<td>1222</td>
<td>8601</td>
</tr>
<tr>
<td>CMAJ</td>
<td>5.9</td>
<td>100</td>
<td>220</td>
<td>1307</td>
</tr>
</tbody>
</table>

©2006 American Medical Association. All rights reserved.
on the same subject in the same issue may have an upward effect, as shown by Opthof. For greater precision, it is preferable to conduct item-by-item journal audits so that any differences in impact for different types of editorial items can be taken into account.

Some editors would calculate impact solely on the basis of their most-cited papers so as to diminish their otherwise low impact factors. Others would like to see rankings by geographic or language group because of the SCI’s alleged English-language bias, even though the SCI covers European—largely German, French, and Spanish—medical journals.

Other objections to impact factors are related to the system used in the JCR to categorize journals. The heuristic methods used by Thomson Scientific (formerly Thomson ISI) for categorizing journals are by no means perfect, even though citation analysis informs their decisions. Recent work by Pudovkin and myself is an attempt to group journals objectively. We rely on the 2-way citational relationships between journals to reduce the subjective influence of journal titles such as the Journal of Experimental Medicine—one of the top 5 immunology journals.

The JCR recently added a new feature that provides the ability to more precisely establish journal categories based on citation relatedness. A general formula based on the citation relatedness between 2 journals is used to express how close they are in subject matter. For example, the journal Controlled Clinical Trials is more closely related to JAMA than at first meets the eye. In a similar fashion, using the relatedness formula one can demonstrate that, in 2004, the New England Journal of Medicine was among the most significant journals to publish cardiovascular research.

Journal Performance Indicators

Many of the discrepancies inherent in JIFs are eliminated altogether in another Thomson Scientific database called Journal Performance Indicators (JPI). Unlike the JCR, the JPI database links each source item to its own unique citations. Therefore, the impact calculations are more precise. Only citations to the substantive items that are in the denominator are included. And it is possible to obtain cumulative impact measures covering longer time spans. For example, the cumulated impact for JAMA articles published in 1999 was 84.5. This was derived by dividing the 31 257 citations received from 1999 to 2004 by the 370 articles published in 1999. That year JAMA published 1905 items, of which 680 were letters and 253 were editorials. Citations to these items were not included in the JPI calculation of impact.

In addition to helping libraries decide which journals to purchase, JIFs are also used by authors to decide where to submit their articles. As a general rule, the journals with high impact factors include the most prestigious. Some would equate prestige with high impact.

The use of JIFs instead of actual article citation counts to evaluate individuals is a highly controversial issue. Granting and other policy agencies often wish to bypass the work involved in obtaining citation counts for individual articles and authors. Allegedly, recently published articles may not have had enough time to be cited, so it is tempting to use the JIF as a surrogate evaluation tool. Presumably, the mere acceptance of the paper for publication by a high-impact journal is an implied indicator of prestige. Typically, when the author’s work is examined, the impact factors of the journals involved are substituted for the actual citation count. Thus, the JIF is used to estimate the expected count of individual papers, which is rather dubious considering the known skewness observed for most journals.

Today, so-called Webometrics are increasingly brought into play, though there is little evidence that this approach is any better than traditional citation analysis. Web “sitations” may occur a little earlier, but they are not the same as “citations.” Thus, one must distinguish between readership or downloading and actual citation in new published papers. But some limited studies indicate that Web sitation is a harbinger of future citation.

The assumption that the impact of recent articles cannot be evaluated in the SCI is not universally correct. While there may be several years’ delay for some topics, papers that achieve high impact are usually cited within months of publication and certainly within a year or so. This pattern of immediacy has enabled Thomson Scientific to identify “hot papers” in its bimonthly publication, Science Watch. However, full confirmation of high impact is generally obtained 2 years later. The Scientist waits up to 2 years to select hot papers for commentary by authors. Most of these papers will eventually go on to become “citation classics.”

Two recent examples of hot papers published in JAMA are those on the benefits and risks of estrogen in postmenopausal women. The first was cited in 132 articles after 6 months, then 776 times in 2003 and 862 times in 2004. The second, more recent, hot paper has already been cited in 300 articles.

Conclusion

Of the many conflicting opinions about impact factors, Hoef fel expressed the situation succinctly:

Impact Factor is not a perfect tool to measure the quality of articles but there is nothing better and it has the advantage of already being in existence and is, therefore, a good technique for scientific evaluation. Experience has shown that in each specialty the best journals are those in which it is most difficult to have an article accepted, and these are the journals that have a high impact factor. Most of these journals existed long before the impact factor was devised. The use of impact factor as a measure of quality is widespread because it fits well with the opinion we have in each field of the best journals in our specialty.

The use of journal impacts in evaluating individuals has its inherent dangers. In an ideal world, evaluators would read each article and make personal judgments. The recent International Congress on Peer Review and Biomedical Publication (http://www.jama-peer.org) demonstrated the dif-
difficulties of reconciling such peer judgments. Most individuals do not have the time to read all the relevant articles. Even if they do, their judgment surely would be tempered by observing the comments of those who have cited the work. Online full-text access has made that practical.

Financial Disclosures: Dr Garfield owns stock in, and occasionally has reviewed for per diem payment from, Thomson Scientific.

Previous Presentations: Presented in part at the 2000 Council of Science Editors Annual Meeting; May 6–9, 2000; San Antonio, Tex; and at the Fifth International Congress on Peer Review and Biomedical Publication; September 16–18, 2005; Chicago, Ill.

REFERENCES
23. Pereneger TV. Relation between online “hit counts” and subsequent citations: prospective study of research papers in the BMJ. BMJ. 2004;329:546-547.