Panel on "Evaluative Measures for Resource Quality: Beyond the Impact Factor."

Eugene Garfield
Chairman Emeritus, Thomson ISI
3501 Market Street, Philadelphia PA 19104
Fax: 215-387-1266 () Tel. 215-243-2205
garfield@codex.cis.upenn.edu
www.eugenegarfield.org

presented at
Medical Library Association Meeting
Philadelphia, May 22, 2007

I first mentioned the idea of an impact factor in *Science* in 1955. [1] Then from 1960-63, the National Institutes of Health supported the experimental *Genetics Citation Index*. This project led to the 1961 *Science Citation Index*[2] which covered about 600 journals covered in *Current Contents*. We 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 a core group of large and highly cited journals had to be covered in the new *Science Citation Index* (*SCI*).

SLIDE #1:

TOP JOURNALS SORTED BY NUMBER OF ARTICLES, 2004

Abbreviated Journal Title	Total Cites	Impact Factor	Articles
J BIOL CHEM	405017	6.355	6585
P NATL ACAD SCI USA	345309	10.452	3084
BIOCHEM BIOPH RES CO	64346	2.904	2312
J IMMUNOL	108602	6.486	1793
BIOCHEMISTRY-US	96809	4.008	1687
J VIROL	74388	5.398	1464
J AGR FOOD CHEM	27992	2.327	1261
CANCER RES	105196	7.690	1253
J NEUROSCI	93263	7.907	1233
BLOOD	97885	9.782	1206
NUCLEIC ACIDS RES	66057	7.260	1160
CIRCULATION	115133	12.563	1129
FEBS LETT	54417	3.843	1112
NEUROSCI LETT	25138	2.019	1101
J CLIN MICROBIOL	35117	3.439	1090
TRANSPLANT P	9048	0.511	1070
CLIN CANCER RES	23585	5.623	1052
BRAIN RES	58204	2.389	1037
J UROLOGY	39589	3.713	1029
ONCOGENE	45546	6.318	1003

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 relied solely on absolute publication or citation counts,[3] so we created the journal impact factor (JIF).

SLIDE #2:
Slide 2 provides a selective list of journals ranked by impact factor for 2004. The Table includes the number of articles published in 2004, the citations to everything published in 2002 plus 2003 (the JIF numerator), and the total citations in 2004 for all articles ever published in the 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 parameters would result in a different ranking

SELECTED TOP BIOMEDICAL JOURNALS SORTED BY IMPACT FACTOR, 2004

Abreviated Journal Title	Total Cites	Impact Factor	Articles	Cites to 2002/3
ANNU REV IMMUNOL	14357	52.431	30	2674
NEW ENGL J MED	159498	38.570	316	28696
NAT REV CANCER	6618	36.557	79	5447
PHYSIOL REV	14671	33.918	35	2069
NAT REV MOL CELL BIO	9446	33.170	84	4876
NAT REV IMMUNOL	5957	32.695	80	4937
NATURE	363374	32.182	878	56255
SCIENCE	332803	31.853	845	55297
ANNU REV BIOCHEM	16487	31.538	33	1640
NAT MED	38657	31.223	168	9929
CELL	136472	28.389	288	17800
NAT IMMUNOL	14063	27.586	130	7531
JAMA-J AM MED ASSOC	88864	24.831	351	18648
NAT GENET	49529	24.695	191	10372
ANNU REV NEUROSCI	8093	23.143	26	972
PHARMACOL REV	7800	22.837	19	1119
NAT BIOTECHNOL	18169	22.355	138	5723
LANCET	126002	21.713	415	22147
ANN INTERN MED	36932	13.114	189	5193
ANNU REV MED	3188	11.200	29	728
ARCH INTERN MED	26525	7.508	282	4257
BRIT MED J	56807	7.038	623	8601
CAN MED ASSOC J	6736	5.941	100	1307

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 chemist Carl Djerassi 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 citations from more than 6000 journals—about 20 million citations from 1.2 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 but hypesters prefer the pseudo-precision of three decimal places.

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] meeting next month in Madrid. 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. It is significantly lower for mathematics than for molecular biology journals. The halflife (ie, number of retrospective years required to find 50% of the cited references) is longer for physiology than physics journals. For some fields, the *JCR*'s
two-year period for calculation of impact factors may or may not provide as adequate 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 usually do not differ significantly.[8,9] Similarly, Hansen and Henriksen[10] 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. Most are 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 considered substantive. 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 during the current year. For that reason, they do not usually 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. The size of the bibliometric literature suggests there are plenty of those, especially editors of low impact journals.

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 crossdisciplinary 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 dozen 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,[11] cited explicitly in over 300,000 papers, or EM Southern's Southern Blot technique, cited in 30,000 articles.[12] 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.[13] Alternatively, the appearance of articles on the same subject in the same issue may have an upward effect, as shown by Opthof.[14] 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.[15]

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. Work by Pudovkin and myself[16] and others 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.[17]

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 the *New England Journal of Medicine* was among the most significant journals that publish cardiovascular research.

Journal Performance Indicators

SLIDE #3:

JAMA, CITATION IMPACT (ALL ITEMS), IN ONE YEAR PERIODS, 1981 TO 2004.

JAMA
CITATION IMPACT (ALL ITEMS)
IN ONE YEAR PERIODS 1981 TO 2004

Source: *ISI Journal Performance Indicators file*, 2004

Rank	Year	Impact	Citations	Papers			
1	1981	29.57	16,291	551			
2	1982	35.53	20,358	573			
3	1983	40.11	22,219	554			
4	1984	35.26	21,791	618			
5	1985	35.05	18,436	526			
6	1986	48.76	24,576	504			
7	1987	44.70	26,688	597			
8	1988	48.40	30,009	620			
9	1989	55.79	34,979	627			
10	1990	54.83	35,968	656	**31,257**	**Citations received 1999-2004**	**=84.5**
11	1991	47.19	30,389	644	**370**	**Articles published in JAMA**	
12	1992	58.48	34,389	588		**in 1999**	
13	1993	65.55	38,349	585			
14	1994	70.54	39,148	555			
15	1995	81.99	45,094	550			
16	1996	60.16	32,908	547			
17	1997	58.19	32,821	564			
18	1998	75.20	37,372	497			
→ 19	1999	84.48	31,257	370	←		
20	2000	56.71	21,040	371			
21	2001	49.98	18,842	377			
22	2002	42.84	16,921	395			
23	2003	19.09	7,311	383			
24	2004	3.34	1,174	351			

Many of the discrepancies inherent in JIFs are eliminated altogether in another Thomson Scientific database called Journal Performance Indicators (JPI).[18] 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, Slide #3 shows that 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 impact 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 systematic evidence that this approach is any better than traditional citation analysis. Web "sitations" may occur a little earlier, but they are not necessarily the same as "citations." Thus, one must distinguish between readership or downloading and actual citation in newly published papers; that is, impact on research. But some limited studies indicate that Web sitation is a harbinger of future citation.[19-23]

The assumption that the impact of recent articles cannot be evaluated in the *SCI* is not universally correct. "Delayed recognition" is a relatively rare phenomenon which Glanzel and I have demonstrated.[24] 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* magazine waits up to 2 years to interview authors of selected hot papers. Most of these papers will eventually go on to become "citation classics."[25]

Conclusion

Of the many conflicting opinions about impact factors, Christine Hoeffel[26] 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, herefore, 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.

This opinion was publisher now ten years ago. I would like to know what Dr. Hoeffel would say today.

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 difficulties of reconciling such peer judgments. Most evaluators do not have the time to read all the relevant articles. In any case, their judgments surely would be tempered by observing in context the comments of those who have cited the work. Online full-text access has made that easier but just as in the days when evaluators relied on author reprints or used libraries that did not solve the problem of finding the time to read them all!

REFERENCES

1. Garfield E. Citation indexes to science: a new dimension in documentation through association of ideas. *Science*. 1955;122:108-111. Available at: http://garfield.library.upenn.edu/essays/v6p468y1983.pdf Accessed October 26, 2005.

2. Garfield E, Sher IH. *Genetics Citation Index*. Philadelphia, Pa: Institute for Scientific Information; July 1963. Available at: http://www.garfield.library.upenn.edu/essays/v7p515y1984.pdf Accessibility verified November 29, 2005.

3. Brodman E. Choosing physiology journals. *Bull Med Libr Assoc*. 1944;32:479-483.

4. Garfield E. Citation analysis as a tool in journal evaluation. *Science*. 1972;178: 471-479. Available at: http://garfield.library.upenn.edu/essays/V1p527y1962-73.pdf . Accessed October 25, 2005.

5. International Society of Scientometrics and Informetrics Web site. Available at: http://www.issi-society.info/ Accessibility verified November 14, 2005.

6. Lock SP. Journalology: are the quotes needed? *CBE Views*. 1989;12:57-59. Available at: http://garfield.library.upenn.edu/essays/v13p019y1990.pdf Accessed October 25, 2005.

7. Pudovkin AI, Garfield E. Rank-normalized impact factor: a way to compare journal performance across subject categories. In: *Proceedings of the 67th Annual Meeting of the American Society for Information Science & Technology*. Vol 41. Silver Spring, Md: American Society for Information Science & Technology; 2004:507-515. Available at: http://garfield.library.upenn.edu/papers/ranknormalizationasist2004published.pdf Accessed October 25, 2005.

8. Garfield E. Long-term vs. short-term journal impact: does it matter? *Scientist*. 1998;12:10-12. Available at: http://www.garfield.library.upenn.edu/commentaries/tsv12(03)p10y19980202.pdf Accessed October 25, 2005.

9. Garfield E. Long-term vs. short-term journal impact, II: cumulative impact factors. *Scientist*. 1998;12:12-13. Available at: http://www.garfield.library.upenn.edu/commentaries/tsv12(14)p12y19980706.pdf Accessed October 25, 2005.

10. Hansen HB, Henriksen JH. How well does journal "impact" work in the assessment of papers on clinical physiology and nuclear medicine? *Clin Physiol*. 1997; 17:409-418.

11. Lowry OH, Rosebrough NJ, Farr AL, et al. Protein measurement with the folin phenol reagent. *J Biol Chem*. 1951;193:265-275.

12. Southern EM. Detection of specific sequences among DNA fragments separated

by gel-electrophoresis. *J Mol Biol*. 1975;98:503-517.

13. Yu G, Wang X-Y, Yu D-R. The influence of publication delays on impact factors. *Scientometrics*. 2005;64:235-246.

14. Opthof T. Submission, acceptance rate, rapid review system and impact factor. *Cardiovasc Res*. 1999;41:1-4.

15. Garfield E. Which medical journals have the greatest impact? *Ann Intern Med*. 1986;105:313-320. Available at: http://www.garfield.library.upenn.edu/essays/v10p007y1987.pdf Accessed October 25, 2005.

16. Pudovkin AI, Garfield E. Algorithmic procedure for finding semantically related journals. *J Am Soc Inf Sci Technol*. 2002;53:1113-1119. Available at: http://www.garfield.library.upenn.edu/papers/pudovkinsemanticallyrelatedjournals2002.html Accessed October 25, 2005.

17. Garfield E. Journal Citation Studies, III: Journal of Experimental Medicine compared with Journal of Immunology: or, how much of a clinician is the immunologist? *Curr Contents Clin Med.* June 28, 1972:5-8. Available at: http://www.garfield.library.upenn.edu/essays/V1p326y1962-73.pdf Accessed October 25, 2005.

18. Thomson Scientific. Journal Performance Indicators. Available at: http://scientific.thomson.com/products/jpi/ Accessibility verified November 14, 2005.

19. Antelman K. Do open-access articles have a greater research impact? *Coll Res Libr*. 2004;65:372-382.

20. Lawrence S. Free online availability substantially increases a paper's impact. *Nature*. 2001;411:521.

21. Vaughan L, Shaw D. Bibliographic and web citations: what is the difference? *J Am Soc Inf Sci Technol*. 2003;54:1313-1322.

22. Kurtz MJ, Eichhorn G, Accomazzi A, et al. The effect of use and access on citations. *Inf Processing Manag*. 2005;41:1395-1402.

23. Perneger TV. Relation between online "hit counts" and subsequent citations: prospective study of research papers in the *BMJ*. *BMJ*. 2004;329:546-547.

24. Glanzel W. and Garfield E. "The Myth of Delayed Recognition -Citation analysis demonstrates that premature discovery, while rare, does occur: Nearly all significant research is normally cited soon after publication" *The Scientist* 18(11): 8-8 June 7 2004. http://garfield.library.upenn.edu/papers/mythdelayedrecognition2004.html

25. Citation Classics. Available at: http://www.citationclassics.org Accessibility verified November 14, 2005.

26. Hoeffel C. Journal impact factors. *Allergy*. 1998;53:1225.