CONTRIBUTIONS to SCIENCE, **5** (1): 63–70 (2009)
Institut d'Estudis Catalans, Barcelona
DOI: 10.2436/20.7010.01.60 ISSN: 1575-6343
*www.cat-science.cat*



# The evolution of the *Science Citation Index**

**Eugene Garfield****
Chairman Emeritus, Thomson Scientific ISI, Philadelphia, Pennsylvania, USA

**Resum.** El *Science Citation Index* (SCI) va ser proposat fa més de 50 anys per facilitar la disseminació i recuperació de literatura científica. El seu cercador era únic pel fet de basar-se en una recerca per cites, però no va ser àmpliament adoptat fins que va estar disponible en la xarxa en l'any 1972. El *Journal Citation Reports*, que apareix com a conseqüència d'aquest en 1975, incloïa a més una classificació basada en el factor d'impacte. No era comú utilitzar els factors d'impacte fins a fa una dècada quan van començar a ser usats com una alternativa per a calcular les freqüències esperades de cites a articles publicats recentment—una aplicació molt polèmica de la cienciometria com eina per avaluar institucions i científics. L'inventor de l'SCI, i la seva base de dades companya, l'SSCI, examinarà la seva història i discutirà el seu ús més recent en la visualització gràfica de microhistories de temes acadèmics. Mitjançant *HistCite*, un programari patentat per a l'anàlisi historiogràfic algorítmic, es parlarà sobre la genealogia del descobriment de Watson-Crick de la estructura de la doble hèlix del DNA i la seva relació amb el treball de Heidelberger, Avery i altres.

**Paraules clau:** *Science Citation Index* · bibliometria · Factor d'Impacte

**Abstract.** The *Science Citation Index* was proposed over 50 years ago to facilitate the dissemination and retrieval of scientific literature. Its unique search engine, based on citation searching, was not widely adopted until it was made available online in 1972. Its by product, *Journal Citation Reports*, became available in 1975 and included its rankings by impact factor. Impact factors were not widely implemented until about a decade ago, when they began to be used as surrogates for expected citation frequencies for recently published papers—a highly controversial application of scientometrics in evaluating scientists and institutions. Here, the inventor of both the *SCI* and its companion, *Social Sciences Citation Index*, review the history of these instruments and discusses their more recent use in graphically visualizing microhistories of scholarly topics. In an example thereof, the patented *HistCite* software for algorithmic historiographic analysis is used to follow the genealogy of the Watson-Crick discovery of the double-helix structure of DNA and its relationship to the work of Heidelberger, Avery, and others.

**Keywords:** *Science Citation Index* . bibliometrics · Impact Factor

## Introduction

The *Science Citation Index* (*SCI*) was first promulgated in *Science* in 1955 as an up-to-date tool to facilitate the dissemination and retrieval of scientific literature. As the older generation of scientists will remember, an already-existing information service known as *Current Contents* was the primordial revolutionary "idea" that made practical realization of the *SCI* possible. Remarkably, *Current Contents* is still published in print every week although its electronic version has been around for more than a decade. Even aficionados of *Current Contents* may not remember the role that the early electronic computer played in making it possible for it to appear each week together with its title word indexes and author address directory. In those days, conventional indexes were normally six months to three years behind the literature, whereas it was estimated that ten million worldwide reprint requests were generated each year thanks to *Current Contents*.

Nevertheless, the success of the *SCI* did not come from its original function as a search engine, but from its subsequent use as an instrument for measuring scientific productivity thanks to its by-product, the *SCI Journal Citation Reports* (JCR) and its impact factor rankings.

The *SCI*'s multidisciplinary database has two purposes: first, to identify what each scientist has published, and second, where and how often the papers by that scientist are cited. Hence, the *SCI* has always been divided into two author-based parts: the Source Author Index and the Citation Index. By extension, one can also determine what each institution and country has published and how often the respective papers are

* Based on the Lecture given by Eugene Garfield at the Institute for Catalan Studies, Barcelona, on 24 January, 2007. A previous version of this article was published in International Microbiology 2007, 10:65–69.

** Correspondence: Eugene Garfield, Thomson Scientific ISI, Philadelphia, Pennsylvania (USA). Tel. +1-2152432205. Fax +1-2153871266. Email: garfield@codex.cis.upenn.edu

cited. This is especially important, since it is remarkable how difficult it can often be to find a complete list of a particular author's publications.

The *Web of Science* (*WoS*)—the *SCI*'s electronic version—links these two functions: an author's publication can be listed by chronology, by journal, or by citation frequency. It also allows searching for scientists who have published over a given period of years. Table 1, for example, shows the resulting list for a search of scientists who have published for 70–85 years. The record is held by Issac M. Kolthoff, followed by molecular biologist Michael Heidelberger, whose last paper appeared in 2004, shortly before he died at the age of 104.

When the *SCI* was launched in 1964, Irving Sher and I had already begun using bibliographic citations to create topological maps, called historiographs, to investigate whether citation indexes could aid in writing mini-histories of scientific topics. More recently, the gigabyte memory capacities of computers made it possible to write a program called *HistCite*—a patented software that was in development for about five years and will be available commercially beginning in February of 2009—which accepts the output of a *WoS* search and automatically generates historiographs. By collecting all the relevant cited papers on a subject in a *WoS* search, *HistCite* represents the collective memory of the citing authors and produces a visual description of the topical history. A key question often arises as to the ability of citation indexing to retrieve all the relevant work on a topic. In the pre-WWII days, and much before the advent of molecular biology, citation practices were not nearly as standardized as they are today, and implicit citation was quite common. As a result, the explicit citation of earlier relevant work could not always be found.

Heidelberger was a pioneering molecular biologist; even before WWII, together with Oswald T. Avery and others at the Rockefeller Institute (including Colin M. Macleod and Maclyn McCarthy) he published some primordial work in the history of DNA. As a matter of fact, this work is a key link in the genealogical history of the Watson-Crick 1953 paper on the double-helix structure of DNA. *HistCite* was used to track the implicit connection between this Watson-Crick paper and the 1944 work of Avery et al. on pneumococcal DNA. As those familiar with the story know and although Jim Watson finally stated a few years ago that he regretted not having done so, Watson and Crick did not cite the 1944 Avery paper in their 1953 paper, since it was rushed into print without the usual reference checks. In order to demonstrate that the significance of Avery's work was indeed known to contemporary workers, we produced a series of *HistCite* files by doing an *SCI* search on the *WoS*. Then, in order to explore the historical connection between the work of Heidelberger and his co-author, Oswald Avery, it was essential that the *ISI* edit thousands of such implicit citations. Figure 1 shows the historiograph created to show links between the work of Heidelberger, Avery, and Watson and Crick.

Having demonstrated how the *WoS* search engine can be used to track the historical developments of scientific topics, we can now turn to the subject of the ubiquitous journal impact factor. The Annual *SCI Journal Citation Reports* were officially launched in 1975, although we had already been producing

**Table 1.** Scientists who have published for 70 years or more

| Scientist | Birth/Death | Publication Years | Years published |
|---|---|---|---|
| Izaak Maurits (Piet) Kolthoff (analytical chemist) | 1894-1993 | 1917-2002 | 86 |
| Michael Heidelberger (organic chemist –immunologist) | 1888-1991 | 1909-1993 | 85 |
| Melvin Guy Mellon ( chemist) | 1893-1993 | 1920-2003 | 84 |
| Ernst Mayr (biologist) | 1904-2005 | 1923-2005 | 83 |
| Michel Eugene Chevreul (chemist) | 1786-1889 | 1808-1889 | 82 |
| Carl S. Marvel (polymer chemist) | 1894-1988 | 1917-1996 | 80 |
| Joel H. Hildebrand (chemist) | 1881-1983 | 1907-1983 | 77 |
| Linus Pauling (chemist) | 1901-1994 | 1923-1998 | 76 |
| John Carew Eccles (neurophysiologist) | 1903-1997 | 1929-1992 | 74 |
| Donald Coxeter (mathematician) | 1907-2003 | 1930-2001 | 72 |
| Charles Scott Sherrington (physiologist) | 1857-1952 | 1882-1952 | 71 |
| Alexander Kossiakoff (engineer) Guided missile expert | 1914-2005 | 1935-2005 | 71 |
| Hans Albrecht Bethe (physicist) | 1906-2005 | 1934-2004 | 71 |
| Norman Hackerman (chemist) | 1912-2007 | 1936-2006 | 71 |
| Michael DeBakey (cardiac surgeon) | 1908-2008 | 1937-2006 | 70 |
| Gerhard Herzberg (chemist) | 1904-1999 | 1924-1992 | 69 |
| Herman Mark (polymer chemist) | 1895-1992 | 1922-1990 | 69 |

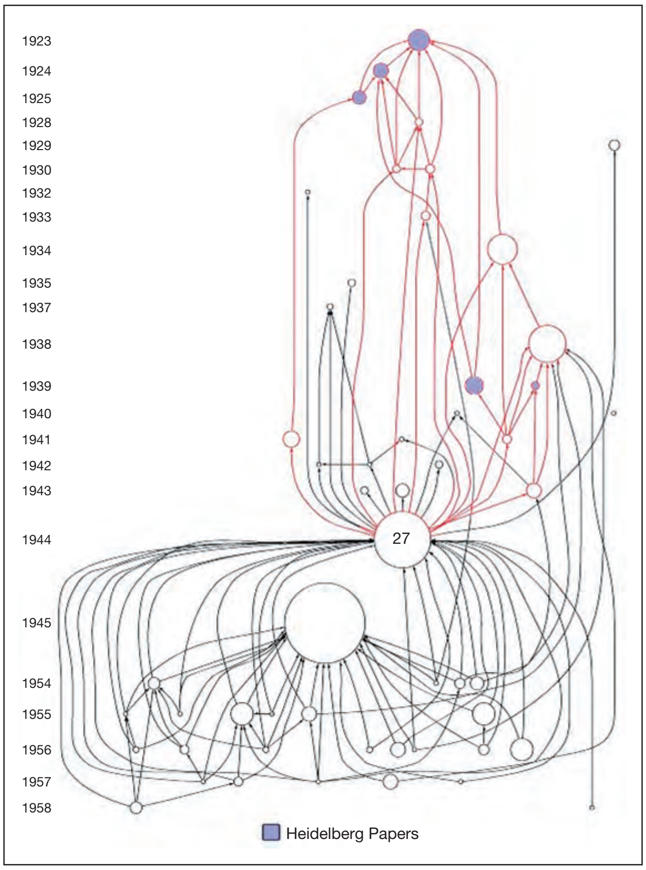


**Fig. 1.** Heidelberger, Avery, Watson-Crick connection.

these data for over a decade. The *JCR* evolved to provide a statistical summation of the Journal Citation Index, which in turn was the result of re-sorting the Author Citation Index: instead of alphabetizing the file by author name, the file was simply sorted by the names of the journals in which papers were published. When this exercise was first performed in the early 1960s, we discovered that the journals already covered in *Current Contents* included those that either produced the most papers or those that were cited the most. But a simple method was needed to compare large journals like *Nature*, *Science*, and *JAMA* with smaller journals like the Annual Reviews. In the early days of *Current Contents*, we had emphasized the fields of molecular biology and biochemistry. We observed that 25% of all citations in the current year's literature were of papers that were only 2–3 years old, so it was decided to use the prior two cited years as the basis for calculating a current-year impact factor, that is, the average number of citations per published paper. However, we also recognized that smaller but important review and specialty journals might not be selected if we depended solely on total publication or citation counts [1]. The journal "impact factor" was created as a method for comparing journals regardless of their size or citation frequency. Figure 2 compares three tables in which life science journals sorted by: (A) most-cited journals in 2008 (B) number of articles published in 2008, and (C) impact factor for 2008, with the consequent appearance of small review journals for the first time.

The term "impact factor" has gradually evolved, especially in Europe, to describe both journal and author impact. This ambiguity often causes problems, since it is one thing to use impact factors to compare journals and quite another to use them to compare authors. Whereas an individual author normally produces a small number of articles on average (although there are some phenomenally productive ones), journal impact factors generally involve relatively large populations of articles and citations. A journal's impact factor is based on two elements: the numerator, which is the number of cites in the current year of any items published in the journal during the previous 2 years; and the denominator, the number of substantive articles (source items) published during 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 or, take into account longer periods of citations and/or sources, but the measure would be less current. It is important to note that correspondence, letters, news stories, obituaries, editorials, interviews, and tributes are not included in *JCR*'s calculation of source items. Nevertheless, since the numerator includes citations to these more ephemeral items, some distortion will result. Ordinarily though, just a small number of journals are affected, and out of those, the effect implies a change of only 5–10% [2]. Also nowadays, the *JCR* includes every citation that appears in the 5000 plus journals that it covers; therefore, discussions of sampling errors in relation to *JCR* are not particularly meaningful.

## Scientometrics and Journalology

Citation analysis has blossomed over the past three decades into the field of scientometrics, which now has both its own International Society for Scientometrics and Informetrics (ISSI, http://www.issi-society.info) and the journal Scientometrics, which was started in 1978. Over 15 years ago, Steve Lock aptly named the application of scientometrics to journal evaluation "journalology" [3].

All citation studies should be normalized to take into accounts variables such as discipline, half life, and citation density [4]. The half-life (number of retrospective years required to find 50% of the cited references) is longer for a physiology journal than that for a physics journal. For some fields, *JCR*'s 2-year-based impact factors may or may not give as complete a picture as would a 5- or 10-year period. Nevertheless, when journals are studied within disciplinary categories, the rankings based on 1-, 7- or 15-year impact factors do not differ significantly [5]. In other words, when journals were studied across fields, the ranking for physiology journals improved significantly as the number of years increased, but the rankings within the physiology category did not change significantly. The citation density is the average number of references cited per source article. Citation density (R/S) is significantly lower for mathematics journals than for molecular biology journals. There is a widespread but mistaken belief that the size of the scientific community that a journal serves significantly affects the journal's 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 in

**(A)**

| Abbreviated Journal Title | Total Cites | Articles | Impact Factor |
|---|---|---|---|
| Nature | 443,967 | 899 | 31.434 |
| P Natl Acad Sci USA | 416,018 | 3508 | 9.380 |
| Science | 409,290 | 862 | 28.103 |
| J Biol Chem | 407,492 | 3761 | 5.520 |
| J Am Chem Soc | 318,252 | 3242 | 8.091 |
| Phys Rev Lett | 310,717 | 3905 | 7.180 |
| Phys Rev B | 250,465 | 5782 | 3.322 |
| New Engl J Med | 205,750 | 356 | 50.017 |
| Appl Phys Lett | 179,925 | 5449 | 3.726 |
| Astrophys J | 177,571 | 2128 | 6.331 |
| J Chem Phys | 164,492 | 2763 | 3.149 |
| Lancet | 148,106 | 289 | 28.409 |
| Circulation | 143,852 | 607 | 14.595 |
| Cell | 142,064 | 348 | 31.253 |
| Angew Chem Int Edit | 139,534 | 1797 | 10.879 |
| J Geophys Res | 129,836 | 2860 | 3.147 |
| Cancer Res | 125,341 | 1228 | 7.514 |
| J Immunol | 123,910 | 1860 | 6.000 |
| Blood | 122,032 | 1237 | 10.432 |
| J Neurosci | 120,933 | 1438 | 7.452 |

**(B)**

| Abbreviated Journal Title | Total Cites | Articles | Impact Factor |
|---|---|---|---|
| J Biol Chem | 407,492 | 3761 | 5.520 |
| P Natl Acad Sci USA | 416,018 | 3508 | 9.380 |
| J Am Chem Soc | 318,252 | 3242 | 8.091 |
| J Chem Phys | 164,492 | 2763 | 3.149 |
| J Immunol | 123,910 | 1860 | 6.000 |
| Angew Chem Int Edit | 139,534 | 1797 | 10.879 |
| J Agr Food Chem | 51,062 | 1670 | 2.562 |
| J Neurosci | 120,933 | 1438 | 7.452 |
| Biochemistry US | 94,645 | 1437 | 3.379 |
| Org Lett | 46,502 | 1403 | 5.128 |
| Nanotechnology | 16,291 | 1397 | 3.446 |
| Macromolecules | 80,559 | 1379 | 4.407 |
| J Virol | 86,021 | 1293 | 5.308 |
| Blood | 122,032 | 1237 | 10.432 |
| Cancer Res | 125,341 | 1228 | 7.514 |
| Brain Res | 56,664 | 1187 | 2.494 |
| Zootaxa | 2639 | 1118 | 0.740 |
| World J Gastroentero | 10,822 | 1112 | 2.081 |
| Neurosci Lett | 28,223 | 1080 | 2.200 |
| Nuclei Acids Res | 86,787 | 1070 | 6.878 |

**(C)**

| Abbreviated Journal Title | Total Cites | Articles | Impact Factor |
|---|---|---|---|
| CA-Cancer J Clin | 7522 | 19 | 74.575 |
| New Engl J Med | 205,750 | 356 | 50.017 |
| Annu Rev Immunol | 15,519 | 24 | 41.059 |
| Nat Rev Mol Cell Bio | 19,628 | 84 | 35.423 |
| Physiol Rev | 17,865 | 40 | 35.000 |
| JAMA-J Am Med Assoc | 114,250 | 225 | 31.718 |
| Nature | 443,967 | 899 | 31.434 |
| Cell | 142,064 | 348 | 31.253 |
| Nat Rev Cancer | 18,908 | 85 | 30.762 |
| Nat Genet | 61,812 | 215 | 30.259 |
| Annu Rev Biochem | 16,889 | 31 | 30.016 |
| Nat Rev Immunol | 15,775 | 86 | 30.006 |
| Nat Rev Drug Discov | 10,062 | 62 | 28.690 |
| Lancet | 148,106 | 289 | 28.409 |
| Science | 409,290 | 862 | 28.103 |
| Nat Med | 48,632 | 141 | 27.553 |
| Annu Rev Neurosci | 10,132 | 23 | 26.405 |
| Nat Rev Neurosci | 15,642 | 71 | 25.940 |
| Nat Immunol | 25,245 | 133 | 25.113 |
| Cancer Cell | 12,985 | 78 | 24.962 |

**Fig. 2.** Life science journals sorted by (**A**) most-cited journals in 2008 (**B**) number of articles published in 2008, and (**C**) impact factor for 2008.

most fields are not well cited, whereas some articles in small fields may have unusual impact, especially when they have cross-disciplinary repercussions.

It is well known that there is a skewed distribution of citations in most fields; according to the 80/20 rule, 20% of the articles may account for 80% of the citations. To reiterate, the average number of citations per article and the immediacy of the citations—and not the number of authors or articles in the field—are the significant elements [6]. The size of the field, however, will generally 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. Review papers are generally cited about twice as often as other papers, but publishing mediocre review papers will not necessarily boost your journal's impact. Just as a curiosity, Table 2 shows a short list of super-cited papers in the life sciences. The Lowry paper was recently discussed in *Journal of Biological Chemistry* [7], but the authors failed to mention Lowry's own commentary on this, the most-cited paper in the history of science, where he himself noted that it was not his most important paper [8].

The skewness of citations is repeated as a mantra by critics of the impact factor. Some editors would like to see impacts calculated solely on the basis of the most-cited papers so that otherwise low-impact factors can be ignored. However, since most journals experience this skewness, this should not significantly affect journal rankings. Others would like to see rankings by geographic area because of *SCI*'s alleged English-language bias. Europhiles would like to be able to compare their journals by language or geographic groups, especially in the social sciences and humanities. The time required to referee manuscripts may also affect impact: if manuscript processing is delayed, references to articles that are no longer within the *JCR* 2-year window will not be counted [9,10]. Alternatively, the appearance of articles on the same subject in the same issue of a journal may have an upward effect. 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 [11].

Other objections to impact factors are related to the system used in *JCR* to categorize journals. In a perfect system it ought to be possible to compare journals with identical profiles. But in fact, there are rarely two journals with identical semantic or bibliographic profiles. *ISI*'s heuristic, somewhat subjective methods for categorizing journals are by no means perfect, even though their specialists do use citation analysis to support their decisions. There have been recent attempts to group journals more objectively, relying on two-way citational relationships between journals to reduce the subjective influence of journal titles [12], e.g., citation analysis proved that the *Journal of Experimental Medicine* was a leading immunology journal, and even nowadays continues to be among the top five immunology journals, based on its impact factor. *JCR* recently added a new feature that provides users with the ability to more pre-

**Table 2.** Most cited papers in the life sciences, through July 2005

| Authors | Title | Source | Year | Volume | Page | Hits |
|---|---|---|---|---|---|---|
| Lowry OH, Rosebrough NJ, Farr AL, Randall RJ | Protein Measurement with the Folin Phenol Reagent | Journal of Biological Chemistry | 1951 | 193 | 265 | 293,328 |
| Laemmli UK | Cleavage of Structural Proteins During Assembly of Head of Bacteriophage T4 | Nature | 1970 | 227 | 680 | 192,022 |
| Bradford MM | Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding | Analytical Biochemistry | 1976 | 72 | 248 | 120,179 |
| Sanger F, Nicklen S, Coulson AR | DNA Sequencing with Chain-Terminating Inhibitors | Proceedings of the National Academy of Science USA | 1977 | 74 | 5463 | 63,909 |
| Chomczynski O, Sacchi N | Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate Phenol Chloroform Extraction | Analytical Biochemistry | 1987 | 162 | 156 | 55,987 |
| Towbin H, Staehelin T, Gordon J | Electrophoretic Transfer of Proteins from Polyacramide Gels to Nitrocellulose Sheets – Procedure and Some Applications | Proceedings of the National Academy of Science USA | 1979 | 76 | 4350 | 48,671 |
| Folch J, Lees M, Stanley GHS | A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues | Journal of Biological Chemistry | 1957 | 226 | 497 | 35,646 |
| Southern EM | Detection of Specific Sequences among DNA Fragments Separated by Gel Electrophoresis | Journal of Molecular Biology | 1975 | 98 | 503 | 31,273 |

cisely establish journal categories based on citation relatedness. Figure 3 shows the general formula for calculating citation relatedness between two journals and the relatedness coefficient expressing the average of the maximum and the minimum. This coefficient reflects how often a journal cites and is cited by each of the journals it is compared to, and it takes into account the sizes of the journals involved (papers published) as well as the number of times each journal cites the other. However, using the *JCR* relatedness method, some journals can be assigned to different *JCR* categories. Using *Circulation* as an example of the journal with the highest impact factor representing cardiology, we find that *NEJM* (a general medicine journal) ranked 7th among the most related journals in this field. Heretofore, one could only guess at the proximity of *NEJM* to this or other topics.

Many discrepancies with journal impact factors are eliminated altogether in another *ISI* database called the *Journal Performance Indicators* (*JPI*, found at http://scientific.thomson.com/products/jpi). This annual compilation covers the period 1981 through the current year. Unlike *JCR*, this database links each source item to its own unique citations, making impact calculations more precise. Only citations to the substantive items are counted in the denominator, and it is possible to obtain cumulative impact measures covering longer time spans. In an analysis of the period 1999–2004, Table 3 shows how 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. In this year, *JAMA* published 1905 items, out of which 630 were letters and 253 were editorials; citations to these items were not included in the *JPI* calculation of impact. In spite of the alleged distortions introduced by counting citations to all "editorial" material in *SCI*, a report by González and Campanario, researchers at the University of Alcalá, demonstrated that the effect, if any, is quite minor [13].

**CALCULATING RELATEDNESS COEFFICIENT OF JOURNAL$_1$ AND JOURNAL$_2$**

$$R_{1>2} = \frac{C_{1>2} \times 10^6}{Ref_1 \times Pap_2}$$

$$R_{1<2} = \frac{C_{1<2} \times 10^6}{Ref_2 \times Pap_1}$$

$$R_{coeff} = \sqrt{R_{1>2} \times R_{1<2}}$$

C = Citations
$Ref_1$ is the number of references cited in Journal 1.
$Pap_2$ is the number of papers published by Journal 2.
$Ref_2$ is the number of references cited in Journal 2.
$Pap_1$ is the number of papers published by Journal 1.

**Fig. 3.** General formula for calculating citation relatedness between two journals and the relatedness coefficient expressing the average of the maximum and the minimum.

In addition to helping libraries decide which journals to purchase, journal impact factors are also used by authors to decide where to submit their articles. As a general rule, the journals with high impact factors also include the most prestigious, although the perception of prestige is a murky subject. Librarians argue that the numerator in the impact-factor calculation is itself even more relevant. Bensman argued that this 2-year total citation count is a better guide to journal significance and cost-effectiveness than is the impact factor [14]. Journal impact can also be useful in comparing expected and actual citation frequency. Thus, when *ISI* prepares a personal citation report it provides data on the expected citation impact not only for a particular journal but also for a particular year, because impact factors can change from year to year.

The use of journal impact factors instead of actual article citation counts to evaluate individuals is a highly controversial is-

**Table 3.** JPI data on *JAMA*. Citation impact (all items) in one year periods, 1981 to 2004 [from 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 |
| 11 | 1991 | 47.19 | 30,389 | 644 |
| 12 | 1992 | 58.48 | 34,389 | 588 |
| 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 | 7311 | 383 |
| 24 | 2004 | 3.34 | 1174 | 351 |

$$\frac{31,257}{370} \qquad \frac{\text{Citacions received 1999-2004}}{\text{Articles published in JAMA in 1999}} = 84.5$$

sue. Granting and other policy agencies often wish to bypass the work involved in obtaining actual citation counts for individual articles and authors. Since recently published articles may not have had enough time to be cited, it is tempting to use the journal impact factor 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 recent bibliography is examined, the impact factors of the journals involved are substituted in lieu of the actual citation count. This practice began about a decade ago, when administrators decided they would estimate the future impact of a recently published paper by incorporating the impact factor for the journal in which the paper is published (for younger scientists especially, many of the papers listed in their C.V. have been published often during the period used to calculate impact, and most of them will not be cited for a few years or more, depending upon the rate at which research on their topic progresses). Thus, the impact factor is used to estimate the expected influence of individual papers, a rather dubious practice taking into account the aforementioned skewness observed for most journals. Today, "webometrics" are increasingly brought into play, although there is little evidence that this 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 research papers. Nevertheless, some studies seem to indicate that web sitation is a harbinger of future citation.

The assumption that the impact of recent articles cannot be evaluated in *SCI* is not universally correct. While there may be several years delay on 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 enable *ISI* 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, yet most of these papers will eventually go on to become citation classics (http://www.citationclassics.org).

Of the many conflicting opinions about impact factors, Hoeffel 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 higher 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." [15]

Obviously, a better evaluation system would involve actually reading each article for quality, but even then difficulties of reconciling peer review judgments would arise. When it comes to evaluating faculty, most people do not have or care to take the time any more. Even if they did, their judgment would surely be tempered by observing the comments of those who have cited the work; this is known as citation context analysis. Fortunately, in the near future, full-text capabilities in the web will make this a more practical task to perform.

## References

[1] Brodman E (1944) Choosing physiology journals. Bull Med Library Association 32:479-483
[2] Garfield E (1972) Citation analysis as a tool in journal evaluation. Science 178: 471-479. Reprinted in Current Contents no. 6 pp. 5-6 (1973); Essays of an Information Scientist (1977) 1:527-544
[3] Lock SP (1989) Journalology: Are quotes needed? CBE Views 12:57-59
[4] Pudovkin AI, Garfield E (2004) Rank-normalized Impact Factor: A way to compare journal performance across subject categories. Annual Meeting, Providence, RI. November 17, 2004. Published in Proceedings of the 67th Annual Meeting of the American Society for Information Science & Technology, 41:507-515
[5] Garfield E (1998) Long-term vs. short-term journal impact: Does it matter? (part II) The Scientist 12:12-13
[6] Garfield E (1976) Is the ration between number of citations and publications cited a true constant? Current Contents no. 6 pp 5-7. Reprinted in Essays of an Information Scientist (1977) ISI Press, Philadelphia, 2:419-421
[7] Kresge N, Simoni RD, Hill RL (2005) The most highly cited paper in publishing history: Protein determination by Oliver H. Lowry. J Biol Chem 280: e25
[8] Lowry OH (1977) Protein measurement with folin phenol reagent. Current Contents/Life Sciences 1:5-7
[9] Yu G, Wang X-Y, Yu D-R (2005) The influence of publication delays on impact factors. Scientometrics 64:235-246
[10] Garfield E (1998) Long-term vs. short-term journal impact: Does it matter? The Scientist 12:10-12
[11] Garfield E (1986) Which medical journals have the greatest impact? Annals of Internal Medicine 105(2), 313-320 (Available in: http://garfield.library.upenn.edu/essasys/v10p007y1987.pdf)
[12] Pudovkin AI, Garfield E (2002) Algorithmic procedure for finding semantically related journals. Journal of the American Society for Information Science & Technology (JASIST) 53:1113-1119
[13] González L, Campanario M (2006) Structure of the Impact Factor of journals included in the Social Sciences Citation Index: Citations from documents labeled "Editorial Material". J Am Soc Infor Sci Technol 58:252-262
[14] Bensman SJ, Wilder SJ (1998) Scientific and technical serials holdings optimization in an inefficient market: a LSU serials redesign project exercise. Library Resources & Technical Services 42:147-242
[15] Hoeffel C (1998) Journal Impact Factors (letter). Allergy 53:1225

## About the author

**Eugene Garfield** is Chairman Emeritus of Thomson Scientific, formerly the Institute for Scientific information.® He is the inventor of *Current Contents*®, *Index Chemicus*®, *Arts & Humanities Citation Index*®, *Journal Citation Reports*®, *Science Citation Index*® and *Social Sciences Citation Index*,® and all now available on the *ISI Web of Knowledge*SM. These last two databases–*SCI*® and *SSCI*®–have provided an objective and quantitative basis for analyzing information flows in scientific communication and fostered the growth of the field of scientometrics. Garfield received a BS in Chemistry and an MS in Library Science from Columbia University, and a Ph.D. in Structural Linguistics from the University of Pennsylvania, where he serves on the Board of Overseers of the Library. His scientific information career began in 1951 when he joined the Welch Medical Indexing Project at Johns Hopkins University, sponsored by the National Library of Medicine. The Welch Project planted the seeds for several major advances in scientific communication and information science that have distinguished his career in information discovery and recovery. He is President and Founding Editor of *The Scientist*,® a monthly magazine for the life sciences professional, available on the web since 1986 (www.the-scientist.com). He also serves on the editorial boards of *Scientometrics* and the *Journal of Information Science*. He is a Past President of the American Society for Information Science and Technology. He is the author of numerous papers on information science. He has lectured widely throughout the world. He is author of the 15-volume series *Essays of an Information Scientist* as well as *Citation Indexing--Its Theory and Application to Science, Technology, and Humanities*, available online at www.eugenegarfield.org.