

## The theory and uses of citation indexing

*Hassan A. El-Munshid, PhD, MD.*

Many colleagues may have come across a mention of the Science Citation Index (SCI) and some might have used it. However, the SCI is such an important tool for getting useful and interesting information from the scientific literature that knowledge of its underlying theory and applications in science, including Biomedicine, is vital for understanding the full range of its uses as well as its limitations. The SCI was pioneered by Eugene Garfield.<sup>1</sup> Garfield is the initiator of Current Contents and the founder and previous Chairman of the Institute for Scientific Information (ISI) in Philadelphia, USA. He also introduced the Social Sciences Citation Index (SSCI) and the Arts and Humanities Citation Index. The SCI was compiled in 1961 and first appeared in print in 1964. In 1988 it became available as CD-ROM. It is currently available on-line under the ISI Web of Science.

Almost all scientific publications for example research articles, reviews, letters, editorials, comments, and so forth cite linked published documents. A citation index is simply built around these links. We are all familiar with the traditional subject index such as Index Medicus but there are certain problems<sup>1</sup> entailed in preparing and using a subject index: 1. specialized knowledge is required to determine subject headings, 2. production of a subject index is expensive and time consuming, which reduces the timeliness of the index, and 3. with the passage of time, word usage may change for example new terms may replace old ones and this may complicate searches of the literature. In comparison, a citation index has the following advantages: 1. By using authors' citations rather than subject headings, a citation index requires less time and expense, 2. Citation indexing can still be used when documents are written in different languages, 3. A citation index can be used not only to search the literature backwards in time but also forwards namely by getting up-to-date citations to a paper published some years back, and 4. Citation links can identify cross-disciplinary developments since authors consistently cite papers outside their own discipline. As first pointed out by Bradford,<sup>2</sup> a specialist in a certain field needs to read articles published not only in journals on his own specialty but also journals on other disciplines. Several subsequent studies confirmed this observation and a consensus emerged that between 500-1,000 different journals are required to obtain 95% of the literature in a given discipline. On the face of it, that means to

produce an index, which covers all disciplines one needs to multiply the number of disciplines by 500-1,000 to get the number of journals to be indexed, an impractical proposition. But since there is considerable overlap between disciplines this situation does not actually arise, and Garfield<sup>1,3</sup> used the 1969 SCI to show that covering the core literature for all scientific disciplines requires no more than 1,000 journals. Originally the SCI covered approximately 3,000 journals. The SCI expanded is an enlarged version of the SCI that indexes 5,900 journals covering 150 scientific disciplines. It is available under the ISI Web of Science, for which a library subscription is required. The ISI Web of Science also includes the SSCI, which covers approximately 1,700 journals and more than 50 disciplines, and the Arts and Humanities Citation Index covering nearly 1,130 journals. The ISI Web of Science thus provides access to current and retrospective bibliographic information, author abstracts and cited references in approximately 8,730 journals. In addition, it includes the ISI Journal Citation Reports (JCR), which entail analysis and evaluation of over 7,000 science and social sciences journals. A wealth of information is available when the ISI Web of Science is accessed. To retrieve articles on a specific research topic, a subject term may be used and the search may be expanded by using key words derived from an article's cited references (Key Word Plus). When a paper or review some years old is available, the literature can be searched backwards and also forwards in time by finding the articles that cited the available article up to the present. An article may be retrieved by using the author's name, journal title or author's affiliation. By using the address of an institution the articles published by the authors affiliated to it over a certain period of time can be obtained for example 50 articles were published from King Faisal University in 2003. The number of publications from a whole country can also be obtained for example 1,588 articles were published from Saudi Arabia in 2002. The citations to an author, whether as first author or coauthor, over a certain period of time up to the present can be obtained. It is provided according to a format (**Table 1**), in which the citations to the author of this article over the period 2000-2004 are given. A hit is a citation and the first column gives the total citations irrespective of the year selections. The last column has been added.

Using citation counts to evaluate scientists and subsequently comparing between individual scientists or applicants for an academic post have incited a lot of discussion. To begin with, approximately 25% of papers are never cited and the average annual citation count for cited papers is only 1.7 times.<sup>1</sup> Another problem is how to apportion citation credit

Table 1 - Example of citations to an author over the period of 2000-2004.

Hits	Cited author	Cited work	Volume	Page	Year	Citations 2000 2004
1	Elmunshid HA	Curr Ther Res Clin E	55	1263	1994	1 in 2003
70	Elmunshid HA	J Physiol-London	298	557	1980	1 in 2000 2 in 2002

in multiauthored papers. Recently Romesburg<sup>4</sup> proposed that in the case of a cited coauthored paper the citation credit for each author should be obtained by dividing the total citations to the paper by the number of authors, rather than crediting each author with the total citations to the paper. For a comparison to be fair it should be between peers for example scientists in the same discipline; comparing between a mathematician and a physiologist for instance is not appropriate since the biomedical literature is much larger than that of mathematics with a larger chance of citation for the physiologist. At this juncture it would be appropriate to ponder what citation counts actually assess. Most probably citations counts provide a measure of the utility or, in a wider sense, the impact of the work of a scientist. As such, citation counts should supplement, not replace, analysis of content and peer review.<sup>1</sup> The validity of citation counts as a measure of scientific merit has been demonstrated by Garfield.<sup>1</sup> He conducted citation studies of Nobel Prize laureates over a number of years and found that their citation counts were much higher than the average. Conversely, lists of highly cited scientists readily picked Nobel Prize winners and those elected to the national academies of sciences. Citation counts are also used to detect highly active or hot areas of research, and highly cited papers in a certain year are usually referred to as hot papers. Furthermore, retrospective citation links and counts have been used as tools for historical research into science.<sup>1</sup>

Citation analysis of scientific journals is included in the ISI Journal Citation Reports (JCR) under the ISI Web of Science. To objectively assess the utility or prestige of scientific journals, the following measurements have been developed:<sup>5</sup> 1. Impact factor is the number of times during the current year a journal was cited in the previous 2 years divided by the total number of articles published by the journal in the same 2 years. 2. Immediacy index is the number of times a journal was cited within the year of publication divided by the total number of articles published by the journal in that year. 3. Cited half-life is the number of years, going back from the current year, that covers 50% of the

Table 2 - A group of journals from the Institute for Scientific Information Journal Citation Reports (2002).

Journal	Impact factor	Immediacy index	Cited half-life
Nature	30.432	7.504	6.9
Nature Medicine	28.740	6.912	4.0
Nature Genetics	26.711	4.811	4.4
Science	28.956	5.839	6.6
N Engl J Med	31.736	8.138	7.2
Lancet	15.397	5.299	6.9
Ann Intern Med	11.414	3.028	8.7
Gastroenterology	13.440	2.429	6.8
Gut	6.323	0.969	6.4
Scand J Gastroenterol	1.847	0.242	7.7
Hepatogastroenterol	0.833	0.031	4.6
Acta Hepato-Ent Belg	0.636	0.111	4.8

citations in the current year to the journal. **Table 2** shows these measurements for 2002 concerning a group of journals that may interest a clinician or a medical physiologist specialized in gastroenterology. Concentrating mainly on the impact factor, it is clear that in general the prestigious multidisciplinary Nature and Science head the list followed by general medical journals and then the specialty journals. The measurements, especially the impact factor, may be used to: 1. Provide objective basis for acquiring journals by a library, and 2. Assist scientists in deciding to which journal a manuscript should be submitted.

In conclusion, the SCI and the SSCI are more than tools for searching the scientific literature. They can also be used for evaluation of scientists and institutions, for improvement of decisions on library management and for studies on the history and evolution of science.

Acknowledgments. I am grateful to Mr. Shoaib A. Quraishi, Librarian, King Fahd University for Petroleum and Minerals Library, for help with navigating through the ISI Web of Science. I am also grateful to Mr. P. Syed Mohamed for typing the manuscript.

Received 19th March 2005. Accepted for publication in final form 5th June 2005.

From the Department of Physiology, College of Medicine, King Faisal University, Dammam, Kingdom of Saudi Arabia. Address correspondence and reprint requests to Dr. Hassan A. El-Munshid, Professor, King Faisal University, PO Box 2114, Dammam 31451, Kingdom of Saudi Arabia. Tel. +966 (3) 8577000 Ext. 2322. Fax. +966 (3) 8575329. E-mail: munshid@yahoo.com

## References

1. Garfield E. Citation Indexing-Its Theory and Application in Science, Technology, and Humanities. A Wiley-Interscience Publication. New York: John Wiley and Sons, Inc; 1979.
2. Bradford SC. Documentation. 2nd ed. London: Lockwood; 1953.
3. Garfield E. Citation analysis as a tool in journal evaluation. *Science* 1972; 178: 471-479.
4. Romesburg HC. If the NFL can do it, so can scientists. *The Scientist* 2003; 17 (20): 60.
5. Garfield E. Use of Journal Citation Reports and Journal Performance Indicators in measuring short and long term journal impact. *Croat Med J* 2000; 41: 368-374.

---

## Analysis of hepatitis C virus core antigenemia in Saudi drug users

*Alhusain J. Alzahrani, BSc, MSc, PhD.*

**H**epatitis C virus (HCV) occurs worldwide with an estimated 170 millions people infected.<sup>1</sup> The majority of infected individuals progress to chronic infection, which can lead to cirrhosis or hepatocellular carcinoma. Enzyme-linked immunosorbent assay (ELISA) is the principal diagnostic test to detect anti-HCV antibodies in blood, recombinant immunoblot assay (RIBA) is used as supplemental assay. Only direct detection of viral RNA or antigens can differentiate ongoing from resolved infection and can help in the follow up after therapy. Therefore, a number of qualitative and quantitative assays for detection of HCV RNA have been developed. Recently, a new assay capable of detecting and quantifying HCV core antigen has been developed. This new assay incorporates at least 2 modifications leading to a theoretically improved sensitivity: 1. an immune complex dissociation step that allows detection of both free and antibody-bound core antigen, and 2. an improved conjugate reagent that provides superior signal amplification.<sup>2</sup>

Screening of blood and blood products for anti-HCV antibodies and HCV RNA has reduced the incidence of post-transfusion HCV infection to negligible levels. However, drug users remain at high risk of acquiring HCV infection. Intravenous drug use (IVDU) was found to be responsible for

approximately 60% of the new cases of HCV infection.<sup>1</sup> The prevalence of antibodies to hepatitis C virus (anti-HCV) in a population of IVDU in Jeddah, Kingdom of Saudi Arabia (KSA) was reported as 74.6%, whereas the prevalence of anti-HCV in drug dependent patients who did not use the intravenous route was 10.5%.<sup>3</sup> On the other hand, Shawky et al<sup>4</sup> reported that high prevalence of hepatitis C virus infection that amounted to 63.9% in drug addicts from the same area of the country. Similarly Fathalla et al<sup>5</sup> found that the prevalence of HCV antibodies among drug users in the Eastern Province of KSA reached only 6.5%. The objective of the present study was to assess HCV core antigen in sera from Saudi drug users and to compare this with anti-HCV antibodies and HCV RNA data from the same patients.

All patients (n=201) who are enrolled in drug rehabilitation program in a drug rehabilitation hospital over a period of one year (October 2003 to September 2004) were included in this study upon their written informed consent. Anti-HCV antibodies were detected by a third-generation ELISA (AxSYM HCV version 3 [Abbott Diagnostics, Chicago, Illinois]) and HCV 3 ELISA test system [Ortho-Clinical Diagnostics, Raritan, New Jersey]) and by a third-generation RIBA (RIBA HCV 3; Ortho-Clinical Diagnostics).<sup>2</sup> A simple enzyme immunoassay for detection and quantification of total HCV core Ag has been recently developed (Ortho-Clinical Diagnostics), and this test was used according to the manufacturer's recommendations described previously.<sup>2</sup> Results were expressed in picograms per milliliter, with a limit of detection established by the manufacturer at 1.5 pg/ml. To detect HCV RNA, the COBAS AMPLICOR HCV test version 2.0 (Roche Molecular Systems, Branchburg, New Jersey) was used according to the manufacturer's instructions. This assay has a limit of detection of 2.0 log<sub>10</sub> IU of HCV genotype 1 RNA per ml of serum.

The study population was 201 Saudi drug users, males, and of mean age of 33 years. Among this cohort of patients 82 were IVDU and 119 non-IVDU. The mean duration of drug use was 10 years. The seroprevalence of HBsAg in the study population was 5.9%, the seroprevalence of HCV antibodies was 35.6%, and seroprevalence of HIV antibodies was 0.99% as determined earlier in our laboratory. Out of the tested 201 samples 68 (33.8%) yielded positive results for HCV core Ag. Among the 68 samples positive for HCV core Ag, 54 (79%) had detectable anti-HCV antibody, and only 14 (21%) were anti-HCV antibody negative. On the other hand, among those 68 samples positive for HCV core Ag, 60 samples (88%) had detectable HCV-RNA by qualitative reverse transcription-