The Evolution of the Science Citation Index Search Engine to the Web of Science, Scientometric Evaluation and Historiography

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 the University of Barcelona
January 24, 2007

The *Science Citation Index* was proposed over fifty 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 including its rankings by impact factor. Impact factors were not widely adopted 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. The inventor of the *SCI* and its companion *Social Sciences Citation Index* will review its history and discuss their more recent use in graphically visualizing micro-histories of scholarly topics. Using the patented *HistCite* software for algorithmic historiographic analysis, 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 will be discussed.

It is now over fifty years since the idea of the *Science Citation Index* (*SCI*) was first promulgated in *Science* magazine in 1955.[i] However, as the older generation of scientists will remember *Current Contents* is the information service that proved to be the primordial revolutionary "idea" which made the 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.

It is ironic that most historians ignore its significance and impact. I believe the main reason for this is its utter simplicity. They overlook how important timing was to its success. Even aficionados of *Current Contents* may not remember the role that the early electronic computer made it possible for *Current Contents* to appear each week together with its title word indexes and author address directory. In those days, conventional indexes were six months to three years behind the literature. At one time it was estimated that ten million worldwide reprint requests were generated each year because of *Current Contents.*

Historians are not impressed by simplicity. They are attracted to the complexity of the *Science Citation Index* and other search engines. Nevertheless, the success of *SCI* as and up-to-date tool for the dissemination and retrieval of information does not fully account for its impact.

The *SCI's* use as a tool in measuring scientific productivity has often overshadowed its primary function as a search engine. Many people think that bibliometrics is its main reason for existing. Although *SCI* was used for scientometrics studies shortly after it appeared, the advent of the *SCI* Journal Citation Reports and its Impact Factor rankings brought it into even greater prominence. This is reflected in the fact that my 1972 *Science* paper on the use of citation analysis in journal evaluation is my most-cited work.[ii]

Multi-Disciplinary Database

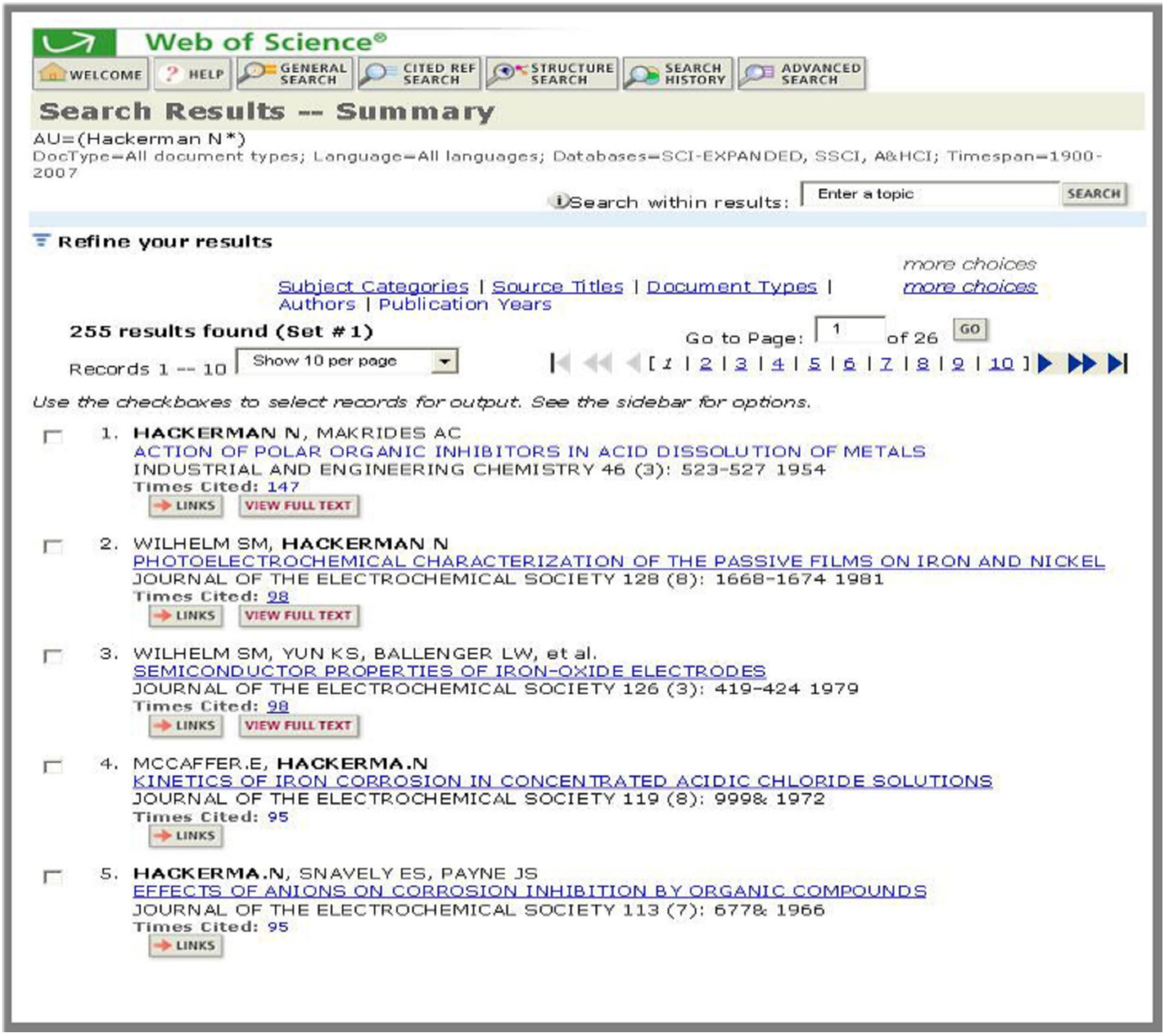
From the launch of the *SCI* multi-disciplinary database I pointed out its two-fold purpose in identifying what each scientist had published, and where and how often the papers by that scientist were cited. Hence, the *SCI* has always been divided into two author-based parts: the Source Author Index and the Citation Index. By extension, one can also determine what each institution and country has published and how often their papers are cited.

There are many ways to determine what an author has published, not the least of which is to obtain his or her C.V. However, it is remarkable how often such a complete list is difficult to find.

SLIDE 1: *WOS* SOURCE INDEX ENTRY FOR NORMAN HACKERMAN

Slide 1 illustrates how *SCI* in its electronic version called the *Web of Science*, links these two functions. An author's publications can be listed by chronology, by journal, or by citation frequency. In Slide 1, Norman Hackerman's five most-cited papers are shown. He has published 255 papers from 1936 to 2006. However, the sorting function in *WOS* permits me to rank his papers by citation frequency. [Do live search on WOS.]

Professor Hackerman is the president emeritus of Rice University in Texas. He is also a highly productive chemist who is now 94 years old. When we met recently, he asked if I knew of any other scientists who, like himself, had published over a seventy-year period. Since it is not possible to answer this question without some special programming, I posted a notice on an electronic bulletin board.



SLIDE 2: SCIENTISTS WHO HAVE PUBLISHED 70 TO 85 YEARS.

Slide 2 shows the resulting list of 16 scientists who have published for 70 to 85 years. The record is held by Isaac M. Kolthoff followed by Michael Heidelberger, a molecular biologist who published his first paper as an organic chemist in 1919. His last appeared in 2004, shortly before he died at the age of 104. There may be others I have not yet heard about. This list is presented here for the first time.

Let's segue to Heidelberger's work as a pioneering molecular biologist. Indeed, even before World War II, together with Oswald T. Avery and others at Rockefeller Institute, (including Colin M. Macleod and Maclyn McCarthy), he published some primordial work in the history of DNA. The work of Avery, Macleod, and McCarthy is a key link in the genealogical history of the Watson Crick 1953 paper on the double helix structure of DNA. I will return to this topic in a few moments.

Scientists who have published 70 years or more

Scientist	Birth/Death	Pub Years	Years Pub
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
Michel Eugene Chevreul (chemist)	1786-1889	1808-1889	82
Carl S. Marvel (polymer chemist)		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	1936-2006	71
Michael DeBakey (cardiac surgeon)	1908	1937-2006	70
Gerhard Herzberg (chemist)	1904-1999	1924-1992	69
Herman Mark (polymer chemist)	1895-1992	1922-1990	69
Ernst Mayr (biologist)	1904-2005	1923-2005	83

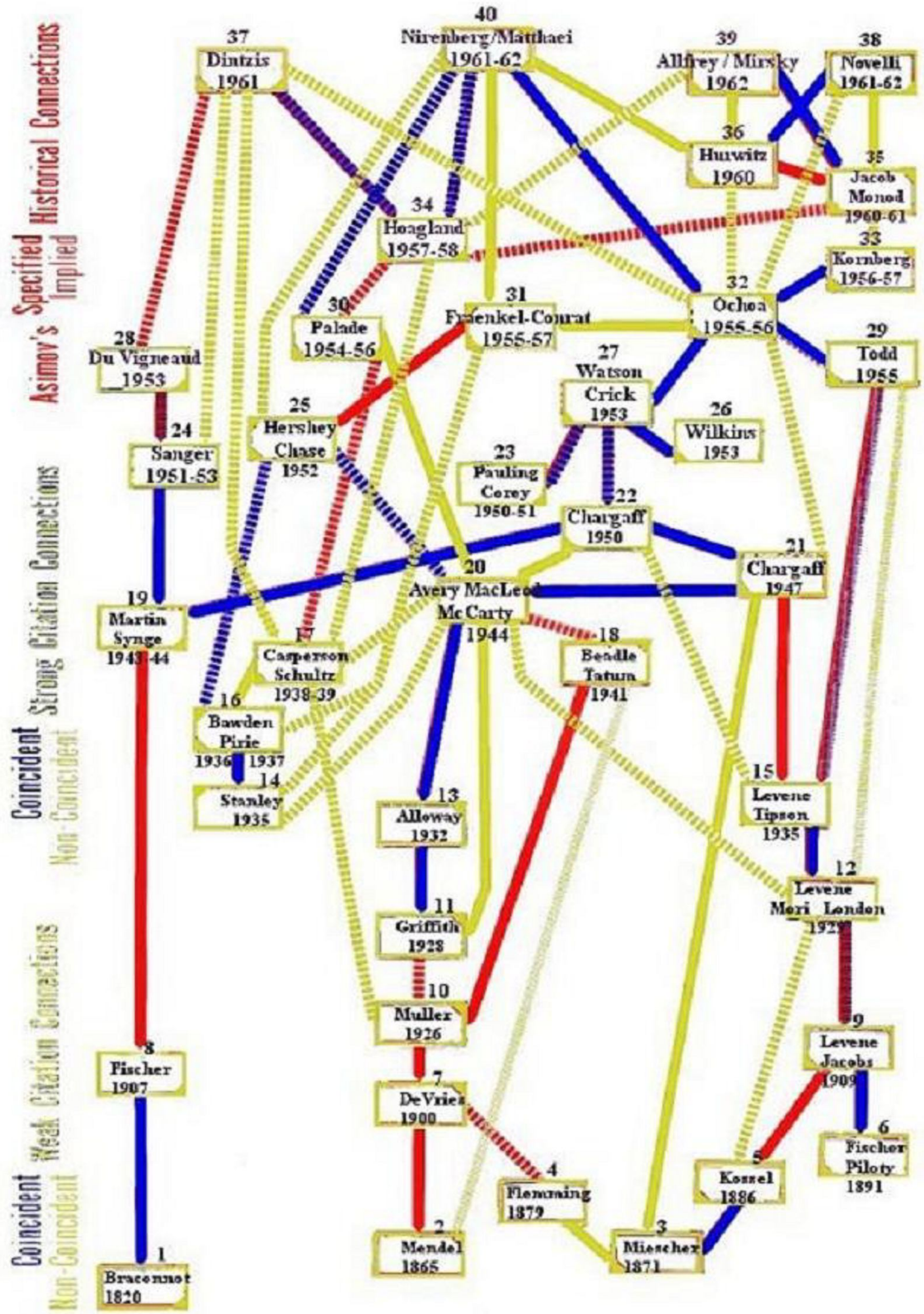
SLIDE 3: HISTORY OF DNA

When we launched the *Science Citation Index* in 1964, Irving Sher and I had already begun using bibliographic citations to create topological maps which we called Historiographs. To put it another way, we investigated whether citation indexes could aid in writing mini-histories of scientific topics. The outcome of that research was our 1964 report on "The Use of Citation Data in Writing the History of Science" which contained the graph shown in Slide 3.[iii] Keep in mind that this genealogical map was drawn manually once the nodes and links were identified in the 1961 *SCI*. Keep in mind that in those days we used IBM punched cards and primitive mainframe computers. I remember visiting Spain in the 1960s where we installed SCI tapes on an IBM 1401 computer with only12K memory.

More recently, gigabyte memories made it possible to write a program called *HistCite* which accepts the output of *a WoS* search and automatically generates historiographs. This patented software has been in development now for about five years and will be available commercially in a month or so. It has been used by dozens of evaluators for a variety of purposes. The process is boiled down to the following.

A topical, author, or citation based search is done on the *Web of Science*. From a few hundred to as many as ten thousand references are retrieved. These records are saved as a marked list on *WOS* and then exported to the *HistCite* software. Within minutes a series of tables and historiographs are generated.

Slide 3 :


Asimov's Specified Historical Connections
Implied
Coincident Strong Citation Connections
Non-Coincident
Coincident Weak Citation Connections
Non-Coincident
40 Nirenberg/Matthaei 1961-62
37 Dintzis 1961
39 Allfrey / Mirsky 1962
38 Novelli 1961-62
36 Hurwitz 1960
35 Jacob Monod 1960-61
34 Hoagland 1957-58
33 Kornberg 1956-57
32 Ochoa 1955-56
31 Fraenkel-Conrat 1955-57
30 Palade 1954-56
29 Todd 1955
28 Du Vigneaud 1953
27 Watson Crick 1953
26 Wilkins 1953
25 Hershey Chase 1952
24 Sanger 1951-53
23 Pauling Corey 1950-51
22 Chargaff 1950
21 Chargaff 1947
20 Avery MacLeod McCarty 1944
19 Martin Synge 1943-44
18 Beadle Tatum 1941
17 Casperson Schultz 1938-39
16 Bawden Pirie 1936 1937
14 Stanley 1935
15 Levene Tipson 1935
13 Alloway 1932
12 Levene Mori London
11 Griffith 1928
10 Muller 1926
9 Levene Jacobs 1909
8 Fischer 1907
7 DeVries 1900
6 Fischer Piloty 1891
5 Kossel 1886
4 Flemming 1879
3 Miescher 1871
2 Mendel 1865
1 Braconnot 1820

SLIDE 4: CITATIONS TO WATSON-CRICK

List of All Records Historiographs Glossary HistCite Guide About

Grand TLCS: 776, Grand TGCS: 17965

Collection span: 1944 - 1958

Citing *Molecular Structure of Nucleic Acids* 1953-1958

Articles from 1953-1958 citing Watson & Crick's 1953 paper, "Molecular Structure of Nucleic Acids" with selected outer references added to the collection

Records: 210, Authors: 262, Journals: 75, Cited References: 7057, Words: 611, Tags: 1

Yearly output | Document Type | Language | Institution | Institution with Subdivision | Country

View: Overview Sorted by publication date Page **1** of 3: [1 2 3]

#	LCR	NCR	Date / Author / Journal	LCS	GCS
			1944		
1	0	37	1 **AVERY OT, MACLEOD CM, MCCARTY M** *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type* III JRNL OF EXPERIMENTAL MEDICINE. 1944 FEB; 79 (2): 137-158	23	1450
			1952		
2	0	18	2 **HERSHEY AD, CHASE M** Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage JOURNAL OF GENERAL PHYSIOLOGY. 1952; 36 (1): 39-56	23	781
			1953		
3	2	20	3 FRANKLIN RE, GOSLING RG The Structure of Sodium Thymonucleate Fibres .1. The Influence of Water Content ACTA CRYSTALLOGRAPHICA. 1953; 6 (8-9): 673-677	11	269
4	3	11	4 **FRANKLIN RE, GOSLING RG** The Structure of Sodium Thymonucleate Fibres .2. The Cylindrically Symmetrical Patterson Function ACTA CRYSTALLOGRAPHICA. 1953; 6 (8-9): 678-685	8	82
5	1	7	5 **SMITH CL** The Breakdown of Desoxyribonucleic Acid Under Deuteron and Electron Bombardment ARCHIVES OF BIOCHEM AND BIOPHYSICS. 1953; 46 (1): 12-17	1	5
6	2	35	6 **WYATT GR, COHEN SS** The Bases of the Nucleic Acids of Some Bacterial and Animal	8	454

SLIDE 5: CHAINED CITATIONS TO WATSON-CRICK

List of All Records Historiographs Glossary HistCite Guide About
Grand TLCS: 3778, Grand TGCS: 70425
Collection span: 1938 - 1958

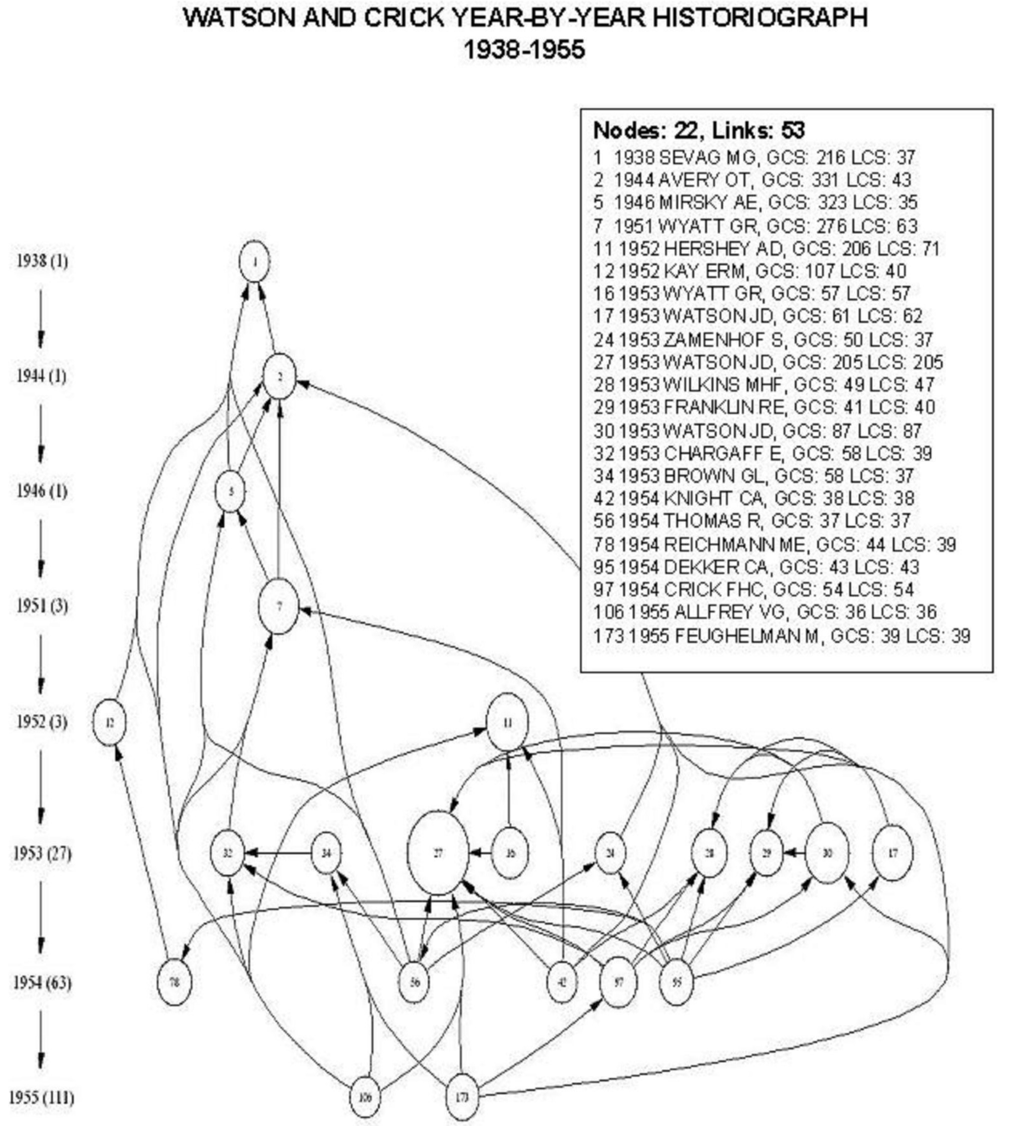
Chain Citing *Molecular Structure of Nucleic Acids* 1953-1958

Papers citing Watson and Crick's *Molecular Structure of Nucleic Acids* from 1953-58 and the papers citing this group 1953-58 and the top most outer references added to the collection. Publication month added to several papers from 1953-54.

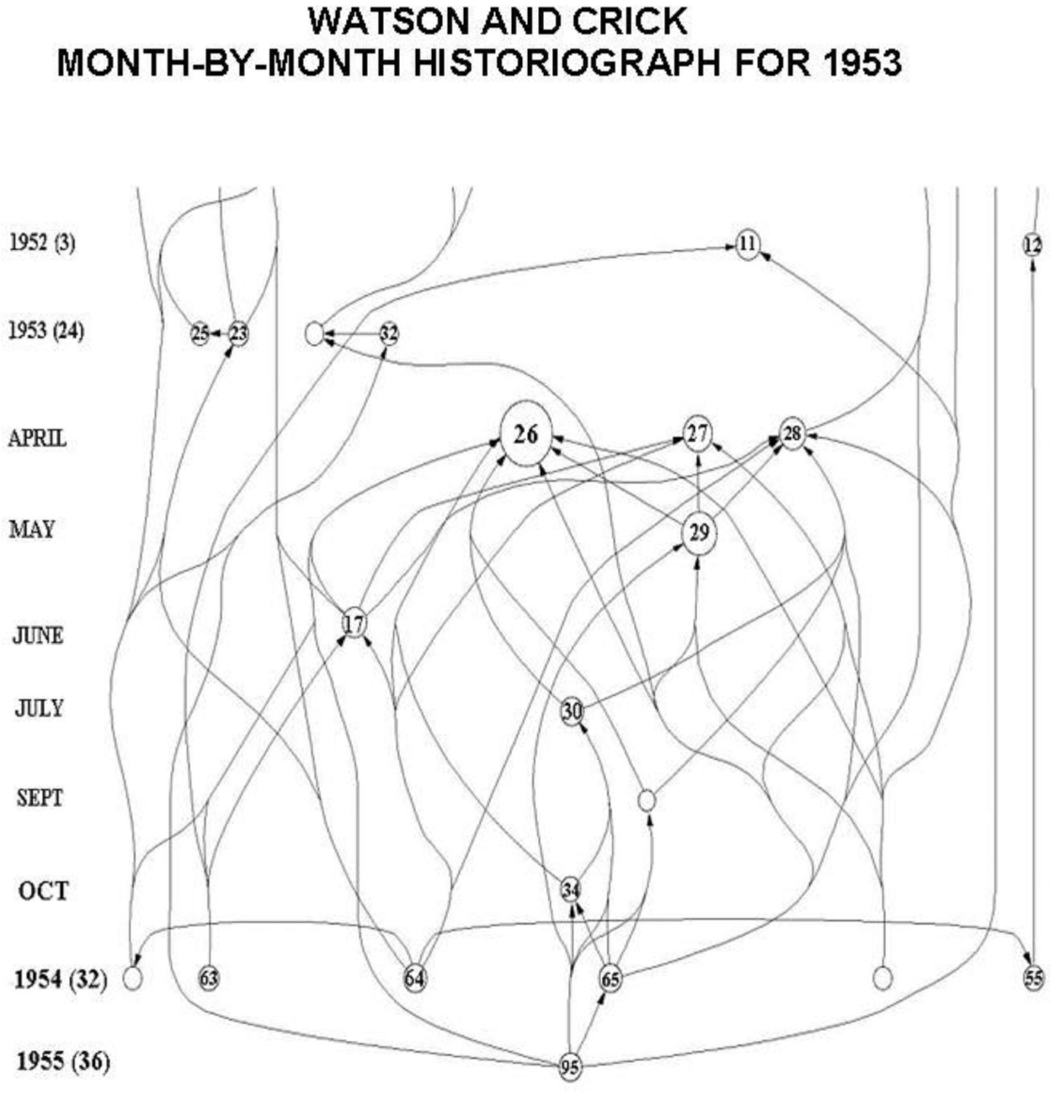
Records: 978, Authors: 1137, Journals: 161, Cited References: 35021, Words: 2044, Tags: 1
Yearly output | Document Type | Language | Institution | Institution with Subdivision | Country
View: Overview Sorted by publication date Page **1** of 10: [1 2 3 4 5 6 7 8 9 10]

#	LCR	NCR	Date / Author / Journal	LCS	GCS
			1938		
1	0	14	1 **SEVAG MG, LACKMAN DB, SMOLENS J** *The isolation of the components of streptococcal nucleoproteins in serologically active form* JOURNAL OF BIOLOGICAL CHEMISTRY. 1938 JUL; 124 (2): 425-436	37	648
			1944		
2	1	37	2 **AVERY OT, MACLEOD CM, MCCARTY M** *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated From Pneumococcus Type III* JOURNAL OF EXPERIMENTAL MEDICINE. 1944 FEB; 79 (2): 137-158	43	1450
			1945		
3	0	11	3 **SCHMIDT G, THANNHAUSER SJ** *A Method for the Determination of Desoxyribonucleic Acid, Ribonucleic Acid, and Phosphoproteins in Animal Tissues* JOURNAL OF BIOLOGICAL CHEMISTRY. 1945; 161 (1): 83-89	34	3897
4	1	34	4 **SCHNEIDER WC** *Phosphorus Compounds in Animal Tissues .1. Extraction and Estimation of Desoxypentose Nucleic Acid and of Pentose Nucleic Acid* JOURNAL OF BIOLOGICAL CHEMISTRY. 1945; 161 (1): 293-303	30	3495
			1946		
5	2	33	5 **MIRSKY AE, POLLISTER AW** *Chromosin, A Desoxyribose Nucleoprotein Complex Of The Cell Nucleus* JOURNAL OF GENERAL PHYSIOLOGY. 1946; 30 (2): 117-&	35	611
			1947		
6	0	40	6 **GULLAND JM, JORDAN DO, TAYLOR HFW** *Deoxypentose Nucleic Acids .2. Electrometric Titration of the Acidic And the Basic Groups of the Deoxypentose Nucleic Acid of Calf Thymus* JOURNAL OF THE CHEMICAL SOCIETY. 1947; (SEP): 1131-1141	31	117

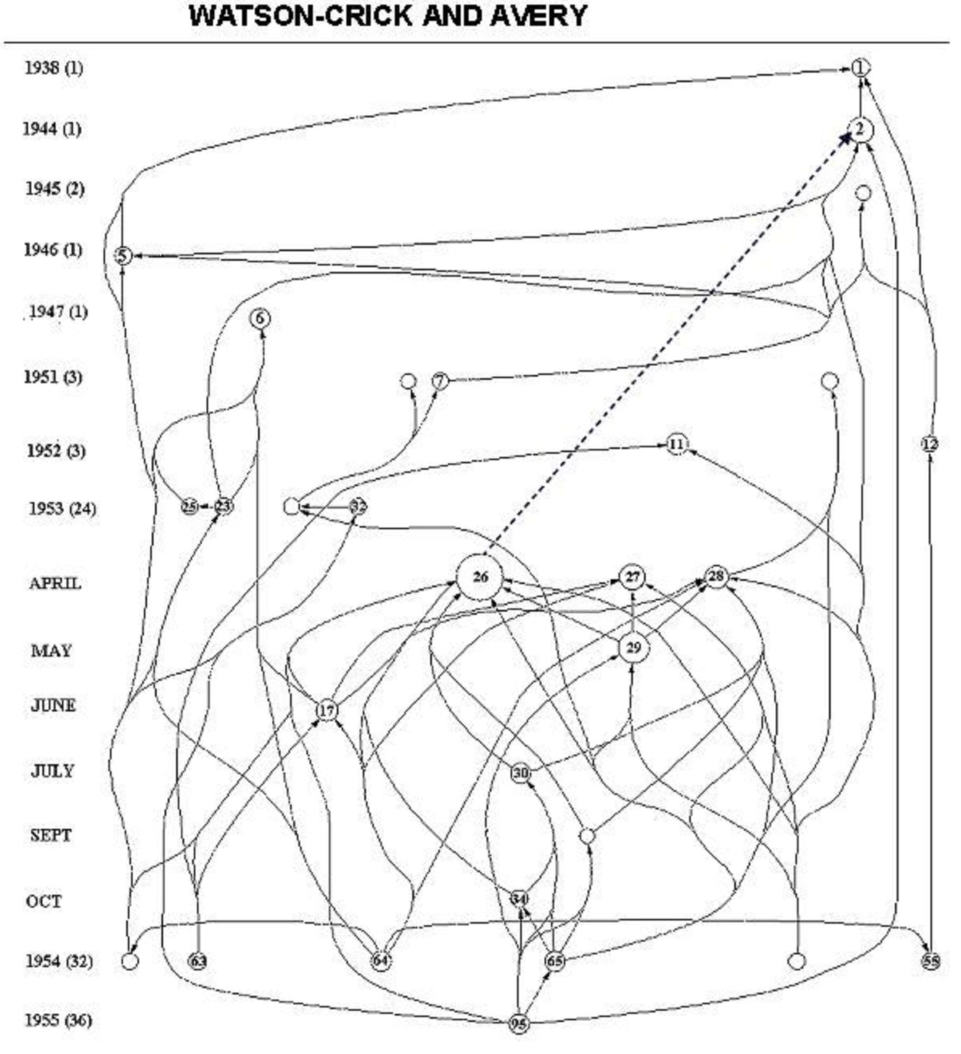
SLIDE 6: WATSON-CRICK – YEAR-BY-YEAR HISTORIOGRAPH



SLIDE 7: WATSON-CRICK MONTH-BY-MONTH HISTORIOGRAPH 1953



SLIDE 8: WATSON-CRICK AND AVERY



This is illustrated in the next series of slides. We used *HistCite* to track the implicit connection between the 1953 Watson-Crick paper on the double helix and the 1944 work of Avery et al on the pneumococcal DNA. As those familiar with this story know, Watson and Crick did not cite the 1944 Avery paper in their 1953 paper. However, a few years ago Jim Watson finally stated his regret that they had not done so.[4] Further, John Maddox, former editor of *Nature* has stated that he would not have allowed the 1953 paper to be published without including the key reference to the Avery work. The Watson-Crick paper was rushed into print without the usual reference e
checks.

A key question often arises as to the ability of citation indexing to retrieve all the relevant work on a topic. So I wanted to demonstrate that the work of significance of the Avery work was known to contemporary workers. To demonstrate the implicit link, however, we did an *SCI* search on *WOS* and retrieved the papers published during the five-year period 1953-1958 which had cited Watson and Crick. Then we added to that collection the group of papers that in turn cited these 200 works. We produced a series of *HistCite* files which are shown in slides 4 to 8.

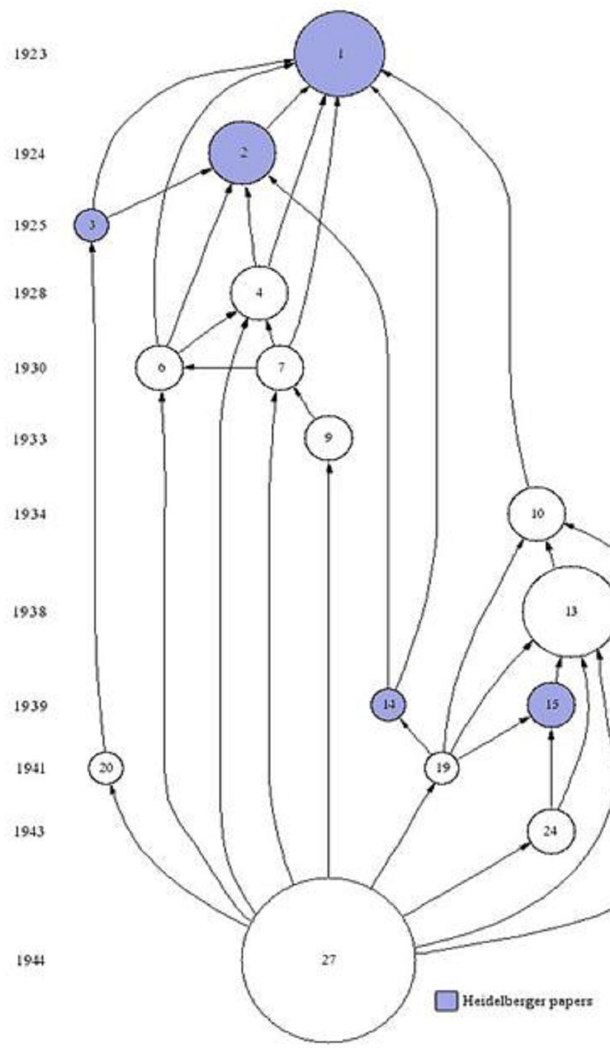
For those of you who would like to see further examples, you can go to the website we have set up at : www.histcite.com and there you will see dozens of files on different authors and subjects.

[If we have access to the WWW I will show the Histcomp URL
http://garfield.library.upenn.edu/histcomp/ .]

SLIDE 9: HEIDELBERGER AVERY CONNECTION

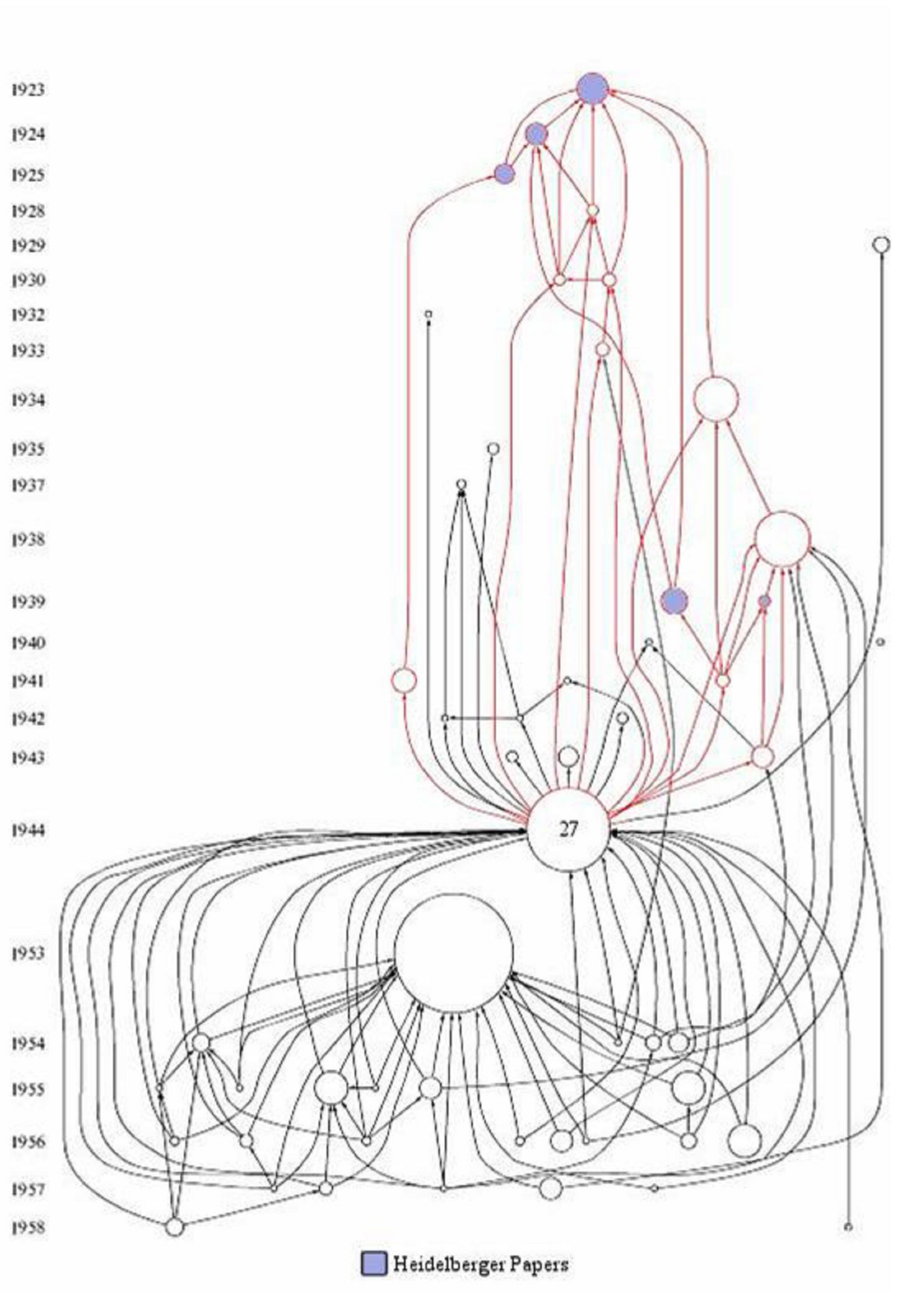
HEIDELBERGER AND AVERY-MACLEOD-MCCARTHY CONNECTION

This collection was created by adding the (19 of 37) papers cited by Avery 1944. Within this group were several citations to 5 papers by Heidelberger. These were added. Finally, the papers from 1953 to 1958 that cited both Avery 1944 and Watson-Crick 1953 were added.



SLIDE 10: HEIDELBERGER, AVERY, WATSON-CRICK

HEIDELBERGER, AVERY WATSON-CRICK



Now as an extension of this exercise, we wanted to explore the historical connection between the work of Michael Heidelberger, mentioned earlier, and his co-author Oswald Avery. One might think there would be a simple and straightforward connection but it turns out otherwise. In the pre-war days before the advent of molecular biology, citation practices were not nearly as standardized as they are today. The classical example of "minimal" citation practice is found in the work of Albert Einstein. Four historiographs of his work can be found on our *Histcite* website. In the 19th and early 20th centuries, the implicit citation was quite common. Therefore, one does not always find an explicit citation of an earlier relevant work. ISI has had to edit thousands of such implicit citations. But by collecting all the relevant citing papers on a subject in a WOS search, the collective memory of the citing authors produces a visual description of the topical history.

Slides 9 and 10, show the links between the work of Avery and Heidelberger and the links between Heidelberger and Watson-Crick.

Journal Impact Factor

Having demonstrated how we can use the WOS search engine to track the historical development of scientific topics, let me turn now to the subject of the ubiquitous journal impact factor. The title of my talk at the last International Congress of Peer Review in Chicago, illustrates 'the agony and the ecstasy of the journal impact factor," and describes my ambiguous feelings on this subject.

We officially launched the Annual *SCI Journal Citation Reports* in 1975. But we were already had producing these data for over a decade. *JCR* evolved from the Journal Citation Index. I illustrated earlier the so-called Author Citation Index. But what is the Journal Citation Index? The *JCI* is the result of resorting the Author Citation Index. Instead of alphabetizing the file by author name, you simply sort the file by the names of the journals in which the papers were published. The *Journal Citation Reports* provides a statistical summation of the Journal Citation Index. .

When we first performed this exercise 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 most. But we needed a simple way to compare large journals like *Nature, Science, NEJM,* and *JAMA* with small journals like the Annual Reviews. In the early days of *Current Contents,* we had emphasized fields of molecular biology and biochemistry. We observed that 25% of all citations in the current year's literature were to papers that were only two to three years old. So we 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. When we did this we obtained results illustrated in the following series of slides.

SLIDE 11: TOP JOURNALS SORTED BY NUMBER OF ARTICLES PUBLISHED IN 2004

In Slide 11, we see the top 20 life science journals sorted by the number of articles published in 2004. *Journal of Biological Chemistry* published 6,500 articles that year.

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

SLIDE 12: MOST-CITED LIFE SCIENCE JOURNALS 2004
In contrast, slide 12 shows the list of journals most-cited in 2004. The *JBC* was cited over 400,000 times that year – this includes citations to any articles in its entire history.

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.[5] We needed a simple method for comparing journals regardless of size or citation frequency. So we created the journal "impact factor."

MOST-CITED JOURNALS, 2004

Abbreviated Journal Title	Total Cites	Impact Factor	Articles
J BIOL CHEM	405017	6.355	6585
NATURE	363374	32.182	878
P NATL ACAD SCI USA	345309	10.452	3084
SCIENCE	332803	31.853	845
J AM CHEM SOC	231890	6.903	3167
PHYS REV LETT	229765	7.218	3575
PHYS REV B	185905	3.075	4964
NEW ENGL J MED	159498	38.570	316
ASTROPHYS J	144264	6.237	2478
J CHEM PHYS	138693	3.105	2772
CELL	136472	28.389	288
LANCET	126002	21.713	415
CIRCULATION	115133	12.563	1129
APPL PHYS LETT	112516	4.308	3731
J IMMUNOL	108602	6.486	1793
J GEOPHYS RES	105601	2.839	2085
CANCER RES	105196	7.690	1253
BLOOD	97885	9.782	1206
BIOCHEMISTRY-US	96809	4.008	1687
J NEUROSCI	93263	7.907	1233

SLIDE 13: LIFE SCIENCE JOURNALS SORTED BY IMPACT FACTOR

Slide 13 shows the life science journals ranked by impact factor. Note the appearance of small review journals.

TOP JOURNALS SORTED BY IMPACT FACTOR, 2004

Abbreviated Journal Title	Total Cites	Impact Factor	Articles
ANNU REV IMMUNOL	14357	52.431	30
CA-CANCER J CLIN	3725	44.515	
NEW ENGL J MED	159498	38.570	316
NAT REV CANCER	6618	36.557	79
PHYSIOL REV	14671	33.918	35
NAT REV MOL CELL BIO	9446	33.170	84
NAT REV IMMUNOL	5957	32.695	80
NATURE	363374	32.182	878
SCIENCE	332803	31.853	845
ANNU REV BIOCHEM	16487	31.538	33
NAT MED	38657	31.223	168
CELL	136472	28.389	288
NAT IMMUNOL	14063	27.586	130
JAMA-J AM MED ASSOC	88864	24.831	351
NAT GENET	49529	24.695	191
ANNU REV NEUROSCI	8093	23.143	26
PHARMACOL REV	7800	22.837	19
NAT BIOTECHNOL	18169	22.355	138
LANCET	126002	21.713	415

.

The term "impact factor" has gradually evolved, especially in Europe, to describe both journal and author impact. This ambiguity often causes problems. It is one thing to use impact factors to compare journals and quite another to use them to compare authors. Journal impact factors generally involve relatively large populations of articles and citations. Individual authors, on average, produce much smaller numbers of articles although some are phenomenally productive. The transplant surgeon Tom Starzl has co-authored over 2,000 articles.[6] Over ten years ago, I attended a celebration of Carl Djerassi's 1000th paper.[7]

While my 1955 paper is considered primordial for citation indexing history, it is my 1972 paper in *Science* on "Citation Analysis as a tool in journal evaluation," that has received most attention from journal editors.[8] That paper was published before the *Journal Citation Reports* existed. We used a quarterly issue of the 1969 SCI to identify the most significant journals of science. I bring this up for an important reason. While our analysis was based on a large sample of literature, the annual *JCR* is not based on a sample. The *JCR* today includes every citation that appears in the 5,000 plus journals that it covers. Therefore, discussions of sampling errors in relation to *JCR* are not particularly meaningful. Furthermore, I myself deplore the quotation of impact factors to three decimal places. ISI uses three decimal places to reduce the number of journals with the identical impact rank. It matters very little whether the impact factor of *JAMA* is quoted as 21.5 rather than 21.455.

A journal's impact factor is based on two elements: the numerator, which is the number of cites in the current year to any items published in the journal in the previous 2 years; and the denominator, the number of substantive articles (source items) 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. A less current impact factor could take into account longer periods of citations and/or sources, but then the measure would be less current. The *JCR* help page provides instruction for computing five-year impact factors.

Scientometrics and Journalology

Citation analysis has blossomed over the past three decades into the field of scientometrics which now has its own International Society of Scientometrics and Informetrics (ISSI).[9] The journal *Scientometrics* was started in 1978. Over 15 years ago, Steve Lock aptly named the application of scientometrics to journals evaluation "journalology."[10]

All citation studies should be normalized to take into account variables such as the discipline, citation density, and half-life.[11]. 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. 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 two-year based impact factors may or may not give as complete a picture as would a five- or ten-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. I reported on this in *The Scientist*.[12,13] seven years ago. 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. Similarly, Hansen and Henrikson[14] reported "good agreement between the journal impact factor and the overall [cumulative] citation frequency of papers on clinical physiology and nuclear medicine."

There are always exceptions to these generalities. Impact critics will usually find them. They also cite all sorts of anecdotal citation behavior which do not represent average behavior. The same can be said about alleged citation errors, most of which are really variants of one kind or another or do not affect impact since only variants in cited journal abbreviations matter in calculating impact. These are all unified prior to issuing the *JCR* each year. And a huge number of author errors or variants are corrected by the ISI system but unseen to the user.

The impact factors reported by *JCR* tacitly imply that all editorial items in *Science*, *Nature*, *JAMA*, *NEJM*, etc. can be neatly categorized. Such journals publish large numbers of items that are not substantive research or review articles. Correspondence, letters, news stories, obituaries, editorials, interviews, and tributes are not included in *JCR*'s calculation of source items (the denominator). But we all know that they may be cited, especially in the current year, but that is also why they don't have a significant effect on the impact calculations. Nevertheless, since the *JCR* numerator includes citations to these more ephemeral items, some distortion will result. But only a small group of journals are affected, if at all. Those that are affected change by 5 or 10%.[8]

The assignment of article 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 versus laboratory studies or, for that matter, practice-based versus research-based articles. All these potential variables provide grist for the critical mill of citation aficionados

Size vs. Citation Density

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 most fields are not well cited, whereas some articles in small fields may have unusual impact, especially where they have cross-disciplinary impact. It is well known that there is a skewed distribution of citations in most fields. The well-known 80/20 rule applies in that 20% of articles may account for 80% of the citations.

To reiterate -- the key determinants in impact are not the number of authors or articles in the field but, rather, the citation density and the age of the literature cited. The average number of citations per article and the immediacy of citations are the significant elements.[15] The size of a 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. Nevertheless, review papers on average are cited about twice the average. Publishing mediocre review papers will not necessarily boost your journal's impact.

SLIDE 14: SUPER CITED PAPERS IN THE LIFE SCIENCES

For your amusement, consider this short list of super-cited papers in the life sciences. Incidentally, since they are all over a decade or more old, they don't affect the calculation of their journal's impact factor. The Lowry paper was recently discussed in *Journal of Biological Chemistry*[16] but the authors failed to mention Lowry's own commentary on this most-cited paper in the history of science.[17] Lowry himself noted that it was not his most important paper.

I have not included here super cited books such as *Molecular Cloning: a Laboratory Manual* by Maniatis and Sambrook which appeared[18] in numerous editions beginning with 1982. They have been cited in at least 150,000 papers. This is my way of reminding those who are book authors, that *SCI*, *SSCI* and *A&HCI* do include citations to books as well as individual chapters of books.

Slide 14:

MOST CITED PAPERS

Through July 2005

Authors	Title	Source	Yr	V	Pg	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 f 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	PNAS USA	1977	74	5463	63,909
CHOMCZYNSKI, P; 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 Polyacrylamide Gels To Nitrocellulose Sheets - Procedure and Some Applications	PNAS 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

SLIDE 15: CITATION FREQUENCY DISTRIBUTIONS

Citation Frequency Distribution 1900-August, 2005

(articles cited at least once)

Number of Citations	Approx # of Items Receive Citations	% of WOS
>10,000	61	0.00%
5,000-9,000	120	0.00%
4,000-4,999	116	0.00%
3,000-3,999	215	0.00%
2,000-2,999	664	0.00%
1,000-1,999	3,887	0.02%
900-999	1,232	0.00%
800-899	1,762	0.01%
700-799	2,614	0.01%
600-699	4,077	0.02%
500-599	6,637	0.03%
400-499	12,557	0.06%
300-399	27,059	0.14%
200-299	74,025	0.37%
100-199	343,269	1.73%
50-99	953,064	4.83%
25-49	2,006,529	10.1%
15-24	2,226,603	11.2%
10-14	2,106,995	10.6%
5-9	3,891,542	19.5%
2-4	4,931,952	24.7%
1	3,343,789	16.7%
Items Cited	**19,938,769**	**100.1%**

For a more realistic view of citation frequencies, slide 15 shows that from 1900-2005, about one half of one percent of cited papers were cited over 200 times. Out of about 38 million source items about half were not cited at all. Keep in mind that "items" includes not only substantive articles but also ephemera mentioned earlier. Therefore, these data provide a distorted picture for high impact journals where the number of uncited publications is much smaller.

The skewness of citations is well known and repeated as a mantra by critics of the impact factor. On the one hand, some editors would like to see impacts calculated solely on the basis of their most-cited papers so that their otherwise low impact factors can be ignored. However, since most journals experience this skewness, that 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* two-year window will not be counted.[19]

Alternatively, the appearance of articles on the same subject in the same issue of a journal may have an upward effect. Opthof[20] showed how journal impact performance can vary from issue to issue.

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.[21]

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 an identical profile. But in fact there rarely are 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. Some might argue that JCR categories are larger than necessary. Recent work by Alexander Pudovkin and myself[22] is an attempt to group journals more objectively. We rely on the two-way citational relationships between journals to reduce the subjective influence of journal titles. Three decades ago, I demonstrated that journal titles can be deceiving. Citation analysis proved the *Journal of Experimental Medicine* was a leading immunology journal.[23] It still is one of the five top immunology journals based on its impact factor.

SLIDE 16: GENERAL INTERNAL MEDICINE CATEGORY SORTED BY IMPACT 2004.

In Slide 16, you see the list of journals in the JCR category "Medicine, General and Internal." There are no surprises here. Few would quarrel with the assignment of these journals to this category, but this tells us little about their actual subject content.

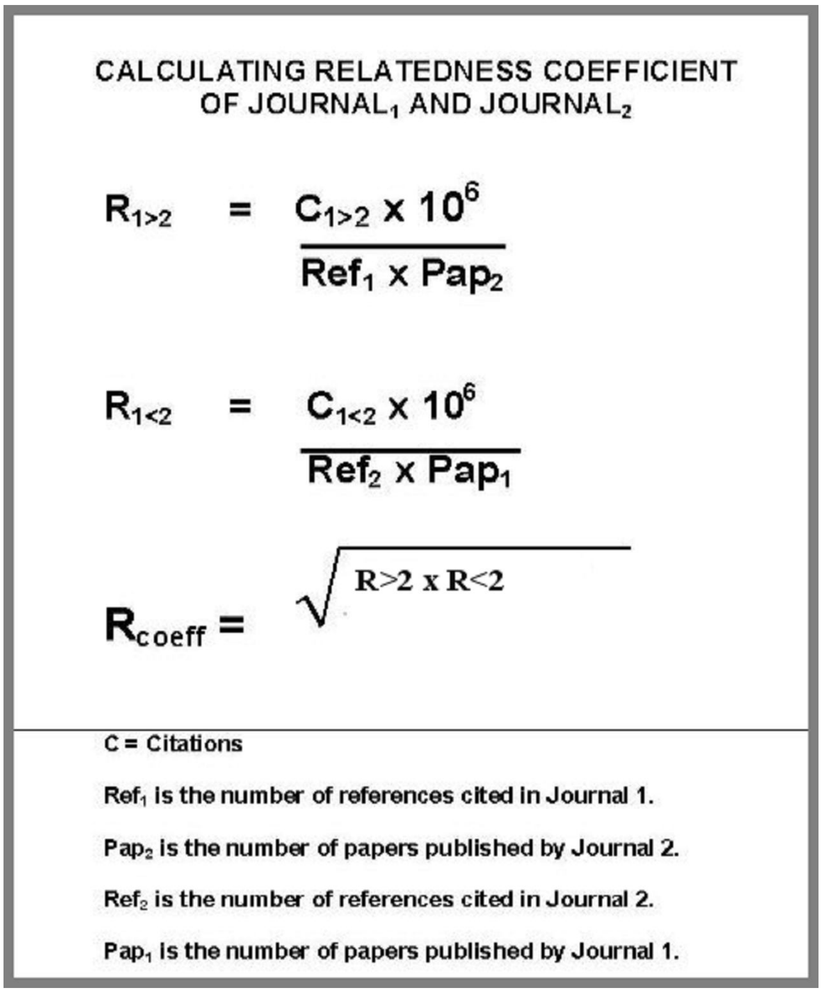
MEDICINE, GENERAL & INTERNAL

Journals sorted by Impact factor

Abbreviated Journal Title	Total Cites	Impact Factor	Articles
NEW ENGL J MED	159498	38.570	316
JAMA-J AM MED ASSOC	88864	24.831	351
LANCET	126002	21.713	415
ANN INTERN MED	36932	13.114	189
ANNU REV MED	3188	11.200	29
ARCH INTERN MED	26525	7.508	282
BRIT MED J	56807	7.038	623
CAN MED ASSOC J	6736	5.941	100
AM J MED	21000	4.179	285
MAYO CLIN PROC	6816	3.746	161
MEDICINE	4255	3.727	30
ANN MED	2626	3.617	79
J INTERN MED	4793	3.590	135
AM J PREV MED	3972	3.188	143
CURR MED RES OPIN	1148	2.928	212
J GEN INTERN MED	4686	2.821	163
QJM-INT J MED	4073	2.580	73
EUR J CLIN INVEST	4332	2.530	110
PREV MED	5372	2.327	287
J PAIN SYMPTOM MANAG	2941	2.187	117

SLIDE 17: CALCULATING RELATEDNESS COEFFICIENTS

JCR recently added a new feature which provides you the ability to more precisely establish journal categories based on citation relatedness. Slide 17 provides the general formula for calculating citation relatedness between two journals and the relatedness coefficient expressing the average of the maximum and minimum.



SLIDE 18: JOURNALS: *JAMA* - RELATED JOURNALS SORTED BY CITATION RELATEDNESS COEFFICIENT

Slide 18 is a list of the 20 journals most related to *JAMA* by the citation relatedness coefficient, which reflects how often *JAMA* cites and is cited by each of the journals listed. The relatedness coefficient takes into account the sizes of the journals involved (papers published) as well as the number of times each journal cites the other.

The top four journals related to JAMA remain the same as in the sort by impact, but many journals have moved up in rank such as *Journal of Family Planning* and *Journal of the American Geriatrics Society.* The checks on the left indicate the journal was not assigned to the General Medicine category.

Slide 18:

JOURNALS MOST RELATED BY CITATION RELATEDNESS TO JAMA

√ = Not in Medicine, General & Internal Category

	Journal	Rcoefficient
	JAMA-J AM MED ASSOC	274.97
	ANN INTERN MED	127.26
	NEW ENGL J MED	123.09
	ARCH INTERN MED	89.85
	J GEN INTERN MED	70.26
√	CONTROL CLIN TRIALS	69.23
√	ADV RENAL REPLACE TH	66.41
√	MED CARE	66.02
	J FAM PRACTICE	64.81
√	HEALTH AFFAIR	64.64
√	J AM GERIATR SOC	53.06
√	CURR CONTR TRIALS C	52.84
√	ACAD MED	52.75
√	INQUIRY-J HEALTH CAR	52.00
	CAN MED ASSOC J	46.98
	AM J MED	46.70
	AM J PREV MED	45.37
√	ARCH PEDIAT ADOL MED	40.25
√	CLIMACTERIC	39.73
√	J AM MED INFORM ASSN	38.28
√	MENOPAUSE	34.55
√	PHARMACOEPIDEM DR S	34.20
√	AM J MED QUAL	33.89
√	ENDOCRIN METAB CLIN	33.89
√	MATURITAS	31.90
√	BLOOD PRESS MONIT	30.20
	FAM MED	30.16

SLIDE 19: *NEJM* RELATED JOURNALS SORTED BY CITATION RELATEDNESS

Performing the identical exercise for the *NEJM*, we see differences that are quite striking. The top four journals are there: *NEJM, Annals of Internal Medicine, JAMA, and Archives of Internal Medicine*, but the next two are cardiology journals, as are 9 of the next 12 journals shown.

While this observation does not affect the categorization of *NEJM* as a general medicine journal, the next slide will demonstrate further that it is relevant to list it in the cardiology category, as well.

JOURNALS MOST RELATED BY CITATION RELATEDNESS TO NEW ENGLAND JOURNAL OF MEDICINE

√ = Not in Medicine, General & Internal Category

	Journal	Rcoefficient
	NEW ENGL J MED	345.24
	JAMA-J AM MED ASSOC	123.09
	ANN INTERN MED	124.85
	ARCH INTERN MED	64.49
	AM J MED	61.13
√	CIRCULATION	57.36
√	J AM COLL CARDIOL	58.15
	MAYO CLIN PROC	47.96
√	CHEST	37.64
√	PROG CARDIOVASC DIS	45.66
	CAN MED ASSOC J	40.31
√	CRIT CARE MED	35.11
√	CURR PROB CARDIOLOGY	36.19
√	J CARD FAIL	34.62
√	EUR HEART J	36.77
√	AM HEART J	37.99
√	AM J CARDIOL	33.90
	AM J MED SCI	27.40
√	MED LETT DRUGS THER	32.97
√	RESUSCITATION	24.79
√	BONE MARROW TRANSPL	22.66
√	GASTROENTEROL CLIN N	24.72
√	CURR OPIN CARDIOL	21.46
	MED CLIN N AM	22.25
√	HEART	22.54

SLIDE 20: JCR CARDIAC & CARDIOVASCULAR SYSTEMS BY IMPACT FACTOR

Here is the listing of the cardiac journals category in the 2004 *JCR*. The ranking by impact factor probably conforms to the general idea of the most prestigious journals in the field.

JCR CATEGORY: CARDIAC & CARDIOVASCULAR SYSTEMS

↓

Rmax Rank	*JCR* Rank	Abbreviated Journal Title	Total Cites	Impact Factor	Articles
3	1	Circulation	115133	12.563	1129
12	2	Circulation Research	35038	9.972	340
2	3	Journal of the American College of Cardiology	40841	9.133	591
4	4	European Heart Journal	10890	6.247	250
16	5	Trends In Cardiovascular Medicine	1497	4.716	53
13	6	Cardiovascular Research	12390	4.575	269
14	7	Journal of Molecular and Cellular Cardiology	7618	4.198	163
7	8	American Heart Journal	14243	3.681	356
17	9	American J of Physiology-Heart and Circulatory Physiology	23887	3.539	652
6	10	Heart	6023	3.271	314
15	11	Journal of Thoracic and Cardiovascular Surgery	15028	3.263	327
5	12	American Journal of Cardiology	29703	3.140	824
19	13	Chest	27826	3.118	654
11	14	Basic Research in Cardiology	1702	3.009	45
	15	European J of Cardiovascular Prevention & Rehabilitation	46	3.000	73
1	16	Journal of Cardiovascular Electrophysiology	4258	2.967	205
8	17	Journal of Cardiac Failure	1213	2.879	79
18	18	Journal of Heart and Lung Transplantation	4023	2.813	220
10	19	European Journal of Heart Failure	1164	2.796	118
9	20	Progress in Cardiovascular Diseases	1327	2.676	31

Source: 2004 *Journal Citation Reports*

SLIDE 21: JOURNALS MOST RELATED TO *CIRCULATION* BY CITATION RELATEDNESS

However, using the *JCR* relatedness ranking method, some journals would be assigned to different *JCR* categories. Using *Circulation*, the highest impact journal in this area, to represent cardiology, we find that *NEJM* 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. However, this analysis also tells us something about the *JCR* placement of the journal *Coronary Artery Disease*. *JCR* assigns it to the category "Peripheral Vascular Disease" but it is in fact the 10th journal in this list.

JOURNALS MOST RELATED BY CITATION RELATEDNESS TO CIRCULATION

Journal	Rmax	Rcirc>j	Rj>circ	Rco-efficient	Rank by Rco-efficient
CIRCULATION	160.16	160.16	160.16	160.16	1
J AM COLL CARDIOL	165.01	85.54	165.01	118.81	2
J CARDIOVASC ELECTR	220.69	27.68	220.69	78.16	3
AM J CARDIOL	156.28	32.9	156.28	71.71	4
EUR HEART J	159.56	31.57	159.56	70.97	5
AM HEART J	139.48	30.65	139.48	65.38	6
→ NEW ENGL J MED	170.03	170.03	19.35	57.36	7
PROG CARDIOVASC DIS	124.73	24.96	124.73	55.80	8
J CARD FAIL	128.67	20.57	128.67	51.45	9
→ CORONARY ARTERY DIS	170.9	14.68	170.9	50.09	10
CURR PROB CARDIOLOGY	180.95	12.9	180.95	48.31	11
BASIC RES CARDIOL	105.09	21.21	105.09	47.21	12
HEART	145.6	14.54	145.6	46.01	13
PACE	159.27	10.76	159.27	41.40	14
J AM SOC ECHOCARDIOG	144.16	11.7	144.16	41.07	15
CARDIOLOGY	113.14	14.54	113.14	40.56	16
CURR OPIN CARDIOL	142.56	11.14	142.56	39.85	17
CARDIOVASC DRUG THER	112.77	10.89	112.77	35.04	18
CATHETER CARDIO INTE	164.94	6.94	164.94	33.83	19
J CARDIOV MAGN RESON	153.61	7.44	153.61	33.81	20
J INTERV CARD ELECTR	173.06	6.45	173.06	33.41	21
J NUCL CARDIOL	169.8	6.45	169.8	33.09	22
EUR J HEART FAIL	123.25	8.31	123.25	32.00	23
CLIN CARDIOL	115.94	7.09	115.94	28.67	24
INT J CARDIOL	125.83	5.56	125.83	26.45	25
J ELECTROCARDIOL	119.84	5.45	119.84	25.56	26

SLIDE 22: JPI DATA ON 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
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	7,311	383
24	2004	3.34	1,174	351

31,257 / 370 — Citations received 1999-2004 / Articles published in JAMA in 1999 = 84.5

Many of the discrepancies with journal impact factors are eliminated altogether in another ISI database called the *Journal Performance Indicators (JPI)*.[24] This annual compilation now covers the period 1981 to 2004. Unlike *JCR*, the database links each source item to its own unique citations. Therefore, the impact calculations are 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. 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.

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

In 1999, *JAMA* published 1905 items of which 680 were letters, and 253 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 recent report by Gonzalez and Companario at the University of Alcala demonstrates that the effect, if any, is quite minor. [25]

SLIDE 23: MYCOLOGY JOURNALS EFFECT OF TIME ON IMPACT RANKINGS FOR ONE, FIVE, AND 24 YEAR PERIOD.

EFFECT OF TIME ON IMPACT RANKINGS OF MYCOLOGY JOURNALS
Ranks for one, five, and 24 year period

Rank	2004 Impact Factor	Impact 2000-2004	Impact 1981-2004
1	Fungal Genetics/Biol. (3.05)	Fungal Genetics/Biol. (5.81)	Yeast (17.53)
2	Yeast (1.94)	Yeast (5.13)	Experimental Mycology (14.36)
3	Mycorrhiza (1.74)	Medical Mycology (4.53)	J. Med. Veter. Mycol. (12.76)
4	Medical Mycology (1.45)	Mycorrhiza (3.37)	Fungal Genetics/Biol. (9.70)
5	Mycologia (1.43)	Mycologia (3.20)	Mycologia (8.46)
6	Fungal Diversity (1.89)	Mycological Research (3.17)	Mycological Research (7.72)
7	Mycological Research (1.13)	Lichenologist (1.95)	Mycorrhiza (7.16)
8	Lichenologist (0.73)	Fungal Diversity (1.87)	Mycopathologia (6.19)
9	Mycopathologia (0.87)	Mycoses (1.63)	Medical Mycology (6.16)
10	Mycoses (0.69)	Mycopathologia (1.53)	Lichenologist (5.90)

From: http://in-cites.com/research/2005/april_25_2005-1.html

EFFECT OF TIME ON IMPACT RANKINGS OF MYCOLOGY JOURNALS
Ranks for one, five, and 24 year period

Rank	2004 Impact Factor	Impact 2000-2004	Impact 1981-2004
1	Fungal Genetics/Biol. (3.05)	Fungal Genetics/Biol. (5.81)	Yeast (17.53)
2	Yeast (1.94)	Yeast (5.13)	Experimental Mycology (14.36)
3	Mycorrhiza (1.74)	**Medical Mycology** (4.53)	J. Med. Veter. Mycol. (12.76)
4	**Medical Mycology** (1.45)	Mycorrhiza (3.37)	Fungal Genetics/Biol. (9.70)
5	Mycologia (1.43)	Mycologia (3.20)	Mycologia (8.46)
6	Fungal Diversity (1.89)	Mycological Research (3.17)	Mycological Research (7.72)
7	Mycological Research (1.13)	Lichenologist (1.95)	Mycorrhiza (7.16)
8	Lichenologist (0.73)	Fungal Diversity (1.87)	Mycopathologia (6.19)
9	Mycopathologia (0.87)	Mycoses (1.63)	**Medical Mycology** (6.16)
10	Mycoses (0.69)	Mycopathologia (1.53)	Lichenologist (5.90)

From: http://in-cites.com/research/2005/april_25_2005-1.html

To illustrate the chronological changes in rankings for a group of related journals, consider the topic of mycology which was reported recently in *inCites* for April 25, 2005.[26]
(http://in-cites.com/research/2005/april_25_2005-1.html) *inCites* is a free ISI news bulletin.

While the journal *Medical Mycology* ranked 4th in 2004, it moved to 3rd place when five years of data were used but 9th when 23 years of data were used. This example seems to contradict the generalization I made when discussing physiology journals.

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 include the most prestigious. The perception of prestige is a murky subject. Some would equate prestige with high impact. However, some librarians argue that the numerator in the impact-factor calculation is itself even more relevant. Bensman[27] argued that this 2-year total citation count is a better guide to journal significance and cost-effectiveness than is the impact factor. This brings us full circle to the first slide I showed you on the most-cited journals.

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 issue. Granting and other policy agencies often wish to bypass the work involved in obtaining actual citation counts for individual articles and authors. And allegedly recently published articles may not have had enough time to be cited, so 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. Thus, the impact factor is used to estimate the expected influence 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 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. But some studies would 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 enabled *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 magazine* 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."[28] .

SLIDE 24: EXAMPLES OF HOT PAPERS

Two recent examples of Hot Papers published in *JAMA* and *NEJM* include papers on coronavirus at http://in-cites.com/hotpapers/2005/may05-cli.html

"A NOVEL CORONAVIRUS ASSOCIATED WITH SEVERE ACUTE RESPIRATORY SYNDROME"

"IDENTIFICATION OF A NOVEL CORONAVIRUS IN PATIENTS WITH SEVERE ACUTE SYNDROME"

HOT PAPER: Citations: 515

Title: A NOVEL CORONAVIRUS ASSOCIATED WITH SEVERE ACUTE RESPIRATORY SYNDROME

Authors: Ksiazek TG; Erdman D; Goldsmith CS; Zaki SR; Peret T; Emery S; Tong SX; Urbani C; Comer JA; Lim W; Rollin PE; Dowell SF; Ling AE; Humphrey CD; Shieh WJ; Guarner J; Paddock CD; Rota P; Fields B; Derisi J; Yang JY; Cox N; Hughes JM; Leduc JW; Bellini WJ; Anderson LJ

Source: N ENGL J MED 348: (20) 1953-1966 MAY 15 2003

Addresses:

Ctr Dis Control & Prevent, Special Pathogens Branch, Natl Ctr Infect Dis, Atlanta, GA 30333 USA
Ctr Dis Control & Prevent, Resp & Enter Virus Brach, Natl Ctr Infect Dis, Atlanta, GA USA
Ctr Dis Control & Prevent, Infect Dis Pathol Act, Natl Ctr Infect Dis, Atlanta, GA USA
Ctr Dis Control & Prevent, Influenza Branch, Natl Ctr Infect Dis, Atlanta, GA USA
Ctr Dis Control & Prevent, Div Bacterial & Mycot Dis, Natl Ctr Infect Dis, Atlanta, GA USA
Ctr Dis Control & Prevent, Off Director, Div Viral & Rickettsial Dis, Natl Ctr Infect Dis, Atlanta, GA USA
WHO, Hanoi, Vietnam.
Queen Mary Hosp, Govt Virus Unit, Hong Kong, Hong Kong, Peoples R China.
Int Emerging Infect Dis Program, Bangkok, Thailand.
Univ Calif San Francisco, San Francisco, CA 94143 USA
Singapore Gen Hosp, Dept Pathol, Singapore, Singapore.
Ctr Dis Control, Dept Hlth, Taipei, Taiwan.

HOT PAPER: Citations: 475

Title: IDENTIFICATION OF A NOVEL CORONAVIRUS IN PATIENTS WITH SEVERE ACUTE RESPIRATORY SYNDROME

Authors: Drosten C; Gunther S; Preiser W; Van Der Werf S; Brodt HR; Becker S; Rabenau H; Panning M; Kolesnikova L; Fouchier Ram; Berger A; Burguiere Am; Cinatl J; Eickmann M; Escriou N; Grywna K; Kramme S; Manuguerra Jc; Muller S; Rickerts V; Sturmer M; Vieth S; Klenk HD; Osterhaus ADME; Schmitz H; Doerr HW

Source: N ENGL J MED 348: (20) 1967-1976 MAY 15 2003

Addresses:

Bernhard Nocht Inst Trop Med, Dept Virol, Natl Reference Ctr Trop Infect Dis, Bernhard Nocht Str 74, D-20359 Hamburg, Germany.
Bernhard Nocht Inst Trop Med, Dept Virol, Natl Reference Ctr Trop Infect Dis, D-20359 Hamburg, Germany.
Univ Frankfurt, Inst Med Virol, D-6000 Frankfurt, Germany.
Univ Frankfurt, Med Clin 3, D-6000 Frankfurt, Germany.
Univ Marburg, Inst Virol, D-3550 Marburg, Germany.
Inst Pasteur, Natl Influenza Ctr No France, Paris, France.
Erasmus Univ, Inst Virol, Rotterdam, Netherlands.

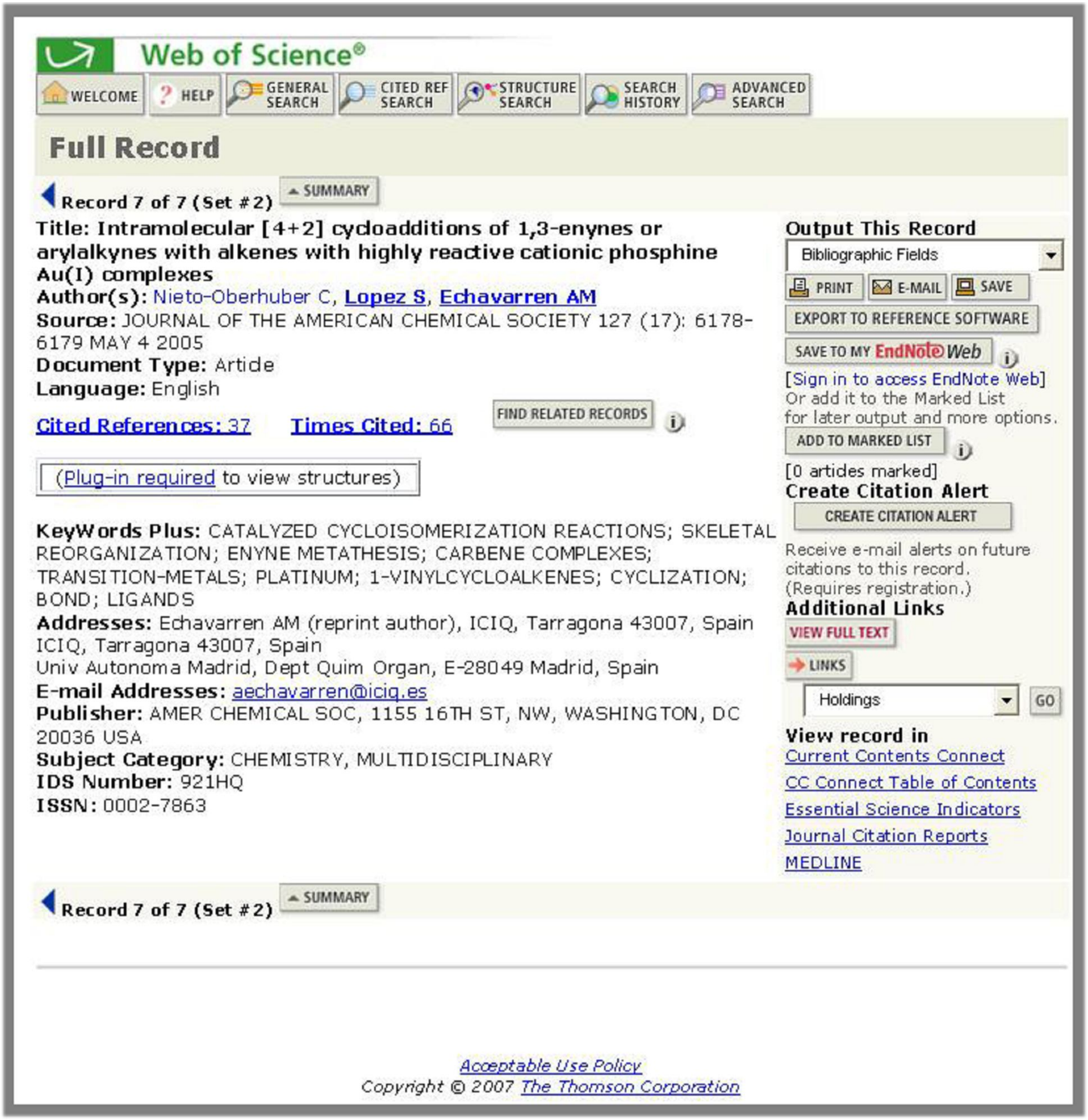
From: http://in-cites.com/hotpapers/2005/may05-cli.html

Slide 25: WHAT'S HOT IN BIOLOGY

WHAT'S HOT IN BIOLOGY...			
Rank	**Paper**	**Citations This Period (Jul-Aug 06)**	**Rank Last Period (May-Jun 06)**
1	D. Altshuler, *et al.* (Int.'l HapMap Consortium), **"A haplotype map of the human genome,"** *Nature*, 437(7063): 1299-1320, 27 October 2005. [63 institutions worldwide] *977UQ	57	1
2	J.D. Fontenot, *et al.*, **"Regulatory T cell lineage specification by the forkhead transcription factor Foxp3,"** *Immunity*, 22(3): 329-41, March 2005. [Howard Hughes Med. Inst., U. Seattle, WA] *912UP	40	†
3	R.L. Levine, *et al.*, **"Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis,"** *Cancer Cell*, 7(4): 387-97, April 2005. [7 U.S. and European institutions] *921CL	38	5
4	C.T. Harbison, *et al.*, **"Transcriptional regulatory code of a eukaryotic genome,"** *Nature*, 431(7004): 99-104, 2 September 2004. [Whitehead Inst., Cambridge, MA; Broad Inst., Cambridge, MA; MIT, Cambridge, MA] *850VC	31	†
5	L.W. Hillier, *et al.* (Int.'l Chicken Genome Seq. Consortium), **"Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution,"** *Nature*, 432(7018): 695-716, 9 December 2004. [50 institutions worldwide] *877UE	29	7
6	K.N. Ferreira, *et al.*, **"Architecture of the photosynthetic oxygen-evolving center,"** *Science*, 303(5665): 1831-8, 19 March 2004. [Imperial Coll., London, U.K.; Japan Sci. Tech. Corp., Nagatsuta] *804EI	28	3
7	T.S. Mikkelsen, *et al.* (The Chimpanzee Seq. and Analysis Consort.), **"Initial sequence of the chimpanzee genome and comparison with the human genome,"** 437(7055): 69-87, 1 September 2005. [23 institutions worldwide] *960AC	27	6
8	A.J. McCoy, *et al.*, *Acta Cryst. D*, **"Likelihood-enhanced fast translation functions,"** 61[4]: 458-64, April 2005. [U. Cambridge, U.K.; Lawrence Berkeley Natl. Lab., Berkeley, CA] *909NW	27	†
9	M. Arrasate, *et al.*, **"Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death,"** *Nature*, 431(7010): 805-10, 14 October 2004. [U. Calif., San Francisco; U. Calif., Los Angeles] *861RE	25	†
10	D.A. Hinds, *et al.*, **"Whole-genome patterns of common DNA variation in three human populations,"** *Science*, 307(5712): 1072-9, 18 February 2005. [Perlegen Sciences Inc., Mountain View, CA; Int'l. Computer Science Inst., Berkeley, CA; U. Calif., San Diego] *900ED	24	†

SOURCE: Thomson Scientific Hot Papers Database
*, †—See legend in the table on page 5

SLIDE 26: HOT PAPERS FROM TARRAGONA, SPAIN



Web of Science®

WELCOME | HELP | GENERAL SEARCH | CITED REF SEARCH | STRUCTURE SEARCH | SEARCH HISTORY | ADVANCED SEARCH

Full Record

Record 7 of 7 (Set #2) SUMMARY

Title: Intramolecular [4+2] cycloadditions of 1,3-enynes or arylalkynes with alkenes with highly reactive cationic phosphine Au(I) complexes
Author(s): Nieto-Oberhuber C, Lopez S, Echavarren AM
Source: JOURNAL OF THE AMERICAN CHEMICAL SOCIETY 127 (17): 6178-6179 MAY 4 2005
Document Type: Article
Language: English

Cited References: 37 Times Cited: 66 FIND RELATED RECORDS

(Plug-in required to view structures)

KeyWords Plus: CATALYZED CYCLOISOMERIZATION REACTIONS; SKELETAL REORGANIZATION; ENYNE METATHESIS; CARBENE COMPLEXES; TRANSITION-METALS; PLATINUM; 1-VINYLCYCLOALKENES; CYCLIZATION; BOND; LIGANDS
Addresses: Echavarren AM (reprint author), ICIQ, Tarragona 43007, Spain
ICIQ, Tarragona 43007, Spain
Univ Autonoma Madrid, Dept Quim Organ, E-28049 Madrid, Spain
E-mail Addresses: aechavarren@iciq.es
Publisher: AMER CHEMICAL SOC, 1155 16TH ST, NW, WASHINGTON, DC 20036 USA
Subject Category: CHEMISTRY, MULTIDISCIPLINARY
IDS Number: 921HQ
ISSN: 0002-7863

Output This Record
Bibliographic Fields
PRINT | E-MAIL | SAVE
EXPORT TO REFERENCE SOFTWARE
SAVE TO MY EndNote Web
[Sign in to access EndNote Web]
Or add it to the Marked List for later output and more options.
ADD TO MARKED LIST
[0 articles marked]
Create Citation Alert
CREATE CITATION ALERT
Receive e-mail alerts on future citations to this record. (Requires registration.)
Additional Links
VIEW FULL TEXT
LINKS
Holdings GO
View record in
Current Contents Connect
CC Connect Table of Contents
Essential Science Indicators
Journal Citation Reports
MEDLINE

Record 7 of 7 (Set #2) SUMMARY

Acceptable Use Policy
Copyright © 2007 The Thomson Corporation

Conclusion

Of the many conflicting opinions about impact factors, Hoeffel[29] 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."

Yes, a better evaluation system would involve actually reading each article for quality but then this entire congress is dedicated to the difficulties of reconciling peer review judgments. When it comes time to evaluating faculty, most people do not have or care to take the time to read the articles any more! Even if they did, their judgment surely would be tempered by observing the comments of those who have cited the work. We call this citation context analysis. Fortunately, new full-text capabilities in the web make this more practical to perform.

I have had to rush through a lot material to save time but hope that I have given you a balanced view of a complex and controversial topic.

We have used journal impact factors to help in the selection process both for *CC* and *SCI* ever since. Since the *Web of Science* now covers over 5,000 journals as the graphs I have shown illustrate, the decision to add journals not yet in the system are often necessarily subjective because one is often making judgments on many low impact journals. While it is not possible to know in advance whether a journal will achieve above average impact, it is possible to estimate based on a variety of known factors.

While the editor or publisher of journals with low impact factors may object to the impact factor, it is usually easy to demonstrate whether a particular journal deserves to be added to the *SCI* or *CC*.

Journal Impact Factor as Surrogate for Citation Rank

However, it is the use of the journal impact factor in evaluating the work of individual scientists which causes the greatest grief. So why is the journal impact factor used for that purpose. For younger scientists many of the papers listed in their C.V.s were published quite recently often during the period used to calculate impact. With certain exceptions, such as "hot papers," these papers will not be cited for a few years or more depending upon the rate at which research on their topic progresses. About a decade ago administrators decided they would estimate the future impact of recently published papers by incorporating the impact factor for the journal in which the paper is published.

This "expected" impact factor is often flawed because within every journal there is a skewed dispersion of citation frequencies . While the average for *Nature* or *JBC* may be high, the skewed distribution must be taken into account. As in most bibliometric studies the 80/20 rule applies. A small percentage of papers accounts for a large percentage of citations.

I have briefly touched on the various aspects of scientific information studies which have been the main focus of my career. I will be glad to answer any questions you may have.
References:
REFERENCES:

1. Garfield, E. "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas." *Science*, 122(3159):108-11 (July 1955). Reprinted in Garfield, E. *Essays of an Information Scientist*, Volume 6, pgs. 468-471. Philadelphia: ISI Press (1984) Back to Document Ref.1

2. Garfield, E. "Citation analysis as a tool in journal evaluation ," *Science*, 178 (4060):471-479 (1972). Reprinted in *Essays of an Information Scientist, Volume* 1 pg. 527. Philadelphia: ISI Press (1977). Back to Document. Ref:2
http://www.garfield.library.upenn.edu/essays/V1p527y1962-73.pdf

3. Garfield, E., I. H. Sher, and R. J. Torpie. "The Use of Citation Data in Writing the History of Science." Published by The Institute for Scientific Information, December 1964. Report of research for Air Force Office of Scientific Research under contract F49(638)-1256. Back to Document. Ref:3
http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf

4. Anonymous. "Genes, Girls, and Honest Jim," *Bio-IT World*, 2(4):28 (2003).
Back to Document. Ref:4

5. Brodman E. "Choosing physiology journals," *Bulletin of the Medical Library Association* 32:479 (1960) Back to Document. Ref:5

6. http://garfield.library.upenn.edu/histcomp/index-starzl.html Back to Document. Ref:6

7. Garfield E. "A tribute to Carl Djerassi: Reflections on a remarkable scientific entrepreneur," *Current Contents* (42): 5-14 (October 18,1982) Reprinted in *Essays of and Information Scientist,* Volume 5, pgs. 721-730 (Philadelphia: ISI Press) (1983).
http://www.garfield.library.upenn.edu/essays/v5p721y1981-82.pdf Back to Document. Ref:7

8. Garfield, E. "Citation analysis as a tool in journal evaluation ," *Science*, 178 (4060):471-479 (1972). Reprinted in:*Current Contents* No. 6, pgs 5-6 (February.7,1973); *Essays of an Information Scientist,* Volume.1, pgs. 527-544 (1977).
http://www.garfield.library.upenn.edu/essays/V1p527y1962-73.pdf Back to Document. Ref:8

Back to Document. Pg.20

9. http://www.issi-society.info/ Back to Document. Ref:9

10. Lock SP. "Journalology": Are the quotes needed?" *CBE Views* 12(4):57-9 (1989).

10. Lock SP. “Journalology”: Are the quotes needed?” *CBE Views* 12(4):57-9 (1989). http://www.garfield.library.upenn.edu/essays/v13p019y1990.pdf Back to Document. Ref:10

11. Pudovkin AI and Garfield E. "Rank-normalized Impact Factor: A way to compare journal performance across subject categories," Presented at the American Society for Information Science and Technology Annual Meeting, Providence, RI. November 17, 2004. Published in *Proceedings of the 67th Annual Meeting of the American Society for Information Science & Technology*, Volume 41, pgs.507-515 (2004) Back to Document. Ref:11 http://garfield.library.upenn.edu/papers/ranknormalizationasist2004published.pdf

12. Garfield E. “Long-term vs. short-term journal impact: Does it matter?” *The Scientist* 12(3):10-2 (February 2, 1998). www.the-scientist.library.upenn.edu/yr1998/feb/research_980202.html Back to Document. Ref:12

13. Garfield E. “Long-term vs. short-term journal impact (part II),” *The Scientist* 12(14):12-3 (July 6, 1998).www.the-scientist.library.upenn.edu/yr1998/july /research_980706.html Back to Document. Ref:13

14. Hansen HB, Henriksen JH. “How well does journal “impact” work in the assessment of papers on clinical physiology and nuclear medicine?” *Clinical Physiology* 17(4):409-18 (1997). Back to Document. Ref:14

15. Garfield E. “Is the ratio between number of citations and publications cited a true constant?” *Current Contents* No. 6, pgs. 5-7 (Feb 9, 1976). Reprinted in *Essays of an Information Scientist*, Volume 2, pgs. 419-421. Philadelphia: ISI Press (1977).. http://garfield.library.upenn.edu/essays/v2p419y1974-76.pdf Back to Document. Ref:15

16. Kresge N, Simoni RD, and Hill RL. “The most highly cited paper in publishing history: Protein determination by Oliver H. Lowry,” *Journal of Biological Chemistry*, 280(28):e25 (July 15, 2005). Back to Document. Ref:16

17a. Garfield E. “Introducing *Citation Classics*: the human side of scientific reports.,” *Current Contents* No. 3, pgs. 6-7 (January 3, 1977). Back to Document. Ref:17a

17b. Commentary by Lowry, OH on ”Protein measurement with folin phenol reagent,” *Current Contents/Life Sciences* (1):7 (January 3, 1977). Back to Document. Ref:17b http://garfield.library.upenn.edu/classics1977/A1977DM02300001.pdf

18. Sambrook J, Fritsch EF, and Maniatis T. *Molecular Cloning : A Laboratory Manual*, 2nd Edition. Cold Spring Harbor: Cold Spring Harbor Press. (1989). Online at www.MolecularCloning.com. Back to Document. Ref:18

19. Yu G, Wang X-Y, Yu D-R. « The influence of publication delays on impact factors,” *Scientometrics* 64(2):235-246 (2005) Back to Document. Ref:19

20. Opthof T. “Submission, acceptance rate, rapid review system and Impact Factor,” *Cardiovascular Research* 41(1):1-4 (1999). Back to Document. Ref:20

21. Garfield E. Which medical journals have the greatest Impact? *Annals of Internal Medicine*, 105(2):313-20 (1986). Available: http://garfield.library.upenn.edu/essays/v10p007y1987.pdf Back to Document. Ref:21

22. Pudovkin AI, Garfield E. "Algorithmic procedure for finding semantically related journals," *Journal of the American Society for Information Science and Technology* (JASIST) 53(13):1113-1119 (November 2002). Back to Document. Ref:22 http://garfield.library.upenn.edu/papers/pudovkinsemanticallyrelatedjournals2002.html

23. Garfield, E. 'Journal Citation Studies. 3. *Journal of Experimental Medicine* compared with *Journal* of Immunology; or how much of a clinician is the immunologist?' *Current Contents* No. *26*, pgs 5-8 (June 28, 1972). Reprinted in *Essays of an Information Scientist*, Volume 1, pgs. Philadelphia: ISI Press (1977). Back to Document. Ref:23

24. http://scientific.thomson.com/products/jpi/ Back to Document. Ref:24

25. Gonzalez L, Companario JM. "Structure of the Impact Factor of journals included in the *Social Sciences Citation Index*: Citations from documents labeled "Editorial Material," *Journal of the American Society for Information Science and Technology, 58(2):252-262(2007).* Back to Document. Ref:25

26. Anonymous. "Journals ranked by impact: Mycology," *inCites* (April 25, 2005). http://in-cites.com/research/2005/april_25_2005-1.html Back to Document. Ref:26

27. Bensman SJ. "Scientific and technical serials holdings optimization in an inefficient market: a LSU serials redesign project exercise," Library Resources & Technical Services 42(3):147-242 (July, 1998). Back to Document. Ref:27

28. www.citationclassics.org Back to Document. Ref:28

29. Hoeffel C. "Journal impact factors" [letter]. *Allergy* 53(12):1225 (December, 1998). Back to Document. Ref:29