Rank-Normalized Impact Factor: A Way To Compare Journal Performance Across Subject Categories

Alexander I. Pudovkin
Russian Academy of Sciences, Institute of Marine Biology, Palchevskogo Street 17, Vladivostok 690041, Russia, Email: aipud@imb.dvo.ru

Eugene Garfield
Institute for Scientific Information, 3501 Market Street, Philadelphia, PA 19104, U.S.A., Email:
Garfield@codex.cis.upenn.edu

Presented at the
American Society for Information Science and Technology Annual Meeting
November 17, 2004
Providence Rhode Island

Journal impact factors have been the subject of considerable controversy ever since their introduction in the seventies. In recent years, *Nature* has made the impact factor a regular matter of controversy. Quite recently there has been a heated discussion of the journal IF in the evaluation of individual scientists and laboratories. In some countries grant application reviews routinely involve the ISI journal IF in considering the applicant's publications (Adam, 2002; Lawrence, 2002; Georgiev, 2003).

When using IF values for evaluatory purposes administrators usually ignore the fact that they greatly differ among subject categories. To overcome the problem of comparing IF across different specialties, Sen and
Marshakova-Shaikevich have suggested using a normalized IF (Sen, 1982; Marshakova-Shaikevich, 1996). However, these normalizations are not quite satisfactory, as they involve either the maximal IF or a few of the highest IFs in each specialty. These "champion" values are not always characteristic of IF values of the majority of journals within the specialty and thus introduce fortuitous elements in the normalized IF.

We suggest a rank normalized IF which involves order statistics for the whole set of journals in a specialty. This normalization procedure, which is similar to percentile ranking, provides more reliable and easily interpretable values we call rank-normalized impact factors or (rnIF).

SLIDE 1.

Calculation of Rank-Normalized Impact Factor

$rnIF_j = (K - R_j + 1)/K$,

where R_j is the *JCR* rank of journal j

and

K is the number of journals in its specialty category.

For the journal GENETICS rank-normalized IF will be

$$rnIF_{Genetics} = (114\text{-}17\text{+}1)/114 = 0.860$$

For any journal j the rank normalized impact factor $rnIF_j = (K - R_j + 1)/K$, where R_j is the *JCR* rank of journal j and K is the number of journals in its category. Keep in mind that within each *JCR* category journals are always displayed in descending order. For example, the journal *Genetics* is the 17th from the top in the *JCR* category for *Genetics & Heredity*. In 2000, this category contained 114 journals. Thus, $rnIF_{Genetics} = (114\text{-}17\text{+}1)/114 = 0.860$. The value of rnIF is very easy to interpret: if a journal j has $rnIF_j = X$ it means that 100% x (1 – X) of the journals in its *JCR* category have higher IF values. So, for the journal *Genetics* 14% of the journals in its category have higher IFs. Under the suggested normalization the top journals in each subject category have rnIF equal to 1.0 and the median journals will have rnIF close to 0.5. When a journal is assigned by the *JCR* to two or more different categories we average the rnIF values.

SLIDE 2

This table presents the IF, SnIF, MnIF and rnIF for journals in five *JCR* categories. Data for six journals in each category are given: for 5 journals with the highest IFs and for the median one. It is clearly seen that IF values vary greatly among the disciplines. There is almost an eighteen-fold difference between IFs for the top journals in the biochemistry/molecular biology category (top IF = 43.4) and the agronomy category (top IF = 2.4). Median IFs for these categories differ less, but nevertheless quite significantly -- almost 4-fold. Variation of the normalized Sen and Marshakova-Shaikevich IF values for journals occupying the same rank position in different categories is also considerable. See coefficient of variation (C.V.) column in Table 1. Our rnIFs are much less variable. The C.V. of rnIF varies from 0.6% to 2.4%, which greatly contrasts with the C.V. values of the *JCR* IF which vary from 57.1% to 86.0%, or of SnIF: from 12.3% to 77.8%, or MnIF: from 2.4% to 22.22%. The other advantage of rnIF is its straightforward interpretation. For example, consider the second highest journal in each *JCR* category. Sen's nIF varies from 63.7 to 100. Marshakova-Shakevich's nIF varies from 92.26 to 137.58. Thus, it is difficult to judge the status of a journal in its subject category by its nIF values. Our rnIFs are more transparent in their meaning. They indicate the proportion of journals in their subject category, which have higher IF values. Thus rnIFs for the journals ranked 2nd in each category range from 0.982 to 0.997, which means there are only .8% to 0.3% of journals with higher IFs.

Average values and coefficients of variation of *JCR* IFs, SnIF, MnIF, and rnIF for 6 journals in 5 different JCR categories (Agronomy, Genetics & Heredity, Biotechnology & Applied Microbiology, Biochemistry & Molecular Biology, Physics, multidisciplinary). Journals in each category include the 5 highest and the median when sorted by IF.

Journal Rank		Agronomy	Genetics & Heredity	Biotechnology & Applied	Biochemistry & Molecular	Physics, Multidisciplinary	Mean	Standard Deviation	C.V.% Coefficient Variation
	Number of journals in JCR category	57	114	134	310	69			
1	IF	2.419	30.910	13.810	43.429	12.774	20.668	16.320	78.96
	SnIF	100	100	100	100	100	100	0	0
	mIF	1	1	1	1	1	1	0	0
	MnIF	116.07	146.04	164.62	153.62	191.95	154.46	24.712	16.00
2	IF	2.418	19.676	11.542	32.440	9.000	15.015	11.536	76.83
	SnIF	100.00	63.7	83.6	74.7	70.5	78.5	14.0	17.83
	mIF	0.982	0.991	0.992	0.997	0.986	0.990	.006	0.60
	MnIF	116.03	92.96	137.58	114.75	135.24	119.31	16.21	13.58
3	IF	2.358	13.810	7.615	27.905	8.756	12.089	9.733	80.51
	SnIF	97.4	44.7	55.1	64.2	68.5	66.0	19.8	30.00
	mIF	0.965	0.982	0.985	0.994	0.971	0.979	0.012	1.23
	MnIF	113.15	65.25	90.77	98.71	131.57	99.89	22.20	22.22
4	IF	1.588	13.450	6.796	26.300	7.110	11.049	9.506	86.03
	SnIF	65.6	43.5	49.2	60.6	55.7	54.9	8.8	16.03
	mIF	0.947	0.974	0.978	0.990	0.956	0.969	0.018	1.86
	MnIF	76.20	63.55	81.01	93.03	106.84	84.12	14.77	17.56
5	IF	1.313	12.912	5.964	18.195	6.462	8.969	6.607	73.66
	SnIF	54.3	41.8	43.2	41.9	50.6	46.4	5.7	12.28
	mIF	0.930	0.965	0.970	0.987	0.942	0.959	0.023	2.40
	MnIF	63.00	61.00	71.09	64.36	97.10	71.31	13.33	18.70
Median	IF	0.500	1.964	0.973	1.882	0.671	1.198	0.684	57.10
	SnIF	20.7	6.4	7.0	4.3	5.2	8.72	6.78	77.75
	mIF	0.509	0.504	0.504	0.502	0.507	0.505	0.003	0.59
	MnIF	23.99	9.28	11.60	6.66	10.08	12.32	6.05	49.11

SLIDE 3

To verify the effectiveness of the proposed normalization scheme we used bibliographic data on the top cited scientists in seven different specialties. ISI regularly publishes data online for the most-cited authors worldwide (see http://www.isihighlycited.com for the latest ten-year period). These data are freely available to all users. We retrieved bibliographic information on the five most recent papers of an arbitrarily chosen person in each of seven specialties in that database: Physics, Animal & Plant Sciences, Molecular Biology & Genetics, Engineering, Immunology, Pharmacology, and Neurosciences -- 35 papers in all, published in 28 journals in 1996-2001. For each journal we determined the rnIF and the two other Sen and Marshakova-Shakevich nIFs. All the information necessary for computation of the three normalized IFs was taken from the 2000 edition of *JCR*. Note that the categories in these two databases, *JCR* vs. *ISI's Highly-Cited*, are not identical. For example, a physicist may have published in journals that are assigned to one or more *JCR* physics categories whereas there is only one physics category in *ISIHighlyCited.com*.

By definition, all the scientists chosen are highly cited. Thus, if our normalization is effective, the average values of rnIF among these scientists should be much more similar than those obtained from the *JCR*. This table displays the average values of *JCR* IF, SnIF, MnIF and rnIF for the seven scientists. One can see that the *JCR* IF values are very different among these top scientists. For example, the average IF for the physicist is 1.992 while the IF average for the immunologist is 18.739, almost a ten-fold difference. The difference in our rnIF is much lower: 0.906 and 0.980. The coefficient of variation (C.V.) of the *JCR* IF is 89.9% while it is only 9.0% for rnIF. Normalized values according to Sen and Marshakova-Shaikevich reduce the differences among disciplines, but the variation is still considerable: C.V. values are 37.3% and 33.4%.

We reiterate, the scientists under consideration are the most-cited authors in their respective fields for the last decade. Not surprisingly, and in accord with their high rank, their papers are usually, but not always, published in the most influential journals. This is revealed by our rnIF: only one paper of 35 (2.9%) was published in a journal with IF less than the median (rnIF = 0.492). Thirty papers of the 35 (85.7%) were published in journals with rnIF higher than 0.82 and thus within 18% of the highest IF journals. Using the *JCR* IF values does not produce these easily interpretable results across fields. Unfortunately, the normalization procedures suggested by Sen (1992) and Marshakova-Shaikevich (1996) do not prove to be sufficiently effective.

Evidently, the efficiency of the suggested normalization depends on the quality of the journal categorization provided by the *JCR*. ISI's heuristic categorization procedure is not ideal. Unfortunately, an ideal categorization procedure is not yet available. This was noted in our recent study of journal relatedness (Pudovkin & Garfield, 2002). The more realistic the categorization, the more efficient the suggested normalization across fields will be. For journals assigned to several categories, the averaging of the rnIF values would require knowledge of the relevance of the journals to the categories. Relevancy weight could be used to improve averaging. We have used equal weights since relevancies were not available. As a concrete example, the *JCR* category for neurosciences includes neurology journals. The latter have, on average, IFs lower than the less clinically and more molecularly oriented journals in neuroscience. This accounts in part for the lower rnIF for the neuroscientist. The rnIF for the neuroscientist would be even higher if papers from the neurology journals were not included.

Average values of IF, nIF after Sen, nIF after Marshakova-Shaikevich, and rnIF for the journals where the 5 most recent papers of 7 top cited scientists in each specialty were published. The categories are those used in the *ISIHighlyCited.com* database. Note that these are somewhat different than *JCR* categories.

Specialty in *ISIHighlyCited.com*	IF	SnIF	MnIF)	rnIF
Animal & Plant Sciences	6.746	43.44	61.98	0.854
Engineering	2.816	81.71	98.81	0.930
Immunology	18.739	63.16	110.45	0.980
Molecular Biology & Genetics	11.190	34.64	58.51	0.930
Neurosciences	3.419	30.09	48.35	0.766
Pharmacology	2.990	46.52	53.34	0.806
Physics	1.992	39.68	64.65	0.906
mean	6.842	48.45	70.87	0.886
ratio max/min	9.41	2.72	2.28	1.28
C.V., %	89.9	37.3	33.4	9.0

SLIDE 4

This table shows time variation of the IF values for the journal *Genetics* from 1997 to 2002. One can see that the rnIF is the most chronologically stable one. It ranges from 0.809 in 2002 to 0.864 in 1998. C.V. for the six-year values equals only 2.42%. The *JCR* IF is the next in its time stability, C.V. equals 5.05%. The values of Sen and Marshakova-Shaikevich are more time variable: C.V. values for them are 18.07% and 5.91%.

The rnIF may also be useful when considering the citation rank of a journal among the global set of journals covered by *JCR*. For example, the 2002 IF of *Genetics* is 4.483. The IF tells us that the average paper published in 2000 or 2001 was cited 4.483 times in 2002. Is that a high citation frequency? Of 5876 journals covered by *JCR* in 2002, how many have higher IF? The global rnIF will provide the answer: rnIF = (5876-302+1)/5876 = 0.948. This means that only 5.2% of journals in the global set have an IF values higher than *Genetics*. The global rnIF tells us that this journal is near the 95th percentile of the distribution of the global set of journals.

Impact Factors for *Genetics* and numbers of journals in *JCR* category of "Genetics & Heredity," 1997 to 2002

Year	*JCR* IF	rnIF	SnIF	MIF	Max IF in *JCR* category Genetics & Heredity	Number of journals	Rank of the journal *Genetics*	Weighted average of IF of top 5 journals
1997	4.275	0.833	11.00	23.35	38.85	90	16	18.31
1998	4.450	0.864	11.03	20.39	40.36	103	15	21.83
1999	4.221	0.838	13.75	23.62	30.69	105	18	17.87
2000	4.687	0.860	15.16	22.14	30.91	114	17	21.17
2001	4.803	0.832	16.23	22.36	29.60	113	20	21.48
2002	4.483	0.809	16.78	24.09	26.71	115	23	18.61
mean	4.4865	0.8393	13.992	22.659	32.855	106.7	18.2	19.877
C.V.,%	5.05	2.42	18.07	5.91	16.62	8.97	16.11	9.03

SLIDE 5

Journal IFs and rnIFs for the 10 journals with the highest IFs in 2002

	Journal	IF	rnIF
1	Annual Review of Imminology	54.455	1.0000
2	Annual Review of Biochemistry	36.278	.9998
3	CA Cancer Journal for Clinicians	32.886	.9997
4	New England Journal of Medicine	31.736	.9995
5	Nature	30.432	.9993
6	Science	28.956	.9991
7	Nature Medicine	28.740	.9990
8	Nature Immunology	27.868	.9988
9	Cell	27.254	.9986
10	Nature Genetics	26.711	.9985

This slide shows 10 journals, which had the highest impact factors in 2002. One can see that these values differed significantly. For instance, the IF of *Annual Review of Biochemistry* ranked 2nd , but is "only" 66.6% of the IF value of *Annual Review of Immunology* which ranked 1st. The IF of *Cell* was half the IF of *Annual Review of Immunology* . Though it is absurd to think that Cell is half as important. Rank-normalized IFs show that all of the 10 journals are above the 99th percentile: in other words these are top journals.

Using three IF values (the standard one, the global rnIF, the specialty rnIF) provide a more complete characterization of the citation rank of a journal. The standard IF gives an absolute measure of citation frequency, regardless of discipline, the global rnIF shows its citation rank relative to the global set , and the specialty rnIF shows its citation rank within its specialty. The global rnIF of 0.948 for *Genetics* testifies that it is indeed a globally high impact journal being among the top 5 or 6% of journals in impact. Its specialty rnIF of 0.809 indicates that it is well cited compared to other journals in the *JCR* Genetics & Heredity category. Its higher global value means that the category includes many fast moving journals where the median journal IF is higher than the median of the global set. For slower moving specialties the opposite relation between the global and specialty rnIF will be true.

SLIDE 6

Normalized Ranking by IF and journal size of 5 journals in different specialties

Journal	IF	Size	GLOBAL RANK rnIF	GLOBAL RANK rnSize	SPECIALTY RANK rnIF	SPECIALTY RANK rnSize
Genetics	4.483	454	0.948	0.960	0.809	0.974
Marine Biology	1.672	238	0.749	0.884	0.836	0.959
Plant and Soil	1.290	281	0.657	0.912	0.845	0.945
Education & Psychological Measurement	1.661	64	0.746	0.494	0.933	0.767
Annals of Mathematics	1.905	53	0.790	0.480	0.988	0.565

Rank normalization can be applied not only to impact factors, but to any other characteristics. In this slide we give rank-normalized values, both global and specialty, of IFs and of sizes for 5 journals of different specialties. The 4 values, global rnIF and rnSize, specialty rnIF and rnSize provide much fuller information that is given by "raw" values of IF and size. For instance, it follows from these 4 characteristics, that *Annals of Mathematics* is near the 80th percentile globally, but it is close to top rank within the *JCR* category of Mathematics, being at 99th percentile. Its size is below the global median, but above the median in its subject category.

Evaluation procedures involving journal IF may be especially important for smaller countries, where establishing evaluation committees that include competent experts across many science fields is difficult. In these situations, the term "poor man's citation analysis," used by Anthony van Raan, seems warranted. The success of using citation indicators in the Research Assessment Exercises in the UK indicates that citation methods are preferable to the often arbitrary and uninformed subjective methods of peer review groups – not to mention the expense (Oppenheim, 1977; Norris & Oppenheim, 2003).

For the evaluation of an individual scientist's or a research collective's contribution, we would recommend using *JCR* IF values in combination with rank-normalized ones. The former are indicative of the unquestionable evidence of professional excellence among the applicant's publications (e.g. papers accepted by *Science*, *Nature*, *PNAS*, or other first rate journals). The mere acceptance by a leading international journal has much greater significance in the developing world. However, consideration of rank-normalized IF would prevent the underestimation of productive scientists from less hot, slower moving specialities. Computation of the suggest rnIF is quite simple since all the necessary values are provided by the *JCR*.

References

Adam, D. (2002). The counting house. Nature, 415, 726-729.

Amin, M., & Mabe, M. (2002). Impact factors: Use and abuse. Perspective in Publishing, No. 1, 1-6.

Garfield, E. (February 2, 1998). Long-term vs. short- term impact: Does it matter? The Scientist, 12(3), 11-12. Retrieved May 26, 2004 from http://www.thescientist.com/yr1998/feb/research_980202.html

Garfield, E. (July 6, 1998). Long-term vs. short-term impact: Part II, The Scientist, 12(14), 12-13. Retrieved May 26, 2004 from http://www.the-scientist.com/yr1998/july/research_980706.html

Garg, K.C., & Padhi, P. (1999). Scientometrics of laser literature as viewed through the Journal of Current Laser Abstracts. Scientometrics, 45, 251-268.

Garg, K.C. (2001). Scientometrics of laser research in India and China. In M. Davis (Ed.), Proceedings of the 8th International Conference on Scientometrics and Informetrics Meeting, University of New South Wales, Sydney (ISSI 2001) (Volume 1, p.167-177). Sydney: Bibliometric & Informetrics Group, University of New South Wales.

Georgiev, G.P., (2003). Strong and efficient. Poisk, 10(720), 5-6 (in Russian).

Lawrence, P.A. (2002). Rank injustice: The misallocation of credit is endemic in science. Nature, 415: 835-836.

Marshakova-Shaikevich, I. (1996). The standard impact factor as an evaluation tool of science and scientific journals. Scientometrics, 25, 283-290.

Marshakova-Shaikevich, I. (2003). Bibliometric evaluation of Russian science journals. Vestnik Rossiiskoi Akademii Nauk (The Herald of Russian Academy of Sciences), 73(9), 788-796. (in Russian).

Norris, M., & Oppenheim, C. (2003) Citation counts and the Research Assessment Exercise V – Archeology and the 2001 RAE, Journal of Documentation 59, 709-739.

Oppenheim, C. (1997). The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaelogy, Journal of Documentation, 53, 477-487.

Pudovkin A.I., & Garfield E. (November 2002). Algorithmic procedure for finding semantically related journals, Journal of the American Society for Information Science and Technology 53, 1113-1119. Retrieved May 26, 2004 from http://garfield.library.upenn.edu/papers/pudovkinsemanticallyrelatedjournals2002.html

Seglen, P. O. (February 15, 1997). Why the impact of journals should not' be used for evaluating research, British Medical Journal 314(7079), 498-502.

Sen, B.K. (1992) Normalized impact factor. Journal of Documentation, 48, 318-329.

Warner, J. (2003) Citation analysis and research assessment in the United Kingdom. Bulletin of the American Society for Information Science and Technology, 30, 26-27.