Essays

The Journal Impact Factor as a performance indicator

Frank-Thorsten Krell
Denver Museum of Nature & Science, 2001 Colorado Boulevard, Denver, Colorado 80205, U.S.A.; frank.krell@dmns.org

Abstract

The Journal Impact Factor is the most commonly applied metric for evaluation of scientific output. It is a journal-focused indicator that shows the attention a journal attracts. It does not necessarily indicate quality, but high impact factors indicate a probability of high quality. As an arithmetic mean of data originating from all authors of a journal with a high variance, it is inapplicable to evaluate individual scientists. For quantifying the performance of authors, author-focused citation metrics are to be used, such as the h index, but self-citations should be excluded ("honest h index" h_h). All citation metrics suffer from the incompleteness of the databases they source their data from. This incompleteness is unequally distributed between disciplines, countries and language-groups. The Journal Impact Factor has its limitations, but if those limitations are taken into consideration, it is still an appropriate indicator for journal performance.

Keywords Journal Impact Factor; the honest h index; citation databases; attention; quality; bibliometrics

The "impact factor" is the most commonly used metrical indicator for quality, performance and impact in science, often applied without critical assessment of what it is actually indicating. The impact factor has extensively penetrated academia and academic publishing, which has provoked change in publishing strategies by academic publishers and editors[1,2] and in authors' publishing behaviour.[3,4] Editors and publishers strive to increase their journals' Impact Factors. Authors, often under perceived or real pressure from their administration,[5,6] choose publication venues according to the values of the Journal Impact Factor. "Massaging" impact factors up by means beyond scholarly quality, such as increased self-citation by authors and journals, creating a higher number of mutually referenced papers from the same body of evidence, timing publications to have maximum exposure for accruing citations, and increasing the number of citation-attracting review papers, has become a common practice; as has the misapplication of the Journal Impact Factor for evaluating research performance of single researchers, institutes or other entities. The body of literature dealing with this phenomenon and imminent problems is substantial and growing. Here, I refrain from attempting a comprehensive review of all problems, manipulation techniques and misapplications of the Journal Impact Factor, but will point to a few crucial aspects and misunderstandings of this pervasive metric.

Journal Impact Factor: definition and coverage

What is commonly called "the impact factor" is short for the latest two year Journal Impact Factor calculated annually in the *Journal Citation Reports™* by Thomson Reuters. It is defined as the number of citations within a given year to items published by a journal in the preceding two years divided by the number of citable items published by the journal in these two years.[7] It is the average number of citations a paper of a journal attracts in the two years following its publication.

The database from which these numbers are sourced is Thomson Reuter's *Web of Science* which currently covers almost 12,000 active journals and over 3,000 proceedings volumes[7]. This is up from 8,684 titles in 2000,[8] but it is still only a third of the scientific serials listed in *Ulrichsweb™* which is incomplete itself. For disciplines in which Bradford's law or Garfield's law of concentration[9] apply and most citations refer to a limited number of core journals, this coverage might be exhaustive. Such fields are, eg molecular biology and biochemistry, biological sciences related to humans, chemistry, and clinical medicine.[10] For other disciplines with more equally distributed relevance of journals or higher relevance of book publications, the *Web of Science's* coverage is rather insufficient (eg for natural history,[11] regionally focused science,[12] taxonomy,[13] mathematics, economics, humanities & arts[10]). In general, the Journal Impact Factor considers how often journals are cited in a selective number of journals. By definition, it does not cover the complete impact of a journal.

The portion it misses depends on the discipline of the journal. On pages 126-130 of his book *Citation Analysis in Research Evaluation*, Henk Moed[10] compiled lists of coverage by disciplines and countries. Coverage can be as low as 64% in ecology, 55% in geology, 45% in nursing, 33% in information & library sciences, and 9% in history. Although Moed gives a coverage of 67% for my own research field, zoology, in 2009 *Web of Science* captured only 25.7% of citations of my own papers.[11] Thomas Nisonger, a library and information scientist, found in 2004 that 42.4% of his print citations were retrieved by *Web of Science*[14]. With the expansion of the coverage of *Web of Science*,[8] these percentages will go up, but as long as coverage is selective, some disciplines will be disadvantaged.

What performance does the Journal Impact Factor indicate?

The Journal Impact Factor was created by Irving H. Sher and Eugene Garfield in the 1960s "to help select journals for the *Science Citation Index*".[15] It is a simple index, easy to understand and to calculate, that allows comparing journals of any size in terms of citations they attract. By proxy of citations, it indicates the use of journals in scientific research or, in other words, the attention a journal

receives. Since the purpose of journals is to be read and used in scientific research, the Journal Impact Factor is an apt indicator for journal performance. It is only the short-term performance that the Journal Impact Factor reflects though. Since the second year after publication is the year attracting most citations of any year, if the whole database is considered,[10] this short term performance is indicative for the overall performance of many journals. However, the majority of journals reach their citation peak *after* the window that the Journal Impact Factor considers[16] with most journals attracting 70-90% of all citations after the second year.[17] In some disciplines, papers one or two years old are rarely cited, for example, in my own subfield, taxonomy.[13]

Since 2007, Thomson Reuters has been providing a five-year Journal Impact Factor which slightly mitigates the underestimate of the two years citation window by increasing the impact factor for the majority of disciplines[18] and journals[19] taking into account their peak citedness. Nonetheless, use of the 'classical' two-year Journal Impact Factor continues to dominate evaluation and marketing of journals.

As an arithmetic mean for the whole journal, the Journal Impact Factor cannot predict the performance of single papers. In fact, the variation in number of citations to articles of the same journal can be several magnitudes. Articles of the 1998 volume of *The Lancet* were cited from zero to 2,799 times.[20] The majority of *Nature* papers from the years 2002 and 2003 received under 20 citations in 2004; 2.7% of the papers received over 100 citations with a record holder with 522 citations.[21] In 2009, a single paper attracting 5,624 citations pushed the impact factor of *Acta Crystallographica A* up from under 3 to 49.93, with all other papers of the journal having attracted three or less citations.[22] Such variation renders attempts to use Journal Impact Factors for evaluation of single papers or authors absurd. The Journal Impact Factor reflects performance of a scholarly journal and nothing else. Can we consider this performance as a proxy for quality of the journal?

Quality? Relevance? Attention!

To answer this question, we need to explore what reasons and motives stand behind citations. Citation motives and behaviour have been studied since the 1970s.[2,23-26] Good quality of a paper is never the sole reason for a citation whereas bad quality can be a good reason not to cite a paper or to cite it as a bad example, or to propose corrections of published errors. The primary reason for citing a paper is or should be that it underpins or at least relates in some useful way to the facts one is writing down. If there are only a few sub-standard studies preceding one's own study, they need to be cited. If there are five bad and two good studies available to cite, then the good ones will be chosen. If the authors of one of the good studies are personal competitors or enemies, one might cite the other study. Collaborating teams tend to cite each other, because of early awareness of the others' results, but also because they want to support each other or because they thank each other with citations. Scientists are humans who act socially (or sometimes antisocially), whether they do so subconsciously

or deliberately. Increasing competitiveness in the research environment fosters selfish behavior. While authors in the pre-impact factor times cited their own publications to embed their studies in their broader research program, to draw attention to their own work, or out of self-adulation, now they become increasingly aware that self-citation helps all sorts of citation metrics. Self-citations, at the journal level, became a strategy to improve the Journal Impact Factor of the journal one publishes in (or one edits).[2] At the author level, it improves the standing of the author by increasing author-focused metrics, as long as self-citations are included in the citation analysis (which they should not be[11]).

Even if the choice of references to cite is far from an objective, quality-oriented process, the few studies comparing peer judgment with citation metrics often found positive correlations,[27-29] particularly at the level of research groups and single papers. One has to be cautious though. Baird and Oppenheim[25] aptly stated: "So, does this mean that if an author writes an article, and it is highly cited, then it is important? No it does not. Rather, what it means is the chances are the paper is important. […] In other words, high citation counts mean a statistical likelihood of high quality research." It is unknown and hardly possible to quantify how high the likelihood is. At the journal level, citations are a quality indicator only in a very crude sense, in distinguishing (with a certain, but unknown probability) established, reputable journals from minor quality outlets of the same discipline. A journal with an impact factor of 5 is likely to have attracted and to continue to attract higher quality papers than a lesser used journal in the same discipline with an impact factor of 0.7. A slight difference of impact factors, eg 1.6 and 1.9 are unlikely to have any meaning beyond variability.

To whatever extent quality can be derived from citation counts, it is undeniable that the citation rate gives evidence for the attention a journal attracts. A high attention shows that a journal is useful and predicts that others will want to consult this journal. The purpose of the Journal Impact Factor, to determine which journals will be of interest to most, is fulfilled. By which motives this attention is achieved is primarily irrelevant.

Evaluating single authors

For the evaluation of single authors author-focused indices are to be used, which are calculated on the basis of citations of only the author to be evaluated. It seems that the prerequisite for wide acceptance of such an index is its simplicity, not necessarily its sophistication. For almost every letter of the alphabet, a citation based index has been proposed. Of those *a*-, *b*-, *c*-, *d*-, *e*-, *f*-, *g*-, *h*-, *j*-, *k*-, *L*-, *m*-, *n*-, *p*-, *q*-, *r*-, *t*-, *u*-, *v*-, *w*-, *x*-, *y*-, and *z*-indices, some of them admittedly very new, only the *h*-index[30,31] has gained widespread use. It is probably the most simple, author-focused index, defined as the number of papers of an author with citation number $\geq h$. It has its disadvantages, particularly for younger scientists with lower publication numbers, but it is at least based on the author's publications. Since it can easily be manipulated by strategic self-citations,[32] I suggested, as has Schreiber[33] before, to exclude self-citations from its calculations and use

what I called "the honest *h* index (*h*ₕ)".[11] This is the sort of metrics that should be applied for evaluation of individuals' research performance, not a journal-focused indicator.

Attention fully covered by citations?

The value of those author-focused indices likewise depends on the database from which citations are extracted. The *h* index of the same scientist can easily be three times higher if another database is used.[11,34] Currently, we have only incomplete, but growing databases[11] available: *Web of Science, SciVerse® Scopus, Google Scholar*. As long as a scientist does not compile his own comprehensive list of citations[11] from which citation metrics are calculated, we have to keep in mind that any citation metrics derive from incomplete data sets with an unknown extent of incompleteness. The extent of incompleteness can differ largely depending on, amongst others, discipline, location and language of the scientist.[10,35]

Besides database incompleteness, we also need to keep in mind that citations represent only a part of the attention a publication attracts. Particularly publications targeted at end-users, such as clinical papers for medical practitioners,[25,35] or identification keys for animals or plants, are likely to be frequently used, but not necessarily cited. No correlation was found between the citation count and photocopy requests in certain social work journals.[25] MacRoberts and MacRoberts[36] found that biogeography source papers from which data are derived remain extensively non-cited. Purely citation-based evaluation would lead to a skewed picture of the overall relevance of such papers or whole journals.[35,36] However, other studies[37] show a strong positive correlation between downloads and later citations.

Conclusion

The Journal Impact Factor is an appropriate means to evaluate journal performance since it indicates the attention a journal attracts, with the provision that some types of works are used without getting cited. The Journal Impact Factor, if high, indicates a chance that the journal published high quality papers. For the evaluation of individual researchers, journal-focused metrics are inapplicable. Author-focused metrics, such as the *h* index, are to be used. For any citation-based evaluation, we need to consider the extent of incompleteness of the data source and the circumstances of the entity to be evaluated, namely discipline, location, language-group which influence the number of citations that papers attract.

Competing interests None declared.

Note

Despite the author's intent to refer to current papers, only 10 of the following 38 references would count for the two-year Journal Impact Factor were *European Science Editing* considered as a source journal by *Web of Science*. For the five-year Journal Impact Factor, it would be 21 references. Since *European Science Editing* currently is not considered by *Web of Science*[38], none of these references count for the Journal Impact Factor of the cited journals.

References

1 Brown H. How impact factors changed medical publishing–and science. *British Medical Journal* 2007;334:561-563. doi: 10.1136/bmj.39142.454086.AD

2 Krell F-T. Should editors influence journal impact factors? *Learned Publishing* 2009;23(1):59-62. doi: 10.1087/20100110

3 Steele C, Butler L, Kingsley D. The publishing imperative: the pervasive influence of publication metrics. *Learned Publishing* 2006;19:277-290. doi: 10.1087/095315106778690751

4 Lawrence PA. The mismeasurement of science. *Current Biology* 2007;17(15):R583-R585. doi: 10.1016/j.cub.2007.06.014

5 Adam D. The counting house. *Nature* 2002;415:726-729. doi:10.1038/415726a

6 Abbott A, Cyranowski D, Jones N, Maher B, Schiermeier Q, Van Noorden R. Do metrics matter? *Nature* 2010;465:860-862. doi:10.1038/465860a

7 Hubbard SC, McVeigh ME. Casting a wide net: the Journal Impact Factor numerator. *Learned Publishing* 2011;24:133-137. doi: 10.1087/20110208

8 ThomsonReuters. Web of Science coverage expansion. http://community.thomsonreuters.com/t5/Citation-Impact-Center/Web-of-Science-Coverage-Expansion/ba-p/10663; posted 27 April 2010 [accessed 2011 December 28].

9 Garfield E. Bradford's law and related statistical patterns. In: Garfield, E. *Essays of an Information Scientist. Volume Four 1979–1980.* Philadelphia, PA: ISI Press, 1981:476-483. http://www.garfield.library.upenn.edu/essays/v4p476y1979-80.pdf

10 Moed HF. Citation Analysis in Research Evaluation. Dordrecht: Springer, 2005.

11 Krell F-T. The poverty of citation databases: data mining is crucial for fair metrical evaluation of research performance. *BioScience* 2009;59(1):6-7. doi: 10.1525/bio.2009.59.1.2

12 Martín J, Gurrea P. La Entomología en España y las revistas incluidas en el Science Citation Index. *Boletín de la Asociación Española de Entomología* 2000;24(3-4):139-156. http://www.entomologica.es/index.php?d=publicaciones&num=54&w=1078&ft=1

13 Krell F-T. Why impact factors don't work for taxonomy. Its long-term relevance, few specialists and lack of core journals put it outside ISI criteria. *Nature* 2002;415:957. doi:10.1038/415957a

14 Nisonger TE. Citation autobiography: an investigation of ISI database coverage in determining author citedness. *College & Research Libraries* 2004;65:152-163. http://crl.acrl.org/content/65/2/152.full.pdf+html

15 Garfield E. Journal impact factor: a brief review. *Canadian Medical Association Journal* 1999;161(8):979-980. http://www.ecmaj.ca/content/161/8/979.full.pdf+html

16 Moed HF, Leeuwen TN van, Reedijk J. A new classification system to describe the ageing of scientific journals and their impact factors. *Journal of Documentation* 1998;54(4):387-419. doi: 10.1108/EUM0000000007175

17 Moed HF, Burger, WJM, Frankfort JG, Raan, AFJ van. The application of bibliometric indicators: important field- and time-dependent factors to be considered. *Scientometrics* 1985;8(3-4):177-203. doi: 10.1007/BF02016935

18 Nierop E van. The introduction of the 5-year impact factor: does it benefit statistics journals? *Statistica Neerlandica* 2010;64(1):71-76. doi: 10.1111/j.1467-9574.2009.00448.x

19 Campanario JM. Empirical study of journal impact factors obtained using the classical two-year citation window versus a five-year citation window. *Scientometrics* 2011;87:189-204. doi: 10.1007/s11192-010-0334-1

Fragments of academic publishing in Estonia

Mare-Anne Laane
Lecturer, Tallinn University of Technology, Estonia; mareanne.laane@ttu.ee

This fragmentary text introduces the venue of EASE Tallinn 2012 conference in the context of Estonian academic publishing.

Estonia's printing press came into being in 1631 at the secondary school of Tartu, the predecessor of the University of Tartu, which was founded on 30 June 1632 by the Foundation Decree of Academia Dorpatensis, signed by King Gustav II Adolf of Sweden.

Main building of the University of Tartu (Photographer: Andres Tennus)

The University of Tartu History Museum showcases the history of science and publishing as well as university education from the 17th century to the present day. The museum occupies the former university library, which was built in the choir of the ruins of a former dome cathedral.

University of Tartu History Museum (Photographer: Andres Tennus)

Tallinn University of Technology (founded in 1918), the venue for the 2012 EASE Conference, is the second largest

20 Kostoff, RN. The difference between highly and poorly cited medical articles in the journal Lancet. *Scientometrics* 2007;72(3):513-520. doi: 10.1007/s11192-007-1573-7

21 Campbell P. Escape from the impact factor. *Ethics in Science and Environmental Politics* 2008;8:5-7. doi:10.3354/esep00078

22 Dimitrov JD, Kaveri SV, Bayry J. Metrics: journal's impact factor skewed by a single paper. *Nature* 2010;466:179. doi:10.1038/466179b

23 Bavelas JB. The social psychology of citations. *Canadian Psychological Review* 1978;19(2):158-163. doi: 10.1037/h0081472

24 Bonzi S. Characteristics of a literature as predictors of relatedness between cited and citing works. Journal of the American Society for Information Science 1982;33(4):208-216.

25 Baird LM, Oppenheim C. 1994. Do citations matter? *Journal of Information Science* 1994;20(1):2-15. doi: 10.1002/asi.4630330404

26 Bornmann L, Schier H, Marx W, Daniel, H-D. What factors determine citation counts of publications in chemistry besides their quality? *Journal of Infometrics 2012*[2011];6:11-18. doi:10.1016/j.joi.2011.08.004

27 Rinia EJ, Leeuwen TN van, Vuren HG van, Raan AFJ van. Comparative analysis of a set of bibliometric indicators and central peer review criteria; evaluation of condensed matter physics in the Netherlands. *Research Policy* 1998;27:95-107. doi:10.1016/S0048-7333(98)00026-2

28 Oppenheim C, Summers MAC. Citation counts and the Research Assessment Exercise, part VI: Unit of assessment 67 (music). *Information Research* 2008;13(2) paper 342. http://InformationR.net/ir/13-2/paper342.html [accessed 2011 December 31]

29 Patterson MS, Harris S. The relationship between reviewers' quality-scores and number of citations for papers published in the journal Physics in Medicine and Biology from 2003–2005. *Scientometrics* 2009;80(2):343-349. doi: 10.1007/s11192-008-2064-1

30 Hirsch JE. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the USA* 2005;102(46):16569-16572. doi: 10.1073/pnas.0507655102

31 Bornmann L, Marx W. The h index as a research performance indicator. *European Science Editing* 2011;37(3):77-80.

32 Bartneck C, Kokkelmans S. Detecting h-index manipulation through self-citation analysis. *Scientometrics* 2011;87:85-98. doi: 10.1007/s11192-010-0306-5Open Access

33 Schreiber M. Self-citation corrections for the Hirsch index. *Europhysics Letters* 2007;78: paper 30002. doi:10.1209/0295-5075/78/30002

34 Bar-Ilan J. Which h-index? – A comparison of WoS, Scopus and Google Scholar. *Scientometrics* 2008;74(2):257-271. doi: 10.1007/s11192-008-0216-y

35 Raan AFJ van, Leeuwen TN van, Visser MS. Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics* 2011;88:495-498. doi: 10.1007/s11192-011-0382-1

36 MacRoberts MH, MacRoberts BR. Problems of citation analysis: a study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology* 2010;61(1):1-13. doi: 10.1002/asi.21228

37 Watson AB. Comparing citations and downloads for individual articles. *Journal of Vision* 2009;9(4):1-4. doi: 10.1167/9.4.i

38 Gasparyan AY. Get indexed and cited, or perish. *European Science Editing* 2011;37(3):66.