Is the journal impact factor a valid indicator of scientific value?

Oh H C, Lim J F Y

The ultimate goal of scholarly scientific publishing is to contribute to the advancement of knowledge and its application into improving life. However, this fundamental tenet of research and scholarship may have been forgotten in modern research. While competition among institutions and individuals striving for excellence are the driving force that has accelerated progress in research undertakings, the intense competition for limited research funds has unfortunately led to the perpetuation of a "publish or perish" philosophy within the scientific community. It is this same competition and the need for an "objective" measure of research quality that has led to the introduction and subsequent widespread acceptance of the journal impact factor (JIF) as a measure of the scientific quality of journals. Over the years, the JIF has become probably the most important indicator of research quality. It has induced researchers to publish in a few high-profile journals with high impact factors, with the career advancement of many researchers becoming more dependent on where their articles are published rather than on what they publish. This unquestioning reverence of the JIF may potentially inhibit the overall progress of the research community as a whole by focusing the research community's efforts on an increasingly narrow agenda and by discouraging many erstwhile promising researchers in so-called esoteric spheres to pursue other interests, as their research holds little appeal to journals with high impact factors. It is therefore timely that Kumar et al have highlighted the dangers of ill-informed usage of the JIF and the potential for misuse and abuse.[1] The authors decry the almost exclusive use of the JIF to assess the scientific quality of individual publications and the resultant consequences on scientist behaviour and the overall advancement of science.

Over the years, JIFs have gained popularity and are increasingly accepted as scientific quality measures of journals despite being the subject of considerable controversy in the academic community.[2-6] Based on the notion that a journal is representative of its articles, the impact factors of journals in which an author's articles have been published are commonly employed to evaluate the author's scientific achievement. This same notion has in turn also led to widespread acceptance of JIF-based assessment frameworks in the allocation of research resources and grants, performance evaluation of researchers and academic position appointments in several countries including Canada, Italy, Hungary and Russia. As a general rule, journals with high impact factors include the most prestigious ones. Inevitably, this has led to a common perception within the research community of equating prestige with high impact factor journals and the acceptance of an article for publication by a high impact factor journal as an implied indicator of prestige. In order to capitalise on the growing attention bestowed to JIFs by the research community and governing bodies, it has also become increasingly common among publishers to use JIFs extensively in the marketing of their respective journals. For example, major publishers like Elsevier,[7] Blackwell Publishing[8] (acquired by Wiley in 2007), Bentham Science,[9] and IGI Global[10] have all used their impact factor to portray the scientific significance of their respective journals in their marketing collaterals.

Despite the widespread application of JIFs within the research community or perhaps precisely because of this widespread application, it is meaningful for any research institution or governing body to critically examine this indicator before they jump on the JIF bandwagon and perpetuate the belief that JIF is a rightful measure of scientific value of a researcher's publications. Although an article may be cited due to the criticism of other authors, there is a general consensus within the research community on the direct correlation between the scientific quality and citation count of an article.[11] However, the number of citations that an article receives is dependent on how long it has been available in its journal. In many instances, articles which are recently published may not have had enough time to be cited. As such, research grant-makers, academic institutions and governing agencies often use the impact factors of the journals in which the aforementioned articles are published as surrogate scientific value evaluation measures. Intuitively, such a move is legitimate if: (1) the impact factor of a journal matches closely with the actual citation count of its articles, and (2) journals from different research disciplines with similar scientific value share similar impact factor scores.

Centre for Health Services Research, Singapore Health Services, 226 Outram Road, #03-01, Block A, Singapore 169039

Oh HC, BEng, MEng
Policy Analyst

Lim JFY, MBBS, MPH, MRCSE
Director

Correspondence to:
Dr Jeremy Lim Fung Yen
Tel: (65) 6323 8201
Fax: (65) 6323 2902
Email: jeremy.lim.f.y@singhealth.com.sg

Unfortunately, the impact factor of a journal is not statistically representative of its individual articles. Both Seglen and Garfield reported poor correlation between the impact factor of a journal and the actual citation rates of its individual articles.[12,13] Moreover, there are also variations in the impact factor of journals within a specific research discipline, which is not attributed to differences in the scientific values of their respective articles. For example, a review article tends to be cited more frequently than non-review articles. As a result, a journal with a number of review articles tends to have a higher impact factor compared to another journal in the same discipline but with fewer review articles. Evidently, such a difference in JIFs does not truly reflect disparities in the scientific value of these journals.

In addition, there is also a distinct variation in two key areas of citation kinetics across research disciplines that make meaningful comparison of journals from different fields on the basis of their respective impact factors impossible. First, the mean number of references cited per source article (citation density) varies significantly across research disciplines. For example, the citation density of articles in biochemistry journals is almost twice that of articles in mathematics journals. This has somehow contributed to the higher impact factors of biochemistry journals relative to those of mathematics journals. Second, authors from different research disciplines usually vary in the speed at which they cite articles upon the publication of these articles. This difference in speed of citation is clearly reflected by the half-lives of journals from different fields. The half-life of a journal is defined as the number of retrospective years required to find 50% of the cited articles in the journal. Since the computation of JIF is based on a two-year time window, journals of research disciplines like physics with short half-lives tend to have more favourable impact factor scores than those of disciplines like physiology with longer half-lives. Therefore, the differences in impact factors of journals from different research disciplines may not necessary be attributed to the differences in their scientific values.

Most journals have articles which attract citations by publications from journals of other research disciplines, with some research disciplines receiving many more of such cross-disciplinary journal citations than others. This is clearly illustrated by comparing journals of basic and clinical science. Results of clinical science research tend to be based on findings of basic science research and not vice versa. Thus, articles from journals of basic science are cited more extensively by articles from journals of clinical science as compared to the number of citations received by clinical science articles from basic science publications. This has contributed to the higher impact factors of basic science journals compared to those of clinical science journals. Again, such differences in JIFs are not attributed to the differences in scientific values of the journals involved.

From the above discussion, we submit that the JIF is not an appropriate metric to measure the scientific quality of individual articles. On the contrary, the JIF has evolved from the original intention of using it as a measure of journal performance. In fact, Garfield has warned extensively against employing the impact factor to evaluate individual articles and scientists.[13-15] To extend the role of the JIF so that it also becomes a surrogate scientific value evaluation measure of articles is not only unfair to the authors involved, but it may also induce an unhealthy research culture which may inhibit the overall progress of the research community. In addition, research disciplines with journals that have low impact factors may in turn fail to attract adequate attention from the researcher community. For example, Singapore's government-linked Agency for Science, Technology and Research (A*STAR), which actively nurtures public sector research and development in biomedical sciences, physical sciences and engineering, uses the number of publications in journals with JIFs greater than 3.5 as one of the criteria in performance assessment of its researchers. Although such explicit criteria may promote research activities with eventual publications in journals with high impact factors, it also deters research development in disciplines like medical informatics, nursing, anatomy and morphology, where their respective highest JIFs are below 3.5, according to the latest Journal Citation Report, which is available to all subscribers of ISI (Institute for Scientific Information) Web of Knowledge℠.

Over the years, several approaches have also evolved to address the limitations posed by JIFs in the scientific valuation of publications. For example, many research institutions and universities are assessing the scientific achievements of their researchers and faculty staff based on total citation counts of their publications, the average citation count per publication and *h*-indices (Note: the *h*-index of an author refers to the number of published articles that have at least *h* citations each). Although these measures are independent of the impact factors of the journals where the publications are accepted, they provide experienced researchers with an unfair edge over new researchers. As such, it may be more appropriate if the aforementioned metrics are only compared among researchers with a similar number of years of experience. In addition, there are also other organisations that classify journals into tiers based on their impact factors and the opinion of a panel of experts. With the list of classified journals, the scientific

performance of the research staff is then assessed based on the tier levels of journals in which their publications appear. Although such an approach entails a subjective mechanism of classifying journals, it offers an effective way of addressing the limitations of the JIF attributed to research discipline-induced differences in JIF scores if the journal classification exercise is carried out in a judicious manner by the panel of experts.

A comprehensive scientific evaluation of an article requires a multidimensional assessment, which may entail multiple measures. These include the clinical or policy impact of the article, its number of citations in textbooks, etc, which cannot be captured by JIFs. Such a multidimensional assessment is beyond the scope of any single metric such as the JIF. Due to their inherent differences in operational goals and constraints, there is no one-size-fits-all set of metrics that can meet the scientific assessment needs of all research organisations. Therefore, each organisation has to decide its own set of metrics to be used in the scientific evaluation of its researchers. Intuitively, it is clear that the guiding principle of deciding which metrics to use should be consistent across all research organisations where the ultimate goal of scholarly scientific publishing is preserved and all research disciplines are accorded equal scientific importance. Nevertheless, the determination of the right scientific valuation metrics to use in any research organisation is more of an art than a science which relies heavily on the judgment of its leadership and the organisation's strategic priorities. Since the choice of metrics to use will play a pivotal role in shaping the future research trends and landscape of a research organisation, it is therefore crucial that an informed decision is made with a clear overview of all available metrics, as well as a good understanding of their respective functions and limitations.

REFERENCES

1. Kumar V, Upadhyay S, Medhi B. Impact of impact factor in biomedical research, its use and misuse. Singapore Med J 2009; 50: 752-5.
2. Moed H, van Leeuwen TN. Impact factors can mislead. Nature 1996; 381:186.
3. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997; 314:498-502.
4. Hussain FN, Husnain A. Impact factor: is it an essential tool to assess a medical journal? J Coll Physicians Surg Pak 2007; 17:382-3.
5. Mishra BK, Patel S. On the controversy of impact factor. Curr Sci 2009; 96:324.
6. Singaravelan N. Looking beyond number of publications and journal impact factor. Curr Sci 2009; 96:327-8.
7. Elsevier. Elsevier announces 2007 journal impact factor highlights. Available at: www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_00961. Accessed May 28, 2009.
8. Blackwell Publishing. Blackwell announced high growth in impact factor, the growth in citations of Blackwell journals surpasses other major publishers. Available at: www.eworldwire.com/pressreleases/14997. Accessed May 28, 2009.
9. Bentham Science Publishers. Bentham Science: an introduction. Available at: www.bentham.org/. Accessed May 28, 2009.
10. IGI Global. IGI Global's Journal of Database Management recognized as a top-ten journal in its discipline. Available at: www.igi-global.com/journals/JDM-ThomsonScientific.asp. Accessed May 28, 2009.
11. Jeang KT. H-index, mentoring-index, highly-cited and highly-accessed: how to evaluate scientists? Retrovirology 2008; 5:106.
12. Seglen PO. Casual relationship between article citedness and journal impact. J Am Soc Inf Sci 1994; 45:1-11.
13. Garfield E. The history and meaning of the journal impact factor. JAMA 2006; 295:90-3.
14. Garfield E. How can impact factors be improved? BMJ 1996; 313:411-3.
15. Garfield E. Journal impact factor: a brief review. CMAJ 1999; 161:979-80.