

Impact of the impact factor in biomedical research: its use and misuse

Kumar V, Upadhyay S, Medhi B

ABSTRACT

The impact factor was created in the biomedical research field in order to measure a journal's value by calculating the average number of citations per article over a period of time. It was initially developed to help libraries decide which highly-cited journals to subscribe to. However, at present, it is being misused to judge the quality of a researcher or medical scientist as well as the quality of the work done. It contains serious sources of errors and flaws, resulting in strong biases against culture- and language-bound medical subspecialties. The present article is aimed to highlight the impact of the impact factor in the biomedical research, as well as its use and misuse.

Keywords: impact factor, Institute for Scientific Information, Journal Citation Report, Science Citation Index

Singapore Med J 2009;50(8):752-755

INTRODUCTION

The impact factor (IF) was initially designed by Eugene Garfield in the 1950s. It was introduced to the scientific community as an assessment tool to evaluate the value of a scientific journal, by calculating the number of citations of an article published in a particular journal over a specific time period. The term, impact factor, was only first used in 1961, after publication in the Science Citation Index (SCI) in 1963.⁽¹⁾ Presently, it is popularly referred to as Journal Citation Reports (JCR). In general, the number of citations of a particular article indicates only the mean interest of scientists for that article. Thus, the IF highlights the average interest in an article that gets published in the journal. However, most of the data utilised in the JCR are IF ratios obtained from dividing the number of citations received in one year on articles published the preceding two years (numerator) by the number of papers published over the two previous years (denominator). The JCR's impact calculations are based on original research and review articles, as well as on short notes.

Since the early 1960s, the IF has been used in the scientific community as the primary tool for judging the

quality of research. Presently, there are more than 5,000 journals from various specialties worldwide that are published annually in the JCR of the Institute for Scientific Information (ISI) in Philadelphia, USA.⁽²⁾ The recent trend in the scientific world is for scientists or researchers to have a good number of publications with citations if they want to become widely recognised. Thus, it has become an imperative for scientists to publish their work in journals with a high IF in order to accomplish this aim. Moreover, this criterion has led to the development of a long-held belief that all that counts is the number of publications. By this reasoning, a high number of publications make scientists visible and therefore distinguished in their field.⁽³⁾

The IF is calculated by the ISI, which is a part of the Thomson company, and is basically a by-product of computerised databases of the Social Science Citation Index (SSCI) and the SCI.⁽⁴⁾ The Association of the Scientific Medical Societies in Germany (AWMF) also uses its recommendations to evaluate scientific achievement on the basis of the IF.⁽⁵⁾ The calculation of the IF of a journal for a particular year depends on the average number of citations of an article that was published in that journal over the last two years from all the published articles in that year,⁽⁶⁾ i.e., if the IF of a journal is 2.0 in 2008, this reflects that on average, the articles published in 2006 and 2007 were cited twice among the collection of all ISI-indexed journals published in 2008.

EVALUATION OF THE IMPACT FACTOR

There are currently various criteria for evaluating the IF, but awareness within the scientific community is negligible. The diagnostic purpose of the IF is to evaluate the significance of a publication in the scientific community and also to assist in the smooth functioning of libraries. However, this should not be exclusively used to guide research activities and in the evaluation of the achievements of scientists involved in active research.⁽⁷⁾ Besides this, the IF is also used by scientists to judge the quality of a journal by citing articles from it. Thus, if the same article from a journal has been cited frequently, the significance of the journal will increase (e.g. protein estimation by Lowry et al⁽⁸⁾). Hence, researchers may aim to publish in journals whose quality is based on their IF.

Department of Microbiology, Panjab University, Sector 14, Chandigarh 160014, India

Kumar V, BSc, MSc, PhD
Senior Research Fellow

Department of Physiology

Upadhyay S, BSc, MSc, PhD
Senior Lecturer

Department of Pharmacology, Postgraduate Institute of Medical Education and Research, Chandigarh 160012, India

Medhi B, MBBS, MD, FIMSA
Associate Professor

Correspondence to:
Dr Bikash Medhi
Tel: (91) 172 275 5250
Fax: (91) 172 274 4401
Email: drbikashus@yahoo.com

Since the IF is calculated by the ISI, the institute is responsible for monitoring the citations listed in its databases. These are compared with the original and review articles from the journals listed with the ISI. Therefore, some quotations which cannot be assigned to a specific paper fail to undergo further checks. Journals that are listed by ISI as original journals can make remarkable changes to their IF by various means. The usual trend for an editorial board of a journal is to publish different kinds of articles such as reviews, original research articles, news, obituaries and book reviews, so that they can increase the IF significantly.⁽⁹⁾ Sometimes, a journal may request authors to include references from its own previous publications in order to increase its IF. Thus, there is a great degree of manipulation in the ISI criteria for evaluation of the IF.

The IF itself is a ratio, where the numerator is made up of citations of articles published in original journals over the last two years, while the denominator is made up of original papers and review articles published within the same time period. The IF is calculated by the ISI, which only counts original papers and review articles in the denominator but accepts all other published material for use in the numerator. This has caused various journals to implement a policy of publishing letters to the editor, editorials, congress reports and book reviews in order to increase their IF.

In 2002, Moed conducted a study that showed a systematic error rate, while quoting by SCI inquiry, of approximately 7%.⁽¹⁰⁾ This was due to errors in documentation; e.g. several studies have shown that articles that have been retracted for some reason have continued to be cited.^(11,12) This is also supported by a recent article published in Science that showed that many studies that have been proven to be fraudulent are not even retracted.⁽⁹⁾

There are several errors that may occur when calculating the IF, because some languages may contain uncommon vowels and special letters, e.g., with journals published in the Chinese and Spanish literature.⁽⁶⁾ A major error may occur while collecting the raw data. The ISI does not have the power to correct the citation. In some situations, the author of the paper makes erroneous quotations, which is probably a frequent occurrence,⁽¹³⁾ resulting in the database becoming biased. Thus, the quality of a research paper and the achievements of a research group or an individual scientist cannot be assessed solely on the basis of a journal's IF, since approximately 15% of the articles published in journals are listed by the ISI but account for 85% of quotations from these journals.⁽¹⁴⁾

HOW TO CHECK ON THE VALIDITY OF THE IF

Although the IF is now well-accepted worldwide, there have been some problems associated with it that cause bias. Krell found that 98.5% of quoted papers in selected articles were older than two years.⁽¹⁵⁾ Also, the validity of the IF can be challenged on the basis of one more plausible reason. Interesting articles normally have more citations. However, in the case of controversial articles, they tend to be quoted more frequently, including retracted articles.⁽⁹⁾ Discrepancies have been found in the occurrence of quotations from different parts of the world; e.g. American papers have the highest number of quotations, approximately five times, compared to English papers (four times) and German papers (three times). Thus, it appears to be true that the more one gets published, the more one gets quoted.⁽¹⁶⁾ This fact may explain the well-known Matthew effect in citations,⁽¹⁷⁾ which has been confirmed by analyses in previous studies.⁽¹⁸⁾ It is common practice for scientists to publish their research contributions in a high IF journal as it helps in securing financial assistance (grants, etc) for future research, and also in gaining recognition as a result of a higher number of citations.

Presently, there are no gold standard criteria for evaluating IF in different countries or scientific fields, because particular specialties are compared independently of their sizes. However, the IF varies according to standardisation, so it is also called "the poor man's citation analysis".⁽¹⁹⁾ Despite the above-mentioned hidden secrets that highlight various drawbacks in the evaluation criteria of a journal's IF, it is used very frequently in the scientific world for judging the quality of a journal and of its scientific output.

USES AND MISUSES OF THE IF

At present, the IF is the main marker within the scientific community for evaluating the status of scientific journals as well as that of scientists, on the basis of their publication output, to assess how actively they are engaged in research. For example, the science ministries in South Korea, China and Pakistan now offer cash rewards to their scientists if they are able to publish papers in journals with high IFs such as Nature, Science and Cell. The remuneration amount can be quite impressive, as much as US\$ 50,000 in China. In Pakistan, scientists can receive between US\$ 1,000 and US\$ 20,000 on the basis of their annual cumulative IFs.⁽²⁰⁾ In many institutions, a professor's cumulative IF is the most important criterion for promotion. Often, publication in journals with a high IF, e.g. those covered by JCR, is used as the only evaluation criterion rather than the quantification of the scientific contribution itself. Because

the original idea of citation analysis was developed to protect against the uncritical citation of fraudulent and even disputed data, some have questioned the usefulness of the IF, stating that it actually represents popularity rather than prestige.⁽²¹⁾

Rey-Rocha et al have shown discrepancies in the IF between scientists or research groups in English and in non-English speaking countries.⁽²²⁾ In most non-English speaking countries, research is not published in high IF journals, and sometimes, it is a source of embarrassment for the scientists working in those countries whose journals are not even listed by SCI. For example, Spanish language research publications in domestic journals are not included in SCI.⁽²³⁾ The IF clearly does not reflect the importance of an individual article and thus, it is also unable to clearly define and assess the quality of an individual author's contribution in that article. Yet, the scientific community persists in using the IF to judge the quality of a researcher.⁽²⁴⁾

Eugene Garfield, the inventor of the IF, emphasised that its potential value lies primarily in the management of library journal collections – to help them determine their optimum make-up, providing a solid basis for a cost-benefit analysis of subscription budgets. In fact, the inventor of the IF never predicted that it would be used in the scientific community as a criterion for judging the quality of a scientist and determining the provision of research grants.⁽²⁴⁾ The IF is often misused, as there are no specifically-defined principles governing its interpretation. The IF is used to measure the importance of journals, as well as a researcher's potential, a use for which it was never intended, and it is also used to make faulty comparisons among journals. Thus, the misuse of the IF is a common problem in the field of research, and scholars have complained about this problem for a long time.⁽¹⁸⁾

PROBABLE FUTURE DIRECTIONS

The IF has been used for evaluating scientists and the quality of their work, despite its various drawbacks in evaluation, in the context of culture and the main languages used in various countries. This is producing unfairness in the evaluation of scientists due to its faulty usage on the basis of language, e.g. among non-Anglo-American countries. There should be a gold standard method to overcome the bias, in terms of non-English language usage and cultural differences, as well as to have a fair selection system and prevent any errors in the collection of raw data by the ISI.

Another important consideration that should be made is regarding the transparency of the IF calculation. When the calculation of the IF is less transparent, it can misguide scientists as well as the various agencies that

provide funds to carry out research. For example, a few recent papers quoted by the ISI are not included in the calculation of a journal's IF by the agency. As the ISI is a private company, the decision to consider a journal as an original one cannot be controlled scientifically. It must be further taken into account that only a very small group of journals is considered for the calculation of the IF and that the selection criteria are rather diffuse. The test procedure is not at all suitable for evaluating the achievements of individual persons and research groups. Thus, there is a need for careful usage of the IF in the future, including in correlating it with the scientific achievements of the scientists.

Some new methods have recently been developed that may help in updating or modifying the methods of evaluation of the IF in the future. For example, Hirsch has developed a new method called the *h*-index, which aims to evaluate the impact of individual scientists.⁽²⁵⁾ Bollen et al have introduced a new parameter which takes into consideration that if journal A is cited ten times more frequently in journal B than in any other journal, then it should transfer ten times more prestige to journal B.⁽²¹⁾ This study noted that the IF, which clearly measures popularity, is not without its value, because it is scientists' peers who are citing their papers. Hence, Bollen et al have invented a new parameter called the Y-factor, in which multiplication of the Page Rank factor is done using the IF. Using these weighting methods, the status of a journal for a particular year can be re-analysed and the rankings of the top journals re-evaluated.⁽²⁶⁾

CONCLUSIONS

Since its invention, the IF has become one of the most depended-upon criteria for the judgment of the quality of scientific production. This procedure has been much simplified, but instead of using only the IF as a criterion for judging the quality of a research group or scientist, other tools like the half-life of the scientific papers can also be used to evaluate the quality of the research paper.⁽²⁶⁾ However, this procedure is more demanding than other possible techniques, besides having systematic errors. There are already some initiatives under way for the development of a "new system which is based on online peer review pre-print publishing", e.g. the Public Library of Life Sciences (PLoS), as this is faster than traditional printed journals.⁽²⁷⁾ Another new step has been taken by scientific societies like The American Society for Pharmacology and Experimental Therapeutics (ASPET), which started the Abel number in order to provide recognition to scientists in the scientific community. Thus, there is a need to set a gold standard criterion for

evaluation, which begins and ends with knowledgeable readers who decide on the importance of a paper after reading it.

It is not the intent of this article to simply highlight the misuse of the IF in the judgment of scientific journals and scientists, but rather, to emphasise that the IF should be used for its original purpose as intended by its creator, Eugene Garfield. We hope that this article contributes to the discussion of the IF, raising questions and motivating the expression of different viewpoints with the intention of rationalising its use in the complex process of evaluating scientific production in biomedical research.

REFERENCES

- Garfield E. How can impact factors be improved? *BMJ* 1966; 313: 411-3.
- Rey-Rocha J, Martín-Sempere MJ, Martínez-Frías J, López-Vera F. Some misuses of journal impact factor in research evaluation. *Cortex* 2001; 37:595-7.
- Brähler E, Beutel M, Decker O. [Deep impact-evaluation in the sciences]. *Soz Präventivmed* 2004; 49:10-4. German.
- Brähler E, Decker O. Der Impaktfaktor: Erbsenzählerei oder ein valides Instrument zur Leistungsmessung? *J Dtsch Ges Plast Wiederherstellungschir* 2002; 14:9-12. German.
- Frömter E, Brähler E, Langenbeck U, Meenen NM, Usadel KH. [The AWMF model for the evaluation of published research papers in medicine. Arbeitsgemeinschaft der wissenschaftlichen medizinischen Fachgesellschaften (Working Group of the Scientific Medical Specialty Societies)]. *Dtsch Med Wochenschr* 1999; 124:910-5. German.
- Brähler E, Decker O, Bokenhagen A. Das Wahre, Schöne, Gute oder schöne, gute Ware? *Psyche* 2001; 54:1245-52. German.
- Garfield E. Uses and misuses of citation frequency. *Curr Contents* 1985; 43:3-9.
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. *J Biol Chem* 1951; 193:265-75.
- Smith R. Commentary: the power of the unrelenting impact factor--is it a force for good or harm? *Int J Epidemiol* 2006; 35:1129-30.
- Moed HF. The impact-factors debate: the ISI's uses and limits. *Nature* 2002; 415:731-2.
- Pfeifer MP, Snodgrass GL. The continued use of retracted, invalid scientific literature. *JAMA* 1990; 263:1420-3.
- Gabehart ME. An analysis of citations to retracted articles in the scientific literature. Available at: hdl.handle.net/1901/199. Accessed September 3, 2006.
- Adam D. The counting house. *Nature* 2002; 415:726-9.
- Seglen PO. Why the impact factor of journals should not be used for evaluating research? *BMJ* 1997; 314:497-502.
- Krell FT. Why impact factors don't work for taxonomy? *Nature* 2002; 415:957.
- EU eliminates citation gap with America. *Nature* 1997; 387:537.
- Merton RK. The Matthew effect in science. The reward and communication systems of science are considered. *Science* 1968; 159:56-63.
- Bonitz M, Schammorst A. [Science and economy]. In: Parthey H, Spur G, eds. [Science and Innovation]. *Wissenschaftsforschung Jahrbuch 2001*. Berlin: Gesellschaft für Wissenschaftsforschung, 2002: 85-96. German.
- Adams J. Benchmarking international research. *Nature* 1998; 396: 615-18.
- Al-Awqati Q. Impact factors and prestige. *Kidney Int* 2007; 71:183-5.
- Bollen J, Rodriguez MA, Van de Sompel H. Journal status. Available at: www.arxiv.org/abs/csGL10601030. Accessed June 12, 2006.
- Rey-Rocha J, Martin-Sempere MJ, Lopez-Vera F, Martinez-Frias J. English versus Spanish in science evaluation. *Nature* 1999; 397:14.
- Kimley S. Limitations of Science Citation Index data in evaluating journals and scientists in Geology. *Proceedings of the 28th Meeting of the Geosciences Information Society, October 25-28, 1993*. Boston: Geosciences Information Society, 1994: 23-31.
- John E. Measuring journals. *Notices* 2006; Version 14.0.
- Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 2005; 102:16569-72.
- Linardi PM, Coelho PM, Coster HM. The "impact factor" as a criterion for the quality of scientific production is a relative, not absolute measure. *Braz J Med Biol Res* 1996; 29; 555-61.
- Coelho PM, Antunes CM, Coster HM, et al. The use and misuse of the "impact factor" as a parameter for evaluation of scientific publication quality: a proposal to rationalize its application. *Braz J Med Biol Res* 2003; 36:1605-12.