

Self-citation and the journal impact factor

The goal of all investigators is to have the results of their studies published, so that scientific knowledge can be expanded. The research paper is expected to have valid findings, because it must undergo careful scrutiny through the peer-review process before it can be accepted for publication. However, it is well known that the evaluation of scientific quality is a difficult undertaking with variable results, depending on whether the reviewer selected for the task is a true expert in the field or one with a general interest in the area. Although a paper is frequently evaluated by two or more reviewers, it is not uncommon for editors to receive reports that vary widely in their comprehensiveness and attention to study methodology. Thus, what is called peer review in clinical research leaves a lot to be desired, as is witnessed by the growth in evidence-based medicine journals, which subject published papers to a second and more thorough review process using methodology filters. The distillate represents valid evidence of high quality that can guide decision making.

Although publication is an important and much-needed endpoint of the scientific inquiry process, the subsequent use and citation of these published papers by other investigators is equally important. The need for an objective judgement of research outcomes by assessing their value to other investigators was first suggested in 1955, when the concept of the impact factor was originated.¹ In the early 1960s, Irving H. Sher and Eugene Garfield created the journal impact factor as a method to help select journals for *Science Citation Index* (SCI).² The impact factor was proposed as a simple method for comparing journals regardless of their size, based on the belief that importance of a paper is determined by the number of times it is quoted in a given period of time.

The numerator for the calculation is the number of citations in the current year to any item published in a journal in the previous two years, and the denominator is the number of substantive articles published during this period of time.² This ratio of citations to publications approximates the frequency with which articles have been cited in the two years after publication. The data used to calculate the journal impact factor are based on information obtained from citation indexes, the most commonly used one being the SCI, which has been published by the Institute of Scientific Information since 1963. The source material used for this purpose includes all original articles, technical notes, and reviews, but does not include letters or abstracts. The citations are obtained from the reference lists of the published articles.

Despite the fact that it is difficult to define scientific excellence in quantitative terms, the impact factor has, through common usage, become established as the currency of scientific quality. Journal impact factors are being used increasingly by investigators as objective measures of the quality of publications and by administrators in academic institutions evaluating a candidate's application for promotion. This approach is based on the implicit premise that the impact factor of a journal is representative of its constituent articles and, hence, of the articles of the authors.^{3,4} It is not uncommon for investigators to make rash judgements about the quality of work of other investigators simply by perusing the names of the journals in which they publish their work. This is a naïve and potentially dangerous attitude, because it is well established that the most prestigious journals in different specialist areas may have very different impact factors.⁵ Some of this variability is related to the

focus of the journal (scientific vs clinical), its content (review vs original articles), its geographical origin (American vs European), and whether there is free electronic access.⁵

Although the impact factor is simple to calculate and straightforward, its use as a judgement of quality of research has provoked much debate.⁴⁻⁶ Both proponents and opponents acknowledge the limitations regarding the use of impact factors. One major problem is that citation analysis is susceptible to advertising that may increase the number of citations, but have nothing to do with the quality of the research. Impact factors can be manipulated by both editors and authors. Journal self-citation occurs when publications in a journal cite previous publications in the same journal. It is not uncommon to hear stories from colleagues who, having submitted their papers for publication, have received letters from editors, suggesting that references to specific articles in that journal be included in their paper. Critics of the impact factor as a measure of the importance of the journal have argued that journal self-citation is a significant problem, because it artificially inflates the journal impact factor, thereby producing a biased estimate of the value of the journal.

A study was performed to investigate the self-citation frequency of six anaesthesia journals and its possible effect on their impact factors.⁷ The journal self-citing rate was calculated as the proportion of self-citations of the journal to the journal's total number of references. The self-citing rates of the six journals in 1995 and 1996 were then correlated with the journals' impact factors for the year 1997, taken from the SCI of Journal Citation Report for that year. Although there was a wide variation in citations among the six journals, there was a significant correlation between self-citing rates and impact factors. This finding of a high self-citing rate having a positive effect on the journal's impact factor suggests that the current method for impact factor calculations should be modified, by either eliminating self-citations from the calculation or applying a correcting factor for self-citation to the calculation.

The second type of self-citation is author self-citation, which occurs when authors cite their previous publications in their new publication. There are many reasons for author self-citation. The concern about publication length, especially for paper-based journals, has forced authors to cite their previous work when describing their study methods. Authors focussed in a specific area of research often refer to their previous results to expand their hypotheses and provide justification for their subsequent research. In doing so, they may overestimate the importance of their original work, compared to the work of other investigators they could have cited, thus distorting the perception of its importance.⁸ The "publish-or-perish" climate that exists in academic institutions, where career advancement evaluations are based, in part, on quantity of publications and the number of citations of that author's work, fuels the need by investigators to increase the number of self-citations. However, for very busy and productive investigators, a high level of self-citation is inevitable. Consequently, it is difficult to determine where the balance should be struck between acceptable and questionable levels of self-citation.

The effect of author self-citation on the process of scientific inquiry is not well known and is difficult to ascertain, although its prevalence has been investigated. In the scientific literature, it has been observed that self-citations account for 10–20% of all

citations, depending on the field and the extent of development of the area.^{9,10} In contrast, the prevalence of self-citation in the clinical literature is relatively unknown. For this reason, a study was conducted to identify the extent to which this practice occurs in the literature on diabetes mellitus.¹¹ Among 170 clinical journals published in 2000, nearly one-fifth of all citations to articles about diabetes mellitus were author self-citations. This finding is in agreement with the evidence in the scientific literature and suggests that a similar, relative high, rate of author self-citation may be prevalent in other disciplines in medicine, including Obstetrics and Gynecology. The study also demonstrated that there was no association between the methodological rigour of the study and the likelihood of subsequent author self-citation.¹¹ These observations suggest that reasons other than methodological quality of the study are responsible for author self-citation.

It has been argued that self-citation is an important method for authors to promote a scholarly reputation and gain professional credit for their research.¹² Publication plays a vital role in our system of scholarly endeavour that both creates knowledge and distributes rewards. The need to have one's work recognized and cited by others has high value in the academic arena. The use of self-citations is a method to accentuate one's credibility or expertise. Informetric research has noted that authors cite their own work more than others.^{9,13} By situating themselves within the literature they cite, writers enhance their authority and promote the construction of a solid disciplinary identity that increases the likelihood that their work will be accepted.¹²

The effect of journal self-citation and author self-citation on the impact factor of any journal is a cause for concern. The problem is exacerbated when the worth of investigators is gauged by the journal in which their work is published, rather than the work itself. Also, when rewards and allocation of research funding are linked to publication in "top" journals (identified by high impact factor), it becomes more important to ensure that the bibliometric data used to make judgements about the scientific importance of published work are valid and reliable. It is clear that, because citations-based indicators are easily influenced by non-scientific factors, they should not be the sole indicators of the quality of research. Instead,

they should be considered as only one part of a larger peer-review process, in which scientific quality, scientific productivity, and scientific relevance are judged by a group of independent experts.⁶ Such a task requires effort and commitment if it is to be done well. Unfortunately, it is all too easy to rely on the impact factor, which has been in place for decades as a simple metric for assessing the scientific influence of a journal. The debate on this issue will continue until an effective solution can be found.

Salim Daya
Editor

Literature cited

- 1 Garfield E. Citation indexes to science: a new dimension in documentation through association of ideas. *Science* 1955;122:108–11.
- 2 Garfield E. Journal impact factor: a brief review. *CMAJ* 1999;161:979–80.
- 3 Metcalfe NB. Journal impact factors. *Nature* 1995;376:720.
- 4 Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ* 1997;314:498–502.
- 5 Neuberger J, Counsell C. Impact factors: uses and abuses. *Eur J Gastroenterol Hepatol* 2002;14:209–11.
- 6 Lankhorst GJ, Franchignoni F. The impact factor—an explanation and its application to rehabilitation journals. *Clin Rehabil* 2001;15:115–8.
- 7 Fassoulaki A, Paraskeva A, Pupilas K, Karabinis G. Self-citations in six anaesthesia journals and their significance in determining the impact factor. *Br J Anaesth* 2000;84:166–9.
- 8 Kovacic N, Misak A. Author self-citation in medical literature. *CMAJ* 2004;170:1929–30.
- 9 Tagliacozzo R. Self-citations in scientific literature. *J Documentation* 1977;33:265.
- 10 Rousseau R. Temporal differences in self-citation rates of scientific journals. *Scientometrics* 1999;44:521–31.
- 11 Gami AS, Montori VM, Wilczynski NL, Haynes B. Author self-citation in the diabetes literature. *CMAJ* 2004;170:1925–7.
- 12 Hyland K. Self-citation and self-reference: credibility and promotion in academic publication. *J Am Soc Inf Sci Technol* 2003;54:251–9.
- 13 White H. Authors as citers over time. *J Am Soc Inf Sci Technol* 2001;52:87–108.