

Opinion Paper

The Journal Impact Factor: don't expect its demise any time soon

Emmanuel J. Favaloro*

Department of Hematology, Institute of Clinical Pathology and Medical Research (ICPMR), Westmead Hospital, SWAHS, Westmead, NSW, Australia

Abstract

Much emphasis continues to be placed on the Journal Impact Factor (IF), a measure of journal article citation rates, and typically used as a surrogate marker of quality of both the article and journal. The IF is both revered and reviled, and is neither a perfect nor comprehensive measure, having several limitations and being subject to easy manipulation. The IF holds 'power' for journals because it can influence their future success. Furthermore, the perceived utility of the IF has grown way beyond that of its original and still popular use as a surrogate marker of publication 'quality'. The IF is increasingly being used (i) to objectively evaluate the scientific and academic value of scientists across a wide variety of disciplines, (ii) to short-list research projects for future financial support, (iii) to short-list or select applicants for academic promotion, and (iv) by researchers to measure the success of research institutes, research funding, or even entire countries. Accordingly, despite our love-hate relationship with the IF, don't expect its demise any time soon.

Clin Chem Lab Med 2009;47:1319–24.

Keywords: commentary; impact factor; journal quality.

Introduction

Much emphasis continues to be placed on the Journal Impact Factor (IF), a measure of journal article citation rates, and typically used as a surrogate marker of quality of both the article and journal (1). The IF is both revered and reviled, and only partially satisfies our collective urge to 'quantify' the value of our scientific work. The IF is a measure of the frequency with

which the 'average article' in a journal has been cited in a particular year. Specifically, the IF is a ratio between citations and recent citable items that have been published. The IF is calculated by dividing the number of current year citations to the source items published in that journal during the previous 2 years. Thus, the IF for 2008 for a journal is the relative number of times papers published in that journal during the preceding 2 years (i.e., 2006 and 2007) were cited in all Thomson listed publications during 2008 (Table 1). The IF has some value, but it is neither a perfect nor comprehensive measure. The IF is reasonably accessible, and most of us 'seem' to understand it. However, the IF has several limitations and is easily manipulated (2, 3).

Why and how do we use the Impact Factor?

The IF holds 'power' for journals because it can influence the future success of that journal, including subscriptions, submissions (quality and quantity) and advertising revenue. The IF is certainly a handy tool, as it provides a measure of citation rates, and thus gives us an indication of how much attention other researchers and authors are giving to our work. However, the perceived utility of the IF has grown way beyond that of its popular use as a surrogate marker of publication 'quality' and, by inference, a journal's 'quality' (2) as summarized in Table 2. Thus, the IF is also increasingly being used to objectively evaluate the scientific and academic value of scientists across a wide variety of disciplines (2, 4–9). However, it does not adequately reflect the individual's participation in the group's research, nor the relative ability of participants to conceive and design a study protocol, acquire, analyze and interpret data, or draft and critically revise manuscripts. The IF is also increasingly being used by educational and research institutions to measure the success of their research groups or programs, by grant-funding bodies to short-list research projects for future financial support, and by researchers to measure the success of research institutes, research funding, or even entire countries (2, 8–15). Having a large number of papers on your CV may no longer be good enough; these also need to be published in 'good' (i.e., high IF ranking) journals. The fate of academic tenure and promotion, and the fate of future research funding for both individual researchers and their research institutes can thus be decided by IF rankings.

*Corresponding author: Dr. E.J. Favaloro, Department of Hematology, Institute of Clinical Pathology and Medical Research (ICPMR), Westmead Hospital, SWAHS, Westmead, NSW, 2145, Australia
Phone: +(612) 9845 6618, Fax: +(612) 9689 2331,
E-mail: emmanuel.favaloro@swahs.health.nsw.gov.au
Received August 17, 2009; accepted August 26, 2009;
previously published online September 22, 2009

Table 1 Calculation of a journal's Impact Factor (IF) – year 2008 as example.

A = The number of times articles published in a journal in 2006 and 2007 were cited in all Thomson Scientific indexed journals during 2008.

B = The number of 'citable items' published in the journal in 2006 and 2007 (source items).

A/B = 2008 Impact Factor for that journal.

Table 2 Uses for the Impact Factor.

- As a marker of citation rates used to identify the 'quality' of published papers, and by inference, the 'quality' of the journal publishing those papers.
- Used by journal editors and their boards to review and monitor journal performance.
- Used by institutions to monitor or rank research activities by individual researchers or departments within their institutions or between institutions.
- Used by researchers to monitor or rank research compared to peers or across geographic boundaries.
- Used to rank the 'quality' of articles published in selected specialties.
- Used to rank the 'quality' of journals within specialties.
- Used to rank the 'impact' of different specialties.
- Used to 'rank' individuals as potential employees in certain specialties or institutions.
- Used for reshaping journal content or editorial approaches, or to reshape the approach undertaken by certain disciplines in order to be taken more seriously by other disciplines.

The Impact Factor has both strengths and weaknesses

There is little doubt that the IF is an important audit of journal article usage, as it essentially provides a

measure of the level of peer attention being given to articles within journals, and (by extrapolation) the level of attention being given to the journal containing those articles. However, the IF has many limitations, summarized in Table 3, and only tells a very small part of the overall story regarding the utility of a journal and the articles within.

The IF is important as a marker of utility regarding those that use the knowledge gained from reading these published articles and subsequently apply that knowledge to future work, for building additional knowledge, and for subsequent publications. The IF fails to identify the utility of publications for those that read the journal articles for the sake of knowledge itself, or who apply that knowledge to their daily clinical and laboratory practice – i.e., the end users of the publication. This is a significant but under-appreciated limitation of the IF in terms of measuring true journal or article utility. This limitation was made very apparent to me a long time ago when I evaluated IF data to compare publication output (numbers and perceive 'quality') from research institutes vs. general hospital facilities within my own geographic region (16). The main finding of this study was that research institutes tended to target high IF journals, but actually published a relatively small number of papers. However, hospitals tended to provide the vast bulk of medical publications, albeit with a much lower average IF.

At that time, I proposed the use of a cumulative IF score to help assess total comparative quantitative and qualitative publication output. The cumulative IF score was simply the total of all journal IF data as represented by journal papers published by the institutions. Interestingly, the cumulative scores became

Table 3 Some limitations of the Impact Factor.

1. Accessibility of journals and journal listing bias (only journals listed in the commercial Thomson Scientific database are included in the analysis; these listings strongly favor English language journals and primarily those published in the United States).
2. There is an inference, not always valid, that high citation rates reflect the 'high quality' of an article (probably true in most cases, but sometimes an article may have a high citation rate due to ongoing citation of some scientific flaw).
3. There is an inference, not always valid, that a high IF reflects generalized high citation rates or generalized 'high quality' for the journal (i.e., that a journal's IF reflects each article's citation pattern; this is not likely to be the case; also, (i) article citation rates determine the journal's IF, not the other way around, and (ii) there exists a wide range of citation patterns for different articles published within any given journal).
4. Citation bias (there is no correction factor applied against the influence of self-citation in IF calculations; and both authors and journals tend to favor self-citation).
5. Citations have equal ranking (there is no adjustment for where citations appear; thus, a citation in a 'low impact' journal is counted equally as a citation in a 'high impact' journal).
6. Specialty bias (IFs differ according to the research or specialty field); thus, generally it is less valid to compare IFs between journals from different specialty fields.
7. Date of publication issues (papers published at the beginning of the year contribute to IF calculations more than those published at the end of the year because they are in press for a longer period of time).
8. Some high quality articles, and high quality journals, are directed at end-users or medical and scientific practitioners, rather than researchers per se. Such end-user directed papers from these journals may have low citation rates and thus low IF scores (i.e., educational reviews or other papers aimed at medical and/or laboratory practitioners rather than scientific researchers will be cited less frequently in future publications and therefore will not contribute to the journal's IF – in fact they will likely act to reduce the journal's IF).
9. The IF only reflects a very small level of journal and article usage or utility. However, this is considered to be an important aspect of that usage.
10. The IF is calculated over a relatively short period of time (2 years), whereas the life of published papers is much longer than this (10 or more years). Indeed, some papers have very long citation lives, whereas others have very short citation lives. The IF favors the latter.
11. The IF is very easy to manipulate and artificially inflate.

quite comparable between hospitals and associated research institutes. Also, interesting, was the fact that the research institutes published almost exclusively in overseas (i.e., non-local or non-Australian) journals, whereas hospitals published predominantly in overseas journals, but about 30%–50% of their papers were published in local (Australian) journals. This meant that research institutes tended to target overseas journals with higher citation rates (i.e., promoting further research publications), whereas hospitals provided the largest bulk of medical publications and were also the sole providers of local end-user or practitioner educational material. I don't expect the situation to be different today, nor in other geographic localities.

Another important under-recognized limitation to the IF in terms of journal statistics is that it is based on data collected over a relatively short period of time (2 years), whereas the life of published papers is generally much longer. The calculation, by necessity, uses a short time frame. However, the problem is that some papers have very long citation lives, whereas others have very short citation lives due to the continuous advances within biomedical science (2). The IF system thus favors the latter – those with high and quick citation lives, but essentially ignores the former – those in which citations may be delayed ('publication sleepers'), but over time build sufficient interest to give them high composite citation scores. Accordingly, journals that publish papers with high and quick citation lives are favored by the IF system. The fact that some of these high IF papers will be forgotten after their short-term popularity is essentially ignored. Thus, a citation data collection system that shows the long-term citation patterns for articles would be of additional value, particularly if it was freely available. This is one reason that I occasionally utilize Google Scholar (17), although there are certainly limitations to the utility of data obtained from this site (to be discussed later).

Another major limitation of the IF is that it is very easy to manipulate and artificially inflate (2, 3). The simplest way to do this is to manipulate the numerator and/or the denominator in the IF calculation (see Table 1). For example, excluding 'Letters to the Editor' as source items in the denominator, but including the references cited in these published 'letters' in the numerator is one way to manipulate the IF. Another simple way to manipulate the IF is by self-citation. Authors and journals both may have a tendency for self-citation. In some cases, authors will genuinely self-cite because they feel that these citations best reflect what they wish to make reference to. However, it is inherently difficult to avoid a more general urge to self cite articles and thus 'artificially' raise the citation count for your own prior work. In some cases, self-citation can be valid and is useful to help raise attention for your work through increased exposure across a broader publication base. This may be appropriate for material that is likely to have a slow uptake. Thus, 'early' self-citation bias might be more evident for authors papers that have just been pub-

lished, and be more evident during the time period that data is collected for the IF. Journals may also have a tendency to self-cite articles for the same reasons that authors do.

Impact Factor 'wars' and Impact Factor 'games': uses and misuses of the Impact Factor

The importance of the IF to medical and scientific journals can be highlighted by the relative high usage of IF data in journal associated documentation. Most journals cite their IF on websites and within printed journal information. Most journal editorial staff review annual IF 'scores' and most journals, like researchers, actively pursue higher and higher IFs. In some disciplines, the rivalry is such that some accuse others of playing 'games' with the IF, and this can lead to a virtual state of competitive 'war' to gain that increasing share. Several articles of interest related to the IF for journals within my own fields of interest (laboratory medicine, thrombosis and hemostasis) can serve as examples (18–22).

Several researchers have called for changes in the way that the IF is calculated (2). For example, by (i) increasing the citation capture period to better reflect actual usage, given that the utility of most published articles is much longer than 2 years; (ii) ranking citations according to the source journal quality (currently all citations carry equal weight, irrespective of origin); and (iii) taking into account (or removing) self-citations (author self-citation and journal self-citation) from the IF calculation. The first suggestion could easily be accommodated with software. Although the second suggestion has some merit, on balance I would reject its application. This suggestion is 'elitist' and based on the flawed assumption that all published papers within a given journal are of high value just because that journal publishes some highly cited papers. However, I can see some merit with the last suggestion – a citation search for a given author could provide two values – the raw citation count plus the self-citation count (or percentage) listed alongside in brackets. Similarly, the citation or IF search for a given journal, or the IF value generated for that journal, could also be provided as two values – the raw IF value with the self-citation count or percentage also listed alongside in brackets. Other researchers have called for a total overhaul of the way in which the scientific literature is assessed, and to abandon the IF altogether. They suggest a variety of bibliometric data collection methods. I would agree that there is a need to move away from the exclusive dominance and potentially inappropriate use of the IF. It is important to reflect on the limitations of the IF, and remember that it only tells a very small (albeit important) part of the story (recitation rates) regarding journal utility. In particular, the IF tells us very little about any journal's major aspect of usage – the vast regular readership.

The alternate use of the IF to assess academic performance, determine academic promotion, and deter-

mine research funding opens a virtual Pandora's box of issues. This use is opposed by most researchers, except perhaps those benefiting from the process.

Challenges to the Impact Factor

There have been several challenges to the dominance of the IF over the years, as well as several calls for changes to the way that the IF is calculated or used. One recent failed challenge was the so-called (and now defunct) 'Prestige Factor' (23). Purportedly, this was similar to the IF, but focused only on 'original articles' on the premise that only such articles published 'innovative research'. Another 'challenge' that has recently emerged (24) is the 'Articles Rank'. This may be a less serious contender than the Prestige Factor, and is (according to the website) "related to the scientific articles published by researchers in peer-reviewed professional journals". The calculation "is based on quantity, quality of the scientific articles published by researchers in the peer-reviewed professional journals by certain authors" and "our know-how".

I would, however, predict that there are a couple of serious contenders to the dominance of the IF. The first of these is Google Scholar (17). Freely available, Google Scholar offers an open-platform comparison base to the IF (2). However, citation counts are expected to differ from the two citation source sites because: (i) Google Scholar includes citation sources not otherwise included in the Thomson Scientific database, and visa versa; (ii) Google Scholar is limited to web-based citation sources and includes many foreign language citations; (iii) due to the differences in (i) and (ii), Google Scholar tends to undercount citations from years that precede web inclusions (e.g., <1995). Also, compared to the Thomson Scientific database, Google Scholar may 'over-count' citations from more recent years. In addition, Google Scholar currently provides only basic search options and some of the citations may be duplicated or be of 'relatively low impact' (e.g., web-only based, non-peer-reviewed, non-journal). However, Google Scholar is simple to use, user friendly, and provides a useful open-platform comparison base. In addition, it has been noted with respect to the IF that the Thomson Scientific database is limited because it is restricted to (predominantly English language) journals listed on their database.

Another possible contender is SCImago (25, 26). Developed from information contained in the Scopus® (27) database managed by the publisher Elsevier (28), SCImago also is an open access platform (at least currently) that permits assessment of various metrics related to journals and countries, including citation and self-citation data.

Another way to reduce the focus on the IF is to utilize other markers of journal and/or article popularity, such as subscription patterns and download data (2, 21). Some journal articles will have more universal appeal, and will have a high citation rate and high

general interest as evidenced by the number of articles that are downloaded. However, some articles will be primarily of general interest (high number of downloads, but fewer citations), and visa versa. Such is the mix of articles published in a journal and the value of the journal to the scientific and medical community. Focusing on the IF alone undervalues the general readership of journals, and is detrimental to the long-term future of such journals.

Conclusions

The IF is an important indicator of citation rates, and therefore is useful as an indicator of article (and by inference journal) 'quality'. However, the IF has several limitations. It is easily manipulated, and only tells a small part of the story because it overlooks the major component of a journal's usage – the general readership. From my own modest analysis, the usage of a journal article reflected in the IF accounts for only ~0.1% of the journal article's usage (2). Thus, two analogies come to mind that describe the relationship between the usage of a journal that results in future citations which help increase the journal's IF, and the usage that describes the much larger general readership: (i) the former is like the tip of an iceberg, and the latter is like the iceberg's submerged portion. The tip is very important and is that which is most visible, but the tip is only a small part of the whole. As the crew and passengers of the Titanic discovered, those that ignore the submerged portion of an iceberg do so at great peril; (ii) perhaps a more palatable analogy is that the former is like the cream on the cake, and the latter is the cake. The cream might be tastier to some, but without the cake, it is only cream. After all, cream is best consumed with the cake.

Focusing on the IF alone undervalues the general readership of journals, and is detrimental to the long-term future of journals. The value of a journal to the scientific and medical community is a composite of its value to the research community (IF and citation patterns), as well as its value to other scientists and clinicians (the end-users of the knowledge gained and used to improve laboratory and clinical practice).

There are other dangers to over-use and misuse of the IF. The IF is now used extensively, and often inappropriately, to help assess or 'rank' the 'quality' of individual or group work, and to 'rank' specialties, geographical localities, institutes and researchers. I am also concerned by another trend with respect to the potentially inappropriate use of the IF which is restricting data collection for meta-analysis and reviews to that published in so-called 'high impact' journals (2). Although I can understand the intention of researchers restricting data collection to 'high quality' data by restricting their focus on 'high impact' journals only, the significant flaw in this approach is the presumption, totally invalid to me, that all the data in 'high impact' journals are 'high quality' and that all the data in 'low impact' journals are 'low quality'. This assumption is flawed since the opportunity to

publish articles in 'high impact' journals also depends on a variety of circumstances that sometimes have little to do with 'quality'. These circumstances include the scientific relevance of the argument, free-availability to the text of the articles, the judgment of the referees, familiarity with the Editorial Board, etc. Other flaws to limiting the analyses to high impact journals is evidence of serious biases (29). High impact journals are more likely to publish research studies and clinical trials showing clinically or statistically significant findings, and are less likely to publish research studies and clinical trials that do not show significant findings. As a result, these restrictive meta-analyses are more likely to report significant findings.

Will the perception of the IF be as important to us in the future as it is now? Perhaps it may, perhaps not. However, unlike Diamandis (30), I don't expect the demise of the IF any time soon, simply because it has become so pervasive. As I stated above and several times before (2, 20), the Internet has opened up some serious competition, both for traditional journals and for measures of citation. One other element to watch for is the spread of the open-access model, itself a strange beast and subject to misuse. Ideally, these forums should be subject to 'scientific regulation', but this may be akin to trying to control the Internet itself – a tall order indeed. Open-access journals, funded in part or in whole, by publication fees charged to authors, facilitate the ease of citation, and hence will challenge established subscription based journals. There are notable problems with both approaches. Someone, either the subscriber or author, must pay to permit publication. When the author pays the publisher, there will always be questions raised concerning the true quality of the work published, unless the process is appropriately scientifically monitored.

Accessibility to the written word has never been easier, and a good paper is a good paper, irrespective of whether it appears in a high IF journal, or a more moderate IF journal. Similarly, a weak paper is a weak paper irrespective of where it appears. There is a spread of 'quality' in every journal. On balance, the 'high impact' journals may have higher 'net quality', but not all articles within them will necessarily be of 'high quality'. Conversely, the 'lower impact' journals may have a lower 'net quality', but not all articles within them will necessarily be of 'lower quality'. All journals are available, and all articles are available to most. Thus, the IF might become less of an issue with time.

Despite several limitations, the IF is probably the best general marker of citation rate and publication quality that is currently available. The IF is readily (but not freely) available and is well recognized and generally accepted. However, it is probably not understood completely by most journal users (authors, readers and commercial advertisers alike). Alternate attempts at producing effective surrogate markers of journal article 'quality' have not yet emerged, although some contenders that are emerging may provide alternatives to the current marker in the future.

Finally, I am reminded of that famous line from George Orwell's *Animal Farm*, "all animals are equal, but some animals are more equal than others". The IF seems to have grown from something of interest to an obsession. Before we all begin to paraphrase the George Orwell line, "all journals are equal, but some journals are more equal than others", we need to remember that journal quality is not simply a measure of its IF.

A recent publication in *JAMA* (31) has reported on a direct comparison of citations using three of the databases (i.e., Web of Science, Scopus, and Google Scholar) identified within this report, for articles published in general medical journals. The authors of this *JAMA* publication (unsurprisingly) identified that these databases produced quantitatively and qualitatively different citation counts for published articles in the journals evaluated.

References

1. Garfield E. The history and meaning of the journal Impact Factor. *J Am Med Assoc* 2006;295:90–3.
2. Favaloro EJ. Measuring the quality of journals and journal articles: the impact factor tells but a portion of the story. *Semin Thromb Hemost* 2008;34:7–25.
3. Falagas ME, Alexiou VG. The top-ten in journal impact factor manipulation. *Arch Immunol Ther Exp (Warsz)* 2008;56:223–6.
4. Dempsey JA. Impact Factor and its role in academic promotion: a statement adopted by the International Respiratory Journal Editors Roundtable. *J Appl Physiol* 2009 Aug 13. [Epub ahead of print].
5. Adler KB. Impact factor and its role in academic promotion. *Am J Respir Cell Mol Biol* 2009;41:127.
6. Leff AR. Impact factor and its role in academic promotion. *Proc Am Thorac Soc* 2009;6:337.
7. Virchow JC, On behalf of the Respiratory Editors Group. Impact factor and its role in academic promotion. *Respir Med* 2009 Jul 9. [Epub ahead of print].
8. Wolf DM, Williamson PA. Impact factor and study design: the Academic Value of Published Research (AVaRes) score. *Ann R Coll Surg Engl* 2009;91:71–3.
9. Sebire NJ. H-index and impact factors: assessing the clinical impact of researchers and specialist journals. *Ultrasound Obstet Gynecol* 2008;32:843–5.
10. Haeffner-Cavaillon N, Graillot-Gak C. The use of bibliometric indicators to help peer-review assessment. *Arch Immunol Ther Exp (Warsz)* 2009;57:33–8.
11. Lyubarova R, Itagaki BK, Itagaki MW. The impact of National Institutes of Health funding on U.S. cardiovascular disease research. *PLoS One* 2009;4:e6425.
12. Tess BH, Furuie SS, Castro RC, Barreto Mdo C, Nobre MR. Assessing the scientific research productivity of a Brazilian healthcare institution: a case study at the Heart Institute of Sao Paulo, Brazil. *Clinics (Sao Paulo)* 2009; 64:571–6.
13. Cappell MS. A more than one-hundred-fold higher per capita rate of authorship of five democratic nations versus their relatively undemocratic neighboring nations among 6437 articles in 14 medical journals: does democracy and civil liberties promote intellectual creativity and medical research? *Dig Dis Sci* 2009;54:1609–20.
14. Pshetizky Y, Tandeter H, Tabenkin H, Vinker S, Lahad A, Karkabi K, et al. Thirty years of family medicine publications in Israel (1975–2004): what, where, and how much? *J Am Board Fam Med* 2009;22:57–61.

15. Pilc A. The use of citation indicators to identify and support high-quality research in Poland. *Arch Immunol Ther Exp (Warsz)* 2008;56:381–4.
16. Favaloro EJ. Medical research in New South Wales 1993–96 assessed by Medline publication capture. *Med J Aust* 1998;169:617–22.
17. Google Scholar (Beta): <<http://scholar.google.com/>>. Last accessed 16th August, 2009.
18. Bergemann D. The Impact Factor game: counting goals or measuring fair play? *Thromb Haemost* 2006;96:105–6.
19. Lippi G, Franchini M. The growing impact of publications in haemostasis and thrombosis journals. *Thromb Haemost* 2007;97:1049–50.
20. Favaloro EJ. More on the impact factor and thrombosis and haemostasis journals: benefits and limitations. *Thromb Haemost* 2007;98:475–6.
21. Lee C, Kessler C, Barback J. Haemophilia: January 2008. *Haemophilia* 2008;14:1–4.
22. Lippi G, Favaloro EJ, Guidi GC. The impact factor and journals in laboratory medicine. *Clin Lab* 2009;55:49–52.
23. The Prestige (Factor) is gone. Information today news report. <<http://www.infoday.com/it/may02/hane1.htm>>. Last accessed 16th August, 2009.
24. BioMedTalk. <<http://www.biomedtalk.com/showthread.php?t=1432>>. Last accessed 16th August, 2009.
25. Falagas ME, Kouranos VD, Arencibia-Jorge R, Karageorgopoulos DE. Comparison of SCImago journal rank indicator with journal impact factor. *FASEB J* 2008;22:2623–8.
26. SCImago. Journal and country rank. <<http://www.scimagojr.com/>>. Last accessed 16th August, 2009.
27. Scopus homepage: <<http://www.scopus.com/home.url>>. Last accessed 16th August, 2009.
28. Elsevier homepage: <http://www.elsevier.com/wps/find/homepage.cws_home>. Last accessed 16th August, 2009.
29. Etter JF, Stapleton J. Citations to trials of nicotine replacement therapy were biased toward positive results and high-impact-factor journals. *J Clin Epidemiol* 2009; 62:831–7.
30. Diamandis EP. Journal Impact Factor: it will go away soon. *Clin Chem Lab Med* 2009;47.
31. Kulkarni AV, Aziz B, Shams I, Busse JW. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *J Am Med Assoc* 2009;302:1092–6.