Viewpoint

The Journal Impact Factor: Too Much of an Impact?

Tam Cam Ha,[1] *PhD*, Say Beng Tan,[1,2] *PhD*, Khee Chee Soo,[3,4] *MD, FRACS, FAMS*

Abstract

Introduction: The journal impact factor is often used to judge the scientific quality of individual research articles and individual journals. Despite numerous reviews in the literature criticising such use, in some countries the impact factor has become an outcome measure for grant applications, job applications, promotions and bonuses. The aim of this review is to highlight the major issues involved with using the journal impact factor as a measure of research quality. Methods: A literature review of articles on journal impact factors, science citation index, and bibliometric methods was undertaken to identify relevant articles. Results: The journal impact factor is a quantitative measure based on the ratio between yearly citations in a particular journal to total citations in that journal in the previous 2 years. Its use as a criterion for measuring the quality of research is biased. The major sources of bias include database problems from the Institute for Scientific Information and research field effects. The journal impact factor, originally designed for purposes other than the individual evaluation of research quality, is a useful tool provided its interpretation is not extrapolated beyond its limits of validity. Conclusion: Research quality cannot be measured solely using the journal impact factor. The journal impact factor should be used with caution, and should not be the dominant or only factor determining research quality.

Ann Acad Med Singapore 2006;35:911-6

Key words: Bibliometric methods, Journal impact factors, Science citation index

Introduction

The publication of research studies in scientific journals is the mechanism by which the latest discoveries, interesting information, and new knowledge are formally disseminated to the scientific community. The identification and evaluation of research studies of high scientific merit is an important but difficult task. Therefore, quantitative measurements of journal article quality, such as the journal impact factor (JIF), have become increasingly popular as a surrogate measure of scientific quality.

For a particular journal, the JIF is defined as the number of citations within a given year (e.g., 2005) cited to all papers published in that journal during the previous 2 years (i.e., 2004 and 2005), divided by the total number of papers published in that journal during those 2 years. The ratio has been used to judge the quality of individual research articles, as well as the quality of individual journals. In some countries, the JIF has been used as a criterion for the assessment of research funding, in the appraisal of research staff performance, and in considering job promotions and salary bonuses. However, one single factor cannot measure the scientific credibility of journal articles, journal quality, individuals, specific research projects or research institutions.

Indeed, for this and other reasons, there have been a number of major reviews in the literature criticising the use of JIFs as a measure of journal article quality and journal quality.[1-5] Nevertheless, the JIF continues to be used as a surrogate measure of scientific quality in many countries.[6-9] In this review, we summarise the main concerns raised in the literature regarding the use of JIFs as the primary measure of research quality. We argue that this penalises high-quality researchers working in low-impact factor fields, and potentially results in poor research quality.

The major problems associated with citation analysis and ultimately, the use of the JIFs, as raised by various papers

[1] Clinical Trials and Epidemiological Sciences, National Cancer Centre, Singapore
[2] Clinical Trials and Epidemiology Research Unit, Singapore
[3] National Cancer Centre, Singapore
[4] Department of General Surgery, Singapore General Hospital, Singapore
Address for Reprints: Dr Tam Cam Ha, National Cancer Centre, Clinical Trials and Epidemiological Sciences (CTE), 11 Hospital Drive, Singapore 169610.
Email: ha.tam.cam@nccs.com.sg

in the literature, are summarised in Table 1.

Institute for Scientific Information Database Problems

The Institute for Scientific Information (ISI) in Philadelphia, USA, has created a database that continuously encodes all references found in the reference lists of articles from 13,673 journals within the medical and natural sciences (as of 2005).[10] However, these journals are a small fraction of the 126,000 scientific journals in the world (as of 1996).[11]

Moreover, the ISI's database seeks to give sufficient representation to all specialties[10] and, consequently, the journals selected do not necessarily comprise those most often cited. The ISI database has a preference for the English language,[2,12] and is dominated by North American publications.[1,2,11] This language bias is further compounded by the tendency of authors to selectively cite articles in their own national language. It was estimated in 1995 that half of all citations are to USA scientists, who are also prone to cite each other, thereby raising the citation rates of USA scientists to 30% above the world average.[13]

Citation indices gather bibliographic citations only from journal articles, and not from books, book chapters, or conference proceedings. These "non-article" publications are included as cited references but not as citing source items. Within this body of source journals, further selection is made, with only some types of journal contributions included as source items; original articles and review articles are included but not letters. Citations to editorials or letters may be included without these publications being counted as source items, and hence potential citations are considered "for free". As the ISI's database does not correct for self-citations, this leaves the potential for editors, perhaps unintentionally, to artificially inflate the impact factor of their journals by frequently referring to their editorials.

Journals are severely punished for publishing many supplements from meetings, as many of them are included in the denominator of the JIF equation, but not the numerator. The JIF is thus dramatically reduced, despite the educational value of these supplements. However, indiscriminately including all types of publications as source items would unfairly lower the JIF, as the great majority of these items were probably never meant to be cited.

Distinguishing JIF from Article Citation Rate

Even within the core citing journals, 10% of journals account for 90% of all citations.[14] Similarly, a relatively small number of articles within a single journal receive the majority of all citations. The distribution of JIFs of articles within a journal is skewed, and it is wrong to assume that all articles in the journal are of similar quality.[15] In this light, it is not valid to assume that giving articles the average citation value of the publishing journal reflects their actual citation rates. There is no correlation between the JIF and the actual citation rate of the individual article.[16]

Table 1. Major Problems Associated With Citation Analysis and Use of JIFs

Technical ISI* database problems	Research field effects	Reference selection and citer motivation	Problems associated with using the journal impact factor
• Biased towards the English language • Biased sample of journals included in the database • Database coverage different between research fields • Books, conference proceedings, letters not included as source items • Delayed registration of citations • Frequent misprints (up to 25%) • Synonymy (several variants of the same article) • Homonymy (several authors with the same name) • Publishing time penalises disciplines with longer turnover times	• Field size • Field dynamics (expansion or contraction) • Research theme • Inter-field relations (e.g., clinical medicine draws heavily on basic science, but not vice versa) • Research fields with literature that rapidly becomes obsolete are favoured	• Primary criterion for reference selection is not quality but utility in research • Incomplete referencing due to journal space limitations • Reference copying • Flattery (citation of editors, potential referees) • Self-citation • In-house citation (friends and close colleagues) • Review articles heavily cited • Utility in research rather than pure scientific quality is the primary criterion for reference selection	• JIFs are determined by technicalities unrelated to the scientific quality of their articles • JIFs are not statistically representative of individual journal articles • Distribution of citations to articles within same journal not uniform • JIFs correlate poorly with actual citation rates of individual articles • No mechanism to correct for self-citations • Selective journal self-citation: articles tend to preferentially cite other articles in the same journal • JIFs are a function of the number of references per article in research field • Short publication times result in high JIFs • National bias in reference selection favours American journals • Review articles are highly cited, resulting in higher JIFs

ISI: Institute for Scientific Information; JIF: journal impact factor

Research Field Effects

The effects of research fields are complex. A major problem when using JIFs for scientific evaluation is that the factor does not allow for comparisons between different research fields. Citation rates and JIFs may be influenced by the choice of field, the field dynamics, and field size. The choice of research theme will determine, *a priori*, the probability of becoming highly cited. Scientists working in rapidly expanding fields, such as acquired immunodeficiency syndrome (AIDS), are likely to have a higher citation rate compared with those working on childhood osteoperosis. The reason is that AIDS research is a relatively new area and there will be many citers relative to the citable material.

In large research fields, the mean citation rate should be independent of field size. However, the range of citations will likely be wider in a large field, thereby providing better prospects for a few authors to become highly cited.

Recommendations

Proposed Benchmarks of Scientific Merit

Despite the number of biases that may distort the JIF, a number of potential alternatives have been identified in the literature (Table 2); however, no consensus on a workable alternative to the JIF appears to have been found. Appropriate benchmarks to compare journal quality, research quality and the scientific merit of individuals and institutions, are multifactorial.

No single summary measure of scientific quality can be used to assess the credibility of individual journal articles or journal quality. Figures 1 and 2 summarise other factors that should be considered when assessing scientific quality, such as study design, the research question investigated, appropriate statistical methodologies, generalisability to other populations, and any wider applications in the scientific community. The fact that an article has been included in a Cochrane review should also be considered a measure of research quality as, ultimately, results from the Cochrane reviews may determine changes in clinical outcomes.

JIFs should not be used as the only, or the dominant, criterion when evaluating journal article quality, individual scientists or research units. Even when only considering publications, it is worthwhile to examine those of an investigator's journal articles or publication types which are not included in the citation indices. Comparing a researcher's total output to their first-author publications may also be another possibility.

Researchers should concentrate their efforts on high scientific merit. In certain settings, local or national impact may be more important than international impact. Citation

Table 2. Some Proposed Alternatives to the JIF Identified from the Literature

Problem	Proposed alternatives	Comments
Field effects	Divide article citation rates by JIFs	May introduce new bias e.g., by punishing authors publishing in highly cited journals
	Construction of individualised field factors	Not feasible and not very useful[1]
Database biases	A separate database for different languages or specialty areas	• Costly in terms of database and human resources • Cannot rely on a subset of journals as many researchers work in narrow fields and often publish their better papers in general journals
Publishing time bias	Change the citing window from 2 to, for example, 10 years	Does not totally correct for varying publishing times of different journals
Different disciplines	Weighted JIF: multiply the JIF by a coefficient that should neutralise the general differences between disciplines. These can be calculated by comparing the journal's JIF with the top JIF of its discipline	• External comparisons and internal rankings of journals by discipline may give a distorted picture, as only the true specialty journals are taken into account • Not clear how general science journals, which include important articles from any specialty, or other specialties' journals should be dealt with in this respect • Actual allocation by Institute for Scientific Information of specific journals to specialty rankings may not always agree with what specialists themselves consider their most important journals
	Research institutions may consider their own specialty journal groupings, which would correspond better to their fields of interest	• Tailor-made solutions are highly susceptible to arbitrary manipulation and might result in an unworkable situation • Less clear how internal popularity variations of subspecialties within disciplines can be neutralised
Relative impact factor	Compares actual citation counts of a paper with the mean citation count of the journal in which it appeared	This could lessen the social biases between journals and specialties, as it might divert interesting papers to less appreciated periodicals

JIF: journal impact factor

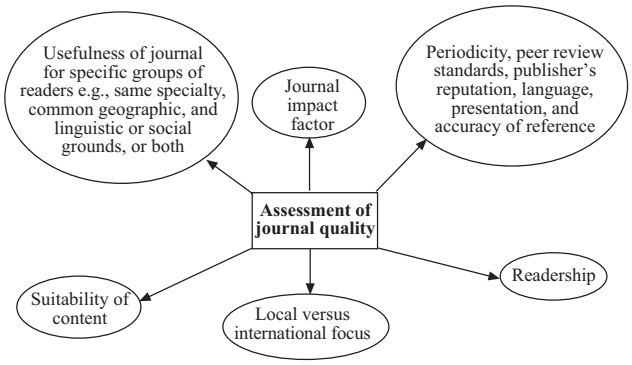

Fig.1. Factors potentially determining journal quality.

analysis is not a short cut and should not be a replacement for thinking. Instead, it is a point of departure for those who are willing to explore every avenue to a thorough evaluation.[17]

Evaluating Scientific Merit Across Specialties

The allocation by the ISI of specific journals to specialty rankings may not always concur with the view of specialists working in the area. For instance, the *International Journal of Leprosy* ranks among the top journals in the tropical medicine category, but is not included in the infectious diseases category. Similarly, *Parasitology Today* and other medical parasitology journals that deal mainly with tropical diseases are not included in the tropical medicine category.[3] If these journals were placed in the suggested categories, their JIF would markedly increase.

Individual research institutions may develop their own specialty journal groupings that may correspond better with their fields of interest; however, such solutions can incline towards their own arbitrary manipulation, and may result in an unworkable solution. It is also less clear how the internal popularity variations of subspecialties can be neutralised. Thus, the citation rates of scientists working in different areas cannot be compared; and this stricture also applies within the same field, but across different subspecialties.

Allocating Research Grants According to Scientific Merit

It is not helpful, for example, for research grant-giving bodies to require the specification of a target number of publications above an arbitrary JIF. It may be more appropriate for grant applicants to specify the target type of journal in which they expect to publish, and the number of such articles: for example, 5 articles in a journal of similar standing to the *British Journal of Cancer* or better. External expert reviewers can then comment on whether these target journals are of sufficiently high scientific quality.

Mis-interpretation and Over-interpretation of Quality Indicators

The ISI is a commercial company whose primary purpose is to provide researchers with access to current research information of high quality. The JIF was derived as a measure for the comparison of individual journals. The use of citation analysis and JIFs is widespread and has become a surrogate measure of research quality. While journals can be compared, over-interpretation may lead to inappropriate conclusions being drawn.

The JIF is also used to gauge the relative importance of individual researchers, research programmes, and even of the institution hosting the research. However, the JIF is just a time-specific citation rate index and nothing more. What is called the JIF should not be misused to evaluate journals or validate the scientific value of a particular researcher or research programme, particularly in making decisions on hiring, research funding and tenure.

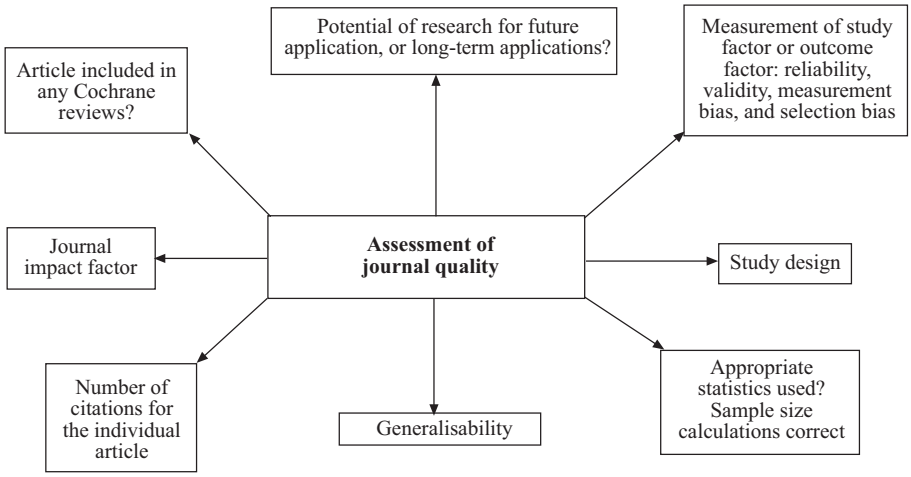

Fig.2. Factors potentially determining research quality.

JIFs have an increasingly influential role, as authors and institutions are often judged and funded based solely on the number of publications in "high-impact" journals. Yet, as a quality indicator of individual and institutions, the JIF is often criticised and is fraught with bias. It is a concern that the editorial decisions of some journals are based not on scientific merit but on financial profit. Quality indicators are never fully correct quantifiers of merit of small research groups or individuals. The more specific the JIF-based assessments or comparisons are, the more they have been challenged. Individual articles' citation rates determine the JIF, and not the converse. The creator of the JIF has stated that it is incorrect to judge an article by the impact factor of the journal.[18]

Effects of JIF on Authors' Behaviour

The results of certain research projects may be more appropriately reported in a local journal with a readership more relevant for the article. However, placing emphasis on journals with high impact factors may induce authors to submit their papers to journals that may not be the most appropriate forum for their work. Many authors believe that publication in a prestigious journal will increase the citations that a paper receives, compared with the same paper in a less prestigious journal. In fact, there is no correlation between the JIF and the frequency with which an article is cited.[2]

Since funding bodies use the JIF to determine the allocation of financial resources to individuals and institutions, it follows that our own scientists would send their best work to journals with high JIFs. This will systematically strengthen journals with high impact factors and remove support from other journals with a second- or third-tier status.

However, despite these valid concerns, JIFs are still widely used in many countries as the primary criterion in assessing research quality. They offer a simple tool for the comparison of research output, but in the end, what is really important? Is it research quantity, or research quality, or patient outcomes?

Journal Citation Rate, Citation Half-life, and Immediacy Index

Citation rates are determined by so many technical factors that pure scientific quality may be a very minor influence. It is tempting to place too much emphasis on this seemingly objective measure of quality. Given the technical biases, vulnerability, distortion, and manipulation of these statistics, citation rates are easily misinterpreted and should be regarded with caution. Citation statistics for articles and journals should never take precedence over the thoughtful analysis of the quality of research, both when reading journal articles, and when deciding where to submit them.

Other measures of a journal's worth include the Index Copernicus, citation half-life and immediacy index. The journal cited half-life is the median age of the articles that were cited in the journal citation reports. A journal with a cited half-life of, say, 7.0 years means that the interval 2000-2006 (inclusive) accounts for 50% of all citations to articles from that journal in 2006. A higher or lower cited half-life does not imply any particular value for a journal, as a primary research journal might have a longer cited half-life than a journal that provides rapid communication of current information. Dramatic changes in cited half-life over time may indicate a change in a journal's format.

The immediacy index is the average number of times an article is cited in the year it is published and indicates how quickly the articles in a journal are cited. The immediacy index is calculated by dividing the number of citations to articles published in a given year by the number of articles published in that year. Because it is a per-article average, the immediacy index tends to discount the advantage of large journals over small ones. Journals that are published more frequently may have an advantage because articles that are published early in the year have a better chance of being cited than articles published later in the year. However, the deficiencies of using parametric analysis to measure scientific quality or journal quality remain.

Index Copernicus is a ranking system, set up by members of the medical community from the Central European Region. This ranking system evaluates journal quality by using 5 groups of standards; scientific quality, editorial quality, technical quality, international availability and frequency-market stability. The overall score derived is considered a measure of journal quality. However, the Index Copernicus system is not widely used, as journals must request to be scored.

Recognition of Limitations of JIFs

JIFs are a relatively simple and cost-effective alternative to true citation analysis. However, the JIF is clearly not the holy grail of quality assessment that some science administrators or highly cited authors may believe it to be. The ISI has been aware of most of these shortcomings from the very beginning and has warned against the use of their tools for individual judgments. These concerns suggest that no counting result or ranking can be foolproof, as many individual factors are in force.

Conclusion

As with other measures of multifaceted phenomena, the transition from qualitative to quantitative measures can produce the drawing of inappropriate conclusions. Users of JIFs need to understand the strengths and weaknesses of

JIFs, and should not over-interpret data from their analysis. It is when data are misused that mistakes occur. In conclusion, "*it is remarkable that scientists may rely upon such a non-scientific method for the evaluation of the scientific quality of a paper as the impact factor of the journal in which is it published*" (Steven Lock, Emeritus Editor of the *British Medical Journal*).[19] As with all measures of quality, any interpretation of the JIF should be guided by a sound knowledge of its limitations.

REFERENCES

1. Seglen PO. Citations and journal impact factors: questionable indicators of research quality. Allergy 1997;52:1050-6.
2. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997;314:498-502.
3. Schoonbaert D, Roelants G. Citation analysis for measuring the value of scientific publications: quality assessment tool or comedy of errors? Trop Med Int Health 1996;1:739-52.
4. Seglen PO. Citation rates and journal impact factors are not suitable for evaluation of research. Acta Orthop Scand 1998;69:224-9.
5. Kurmis AP. Understanding the limitations of the journal impact factor. Bone Joint Surgery Am 2003;85-A:2449-54.
6. Callaham M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. JAMA 2002;287:2847-50.
7. Petroianu A. Quantitative parameters to evaluate the publication of scientific papers (Portuguese). Rev Assoc Med Bras 2003;49:173-6.
8. Kovacic N. Structure of the 2003 impact factor for Croatian medical journal. Croat Med 2004;45:671-3.
9. Leff D. Making an impact: the rise of the impact factor as a measure of journal quality. J Am Diet Assoc 2005;105:29-30.
10. Andersen H. ACTA Sociology in the International Arena – what can the Social Science Citation Index tell us (Danish)? Danish Sociology 1996;2:72-8.
11. Institute for Scientific Information. Available at: http://www.isinet.com. Accessed 15 August 2005.
12. Fava GA, Guidi J, Sonino N. How citation analysis can monitor the progress of research in clinical medicine. Psychother Psychosom 2004;73:331-3.
13. Braun T, Glanzel W, Grupp H. The scientometric weight of 50 nations in 27 science areas, 1989-1993. Part II. Life sciences. Scientometrics 1995;34:207-37.
14. Hamilton DP. Publishing by – and for? – the numbers. Science 1990;250:1331-2.
15. Seglen PO. The skewness of science. J Am Soc Inform Sci 1992;43:628-38.
16. Seglen PO. Causal relationship between article citedness and journal impact. J Am Soc Inform Sci 1994;45:1-11.
17. Garfield E. How to use citation analysis for faculty evaluations, and when is it relevant? Part 2. Current Contents 1983;45:363-72.
18. Garfield E. Journal impact factor: a brief review. CMAJ 1999;161:979-80.
19. Zetterstrom R. Impact factor and the future of Acta Paediatrica and other European medical journals. Acta Paediatr 1999;88:793-6.