A Method for Comparison of Biomedical Publication Quality Across ISI Discipline Categories

Peter Cleaton-Jones, B.D.S., M.B.B.Ch., D.Sc. (Dent.); Glenda Myers, B.A. (H.Dip.Lib.), B.Bibl.Hons, M.Bibl.

Abstract: The purpose of this paper is to offer a method to help to objectively compare quality of publication in biomedical journals in different disciplines with varying ISI Impact Factors (IF). Three methods—the number of journals per ISI Journal Citation Report discipline category/10, the IF/10, and the log (IF+1)/10—were used to calculate an article score. The distribution of article scores were compared across three defined ISI discipline categories: two clinical categories, dentistry (ISI category—dentistry, oral surgery, and medicine, forty-five journals) and medicine (ISI category—medicine, internal and general, 110 journals), and one basic science category, physiology (ISI category—physiology, seventy-four journals). The use of article scores per discipline category enables a reasonable, relative comparison of the quality of biomedical publications of individuals across disciplines for the purpose of promotion or awarding of research grants.

Dr. Cleaton-Jones is Professor and Director, Dental Research Institute of the Medical Research Council, and Ms. Myers is Health Science Librarian and Head, Witwatersrand Health Sciences Library, both at the University of the Witwatersrand, Johannesburg, South Africa. Direct correspondence and requests for reprints to Dr. P. Cleaton-Jones, Dental Research Institute, University of the Witwatersrand, WITS, Johannesburg, 2050 South Africa; +27-11-717-2229 phone; +27-11-717-2121 fax; cleatonjonesp@dentistry.wits.ac.za.

Key words: dental research, journal impact factor, publication quality

Submitted for publication 1/14/02; accepted 4/25/02

The promotion of individuals within academic and research organizations, as well as the awarding of research grants to investigators, is a competitive process. Evaluation of a person's research output is an integral part of the assessment, and even though there is ongoing debate about it, this process should be as fair and objective as possible. Research publications are an important part of the assessment although other research activities have been used as well.[1,2]

Assessment of research publications, particularly their quality, is a complex problem, which, if one is honest, has not yet been solved. There are those who believe that the peer review with careful human judgment is the method of choice.[3,4] Others prefer one or other form of bibliometric analysis ranging from publication numbers[5] through weighted productivity scores[6] to citation analyses,[7] including that pioneered by Garfield.[8]

Generally, in the assessment of an individual's scholarly contributions to the literature, publication counts are almost always used, often combined with subjective impressions of manuscript quality. This may be considered unfair by those being assessed, particularly when publications in different disciplines must be compared.

Two methods of publication quality that are commonly used are publication citations[7] and the Institute for Scientific Information (ISI) journal Impact Factor (IF). While each of these has value, neither is acceptable as absolute indicators of publication quality.[7,9,10] The most common indicator, the ISI IF, is calculated for a particular year for a specific journal according to the formula IF = C/A where C = the total cites in a particular year to articles published in the preceding two years and A = the number of articles published in the preceding two years.

For example, to calculate the IF of the *Journal of Dental Research* in 1999:

$$IF = \frac{\text{1057 cites in 1999 to articles published in 1998, 1997}}{\text{232 articles published in 1998, 1997}} = 4.556$$

The calculation of journal IF is influenced by a couple of factors. Articles in review journals tend to be cited more than those in original research journals because the review articles are often used as reference sources to earlier literature. In addition, articles in journals serving rapidly changing disciplines such as immunology or HIV infection have higher IFs while journals serving smaller research disciplines, or disciplines in which change is slower, have lower IFs.

The University of the Witwatersrand, Johannesburg, South Africa consists of five multidisciplinary faculties with a total enrollment of approximately 20,000 students and 3,000 academic staff. Each faculty decides on promotion and research funding within itself, a process that requires comparison of publications across disparate disciplines. In the Faculty of Health Sciences the broad disciplines are basic sciences (anatomy and physiology), dentistry, medicine, nursing, occupational therapy, pharmacy, and physiotherapy. The faculty promotion and research committees usually consider the IF of publications when making decisions. Because of influences on the calculation of the IF mentioned earlier, individuals working in certain disciplines may be disadvantaged, while those in others may be favored. For example, the highest journal IF in the smaller discipline of dentistry is the *Journal of Dental Research* (IF 4.556), while in the larger discipline of clinical medicine it is the *New England Journal of Medicine* (IF 28.857).[11] However, within their respective disciplines, articles in both journals are highly respected, and both have stringent standards for manuscript acceptance.

There is a need for a method that is simple and reasonably objective for relative, not absolute, comparison of article quality across biomedical disciplines. Ugolini et al.[1,2] realized this need and based their assessment on the IF and ISI Journal Citation Reports (JCR) discipline categories.[11] They placed journal IFs ranked in descending order into ten equal groupings called deciles (Dc) to produce an article score using the formula Dc = Nc/10 where Nc is the number of journals in an ISI discipline category. They rounded the results to the nearest integer and assigned an article score from 10 to 1 in descending order of IF. For example, if there are forty journals in a discipline, the interval of each decile will be 40/10 = 4. The four journals with the highest IF would be given an article score of 10, the next four a score of 9, and so on down to the baseline score of 1 for the final four journals. Each researcher's article may be multiplied by the article score to produce a quality weighting if so desired.

The article scoring method of Ugolini et al.[1,2] is simple and easy to use, but treats all journals equally and does not consider an individual journal's IF. Distribution of IF in a discipline category is exponential, with fewer journals with a high IF than lower IF, which is illustrated in Figure 1 for three ISI

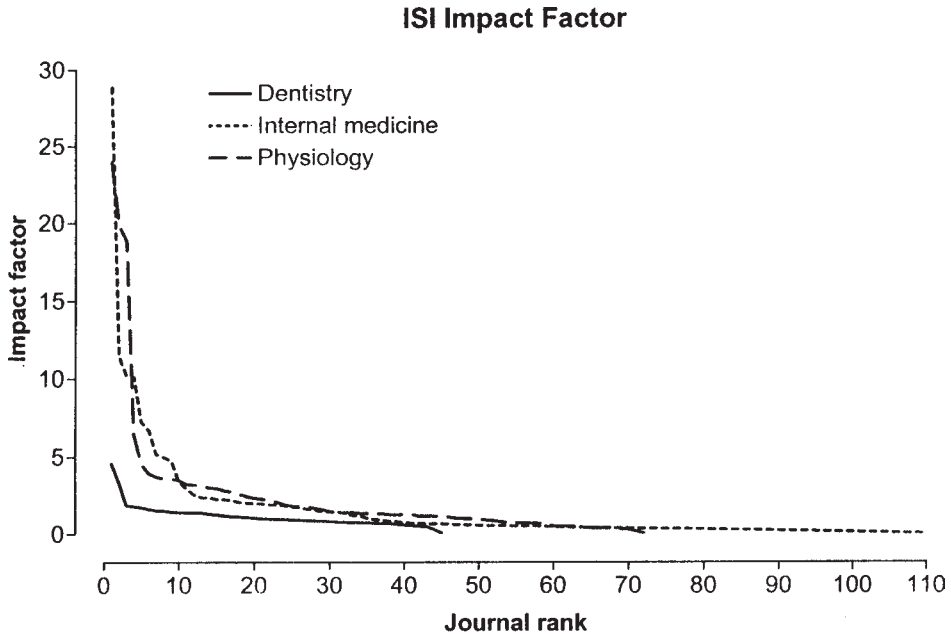

Figure 1. Frequency distribution of journal Impact Factors in three discipline categories

discipline categories. For example, the top four journals in a discipline may have an IF of 20, 12, 1, and 0.4, clearly indicating that they differ in perceived quality through citation frequencies; yet all would receive an article score of 10 with the Ugolini et al.[1,2] article score method.

Our study reports the results of an investigation to see if modification of the Ugolini et al.[1,2] article score method would discriminate "better" between the "quality" of journal articles for relative, not absolute, comparison of research outputs between individuals publishing across ISI discipline categories.

Methods

The 1999 science edition of ISI Journal Citation Reports[11] was the information source; this is the most recent edition at the University of the Witwatersrand. Three biomedical discipline categories were empirically selected to cover two clinical disciplines—dentistry (ISI category for dentistry, oral surgery, and medicine, forty-five journals) and medicine (ISI category for medicine, internal and general, 110 journals)—plus one basic science discipline, physiology (ISI category for physiology, seventy-four journals). Within each category, the journals were sorted in descending IF order and article scores assigned according to three methods of score calculations:

1. *Jn/10.* The number of journals per discipline was divided by 10 and the result used to determine article scores from 10 to 1 according to the method of Ugolini et al.[1,2] For the forty-five journals in the category of dentistry, the four journals with the highest IF would score 10, the next four would score 9, the following four would score 8, and so on down to the last journals, which would score 1.
2. *IF/10.* The highest IF in a category was divided by 10, from which result 10 class ranges were calculated. For example, for dentistry the highest IF is 4.566, thus 4.566/10 = 0.457 with descending class ranges for each decile from 4.566 to 4.109 (score 10), 4.109 to 3.652 (score 9), 3.652 to 3.195 (score 8), down to 0.457 to 0 (score 1). This produced a distribution skewed to the lower article scores.
3. *Log (IF+1)/10*. The log distribution of article scores was used to produce a less skewed distribution of the scores. For each journal log (IF+1) was calculated; 1 was added to each IF to avoid negative values for those journals with an IF less than 1. The log of the highest IF+1 was divided by 10, and ten class ranges were determined as before. For example, for dentistry log (4.566 + 1) = 0.745, therefore log (4.566 + 1) /10 = 0.075, so the log class range for score 10 is 0.745 to 0.671, for score 9 is 0.671 to 0.597, for score 8 is 0.597 to 0.522, and so on down to score 1 at 0.075 to 0.

The results for the three article score calculation methods were plotted using Prism (Version 3.02, GraphPad Software Inc, San Diego, CA, USA), a spreadsheet-based graphing and statistical software package.

Results

Table 1 lists the IF of the top ten journals of the 5,550 journals across all categories in the 1999 Journal Citation Reports data file,[11] as well as in the three ISI discipline categories chosen for our study. The variation across disciplines is clear, as are the high IF of review journals; 4/10 of the highest IF journals in all categories are review journals: Ann Rev Immunol, Ann Rev Biochem, Ann Rev Dev Biol, and Current Opinion Cell Biol. What is also evident in this table is the relatively rapid decrease in IF from the top two to four journals to the rest.

The percentage frequency distribution of journals within article scores by calculation method is shown in bar chart form for dentistry in Figure 2. The slightly alternating pattern across article score deciles produced by dividing the number of journals in a discipline by 10 is due to mathematical rounding of the class interval of 4.5, which places four journals on one score and five in the next.

The formula IF/10 provides a distribution strongly skewed towards the lower scores because the journal with the highest IF sets the class intervals and this IF is more than double that of the next journal. The formula log (IF+1)/10 also produces a distribution skewed towards the lower scores but with a wider spread across the deciles. The percentage of journals within the three ISI discipline categories by calculation method is listed in Table 2. The types of distributions produced by the three calculation methods are clear. Empirically, the most discriminating distribution for cross-discipline comparison is with the log distribution because the distribution is wider across the ten scores.

Table 1. Top ten journals according to their Impact Factors in four ISI categories

Rank	All Categories	IF	Dentistry	IF	Medicine	IF	Physiology	IF
1	Annual Review of Immunology	47.56	Journal of Dental Research	4.556	New England Journal of Medicine	28.857	Physiological Review	23.95
2	Annual Review of Biochemistry	37.11	Critical Reviews in Oral Biology and Medicine	3.353	Journal of the American Medical Association	11.435	Annual Review of Physiology	19.8
3	Cell	36.24	Journal of Clinical Periodontology	1.873	Lancet	10.197	Reviews of Physiology Biochemistry and Pharmacology	18.8
4	Nature Genetics	30.69	Clinical Oral Implants	1.816	Annals of Internal Medicine	10.097	Journal of General Physiology	6.382
5	Nature	29.49	Periodontology 2000	1.729	Annual Review of Medicine	7.219	Journal of Physiology - London	4.552
6	New England Journal of Medicine	28.86	Caries Research	1.624	Archives of Internal Medicine	6.705	Journal of Neurophysiology	3.935
7	Nature Medicine	26.58	Journal of Periodontology	1.537	British Medical Journal	5.143	Journal of Biological Rhythms	3.701
8	Annual Review of Developmental Biology	26.26	Oral Microbiology and Immunology	1.526	American Journal of Medicine	4.977	American Journal of Physiology - Renal	3.59
9	Current Opinion in Cell Biology	25.63	International Journal of Oral and Maxillofacial Implants	1.445	Medicine	4.723	Journal of Vascular Research	3.576
10	Science	24.6	Journal of Periodontal Research	1.418	British Medical Bulletin	3.381	American Journal of Physiology - Cell Physiology	3.485

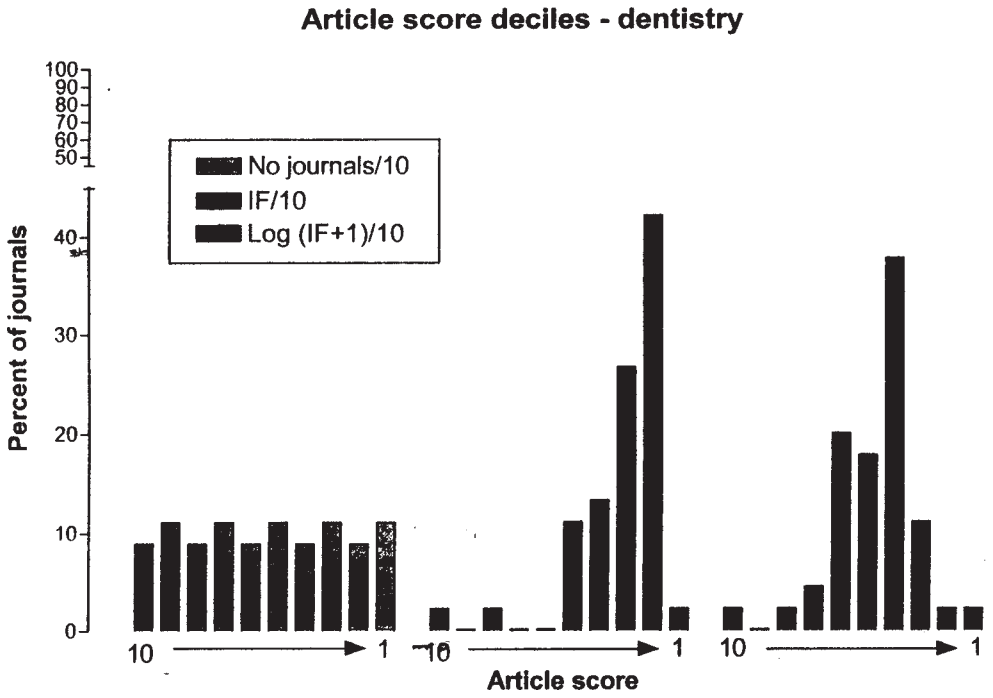

Figure 2. Bar chart to show percentage article score distribution by scoring method

Scores produced with the three calculations are shown in a different manner in Table 3. Here the top ten journals, according to descending IF, in each category have their article scores listed according to calculation method. Based on subjective opinion of the journals listed, the Jn/10[1,2] method rates the journals too highly; IF/10 is too harsh; but log (IF+1)/10 produces a distribution that would fit subjective impressions of journals to judge from the responses of many colleagues to whom the scoring methods have been presented.

Discussion

There is no absolute bibliometric comparison of published article quality that is generally accepted. Many systems have been used with various justifications. In nonmedical disciplines, for example, assessment of articles in economics has been with page counts,[12] and in accounting it has been citation analysis, weighting by peer ranking of journals and article counts.[13] In biomedical disciplines, most publications

Table 2. Percentage of journals in the three ISI discipline categories by article score according to the three calculation methods

	Dentistry Journals n = 45			Medicine Journals n = 110			Physiology Journals n = 74		
Article score	Jn/10	IF/10	Log (IF+1)/10	Jn/10	IF/10	Log (IF+1)/10	Jn/10	IF/10	Log (IF+1)/10
10	10	2	2	10	1	1	10	1	4
9	10	0	0	10	0	0	10	1	0
8	10	2	2	10	0	3	10	1	0
7	10	0	4	10	0	2	10	0	1
6	10	0	20	10	0	2	10	0	3
5	10	11	18	10	0	2	10	0	14
4	10	13	38	10	0	12	10	0	15
3	10	27	11	10	0	10	10	1	31
2	10	42	2	10	3	26	10	20	22
1	10	2	2	10	5	42	10	76	9

Table 3. Article scores of the top ten journals in three ISI discipline categories by the three calculation methods (A = Jn/10, B = IF/10, C = Log (IF+1)/10)

Dentistry	A	B	C	Medicine	A	B	C	Physiology	A	B	C
Journal of Dental Research	10	10	10	New England Journal of Medicine	10	10	10	Physiological Review	10	10	10
Critical Reviews in Oral Biology and Medicine	10	8	8	Journal of the American Medical Association	10	4	8	Annual Review of Physiology	10	9	10
Journal of Clinical Periodontology	10	5	7	Lancet	10	4	8	Reviews of Physiology Biochemistry and Pharmacology	10	8	10
Clinical Oral Implants	10	5	7	Annals of Internal Medicine	10	4	8	Journal of General Physiology	10	3	7
Periodontology 2000	9	5	6	Annual Review of Medicine	10	3	7	Journal of Physiology - London	10	2	6
Caries Research	9	5	6	Archives of Internal Medicine	10	3	7	Journal of Neurophysiology	10	2	6
Journal of Periodontology	9	5	6	British Medical Journal	10	2	6	Journal of Biological Rhythms	10	2	6
Oral Microbiology and Immunology	9	4	6	American Journal of Medicine	10	2	6	American Journal of Physiology - Renal	9	2	5
International Journal of Oral and Maxillofacial Implants	8	4	6	Medicine	10	2	6	Journal of Vascular Research	9	2	5
Journal of Periodontal Research	8	4	6	British Medical Bulletin	10	2	5	American Journal of Physiology - Cell Physiology	9	2	5

on research productivity have been concerned with research output of groups and institutions; they typically use expenditure,[14] research impact,[15] combinations of student/staff numbers and expenditures,[5] and measures such as numbers of authors and byline position.[6,16]

A convenient measure that is readily available for biomedical research is the ISI Impact Factor,[11] which is a way of indirectly measuring reader response to articles in a particular journal. The IF calculation favors general journals over specialist, review journals over journals with no reviews, journals that publish frequently over infrequent journals, and journals in large disciplines over journals in small disciplines.[17] This accounts for the variation in IF values among disciplines.

The principle of using an article score across ten intervals places journal articles into ten "boxes," so to speak. Comparison across disciplines therefore compares the "box" into which a journal article is placed, not the IF of the publishing journal. Relative comparison will show, therefore, that one individual may be publishing in the top score category of his or her discipline compared to another person who is mostly publishing in a much lower category. This should aid decision making on promotion and research grants.

While there are critics of the IF,[10,17-19] as a measure of absolute quality, including the originator of the IF,[9] it is a reasonable basis for relative comparisons across disciplines, particularly since IF are available and the calculation method is clear even if some feel the data base for the calculation is imperfect.[10] Ugolini et al.[1,2] understood this, and their decision to use an article score seems fair. Such a score encourages quality of output and enables comparison between individuals working in the same field and between individuals working in different fields. An advantage of a transparent system of comparison is that it may encourage researchers to publish fewer papers close to the least publishable unit[16] in favor of higher-quality work. This would fulfill Altman's plea for "less research, better research and research done for the right reasons."[3]

Conclusion

Which of the three methods of article score calculation we have presented may be used for relative comparisons is a personal choice. We favor the log (IF+1)/10 calculation because the frequency distribution per article score decile is wider and fits our subjective perception of journal quality better than the other two methods. The calculation is easily done, especially if placed on a spreadsheet that requires the entry only of the highest IF in a category followed by the IF of an individual journal in order to calculate the article score. The article score may be used as a weighting to calculate a total publication score for an individual over a specified period; or one may present results as the proportion of publications in the top 10 percent of journals in a category, that is those with a score of 10, and so on. It must never be used as an absolute measure of the contribution of an individual to knowledge.

Acknowledgment

We gratefully acknowledge the suggestion of Prof. D. Mitchell of the Department of Physiology, University of the Witwatersrand, to use the log method of calculation.

REFERENCES

1. Ugolini D, Parodi S, Santi L. Analysis of publication quality in a cancer research institute. Scientometrics 1997;38:263-74.
2. Ugolini D, Bogliolo A, Parodi S, Casilli C, Santi L. Assessing research productivity in an oncology research institute: the role of the documentation center. Bull Med Libr Assoc 1997;85:33-8.
3. Altman DG. The scandal of poor medical research. BMJ 1994;308:283-4.
4. Taubes G. Measure for measure in science. Science 1993;260:884-6.
5. Davis D. The real world of performance indicators: and review of the use in selected Commonwealth countries. London: Commonwealth Higher Education Management Service (C.H.E.M.S.), 1996:1-21.
6. Ellwein LB, Khachab M, Waldman RH. Assessing research productivity: evaluating journal publication across academic departments. Acad Med 1989;64:319-25.
7. Whitehouse GH. Citation rates and impact factors: should they matter? Br J Radiol 2001;74:1-3.
8. Garfield E. Citation indexes for science: a new dimension in documentation through association of ideas. Science 1955;122:108-11.
9. Garfield E. Fortnightly review: how can impact factors be improved? BMJ 1996;313:411-3.
10. Adam D. The counting house. Nature 2002;415:726-9.
11. Journal Citation Reports on CD-ROM, Science Edition, Institute for Scientific Information, Philadelphia, Penn., USA, 1999.

12. Jin JC, Yau L. Research productivity of the economics profession in East Asia. Economic Inq 1999;37:706-10.
13. Hasselback JR, Reinstein A. A proposal for measuring scholarly productivity of accounting faculty. Iss Account Ed 1995;10:269-305.
14. Ball R, Wilkinson R. The use and abuse of performance indicators in UK higher education. Higher Ed 1994;27:417-27.
15. Kostoff RN. Federal research impact assessment: axioms, approaches, applications. Scientometrics 1993;34:163-206.
16. McAllister PR, Narin R. Characterisation of research papers of US medical schools. J Am Soc Information Sci 1983;34:123-31.
17. Amin M, Mabe M. Impact factors: use and abuse. Perspectives in Publishing. Oxford: Elsevier Science, 2000, No. 1:1-6.
18. Coleman R. Impact factors: use and abuse in biomedical research. Anat Rec (New Anat) 1999;257:54-7.
19. Linde A. Editorial: on the pitfalls of journal ranking by impact factor. Eur J Oral Sci 1998;106:525-6.