Persuasive citations

The discussion forum on citation analysis in the January issue of *TIBS* was revealing in that, looked at side by side, the articles provided compelling evidence for one point upon which the two sets of authors agreed – that a powerful reason behind the selection of citations is the desire of authors to persuade the readers of the validity of their arguments.

One might have thought that in an area such as citation analysis the number of really significant influences would be relatively few and that certain authors and their seminal works would be commonly, if not universally, cited. Yet, of the 17 articles to which Cole referred only one appeared amongst the 25 references listed by MacRoberts and MacRoberts. Admittedly, both authors referred to themselves and to each other, though to different articles and (perhaps not surprisingly to readers of the biochemical literature) considerably more to their own work than that of the other. But while MacRoberts and MacRoberts gave five citations to Eugene Garfield who, at least to readers of *Current Contents*, is a prominent figure in citation analysis, Cole ignored him entirely, presumably because Garfield's work has failed to influence him.

Altogether I found the arguments of both sets of authors quite enlightening and, in view of the strong support each received from cited sources, very persuasive!

A.G. DAWSON

Department of Biochemistry,
University of Technology,
Sydney, Australia.

From bad to worse: evaluation by Journal Impact

MacRoberts and MacRoberts[1] are to be credited for directing the attention of the biochemical community to the increasing misuse of citation analysis in scientific evaluation. They document convincingly that citations are subjective, biased and quite unsuitable as quality indicators. This should come as no surprise: we select our own references primarily because they happen to contain methodology or information relevant to our own work or because they illustrate the particular points we want to make, not because they are particularly original or brilliant – so why should other scientists behave differently?

It seems reassuring that the MacRobertses and Cole[2] agree that citations should not be used for evaluation of individual scientists. However, this advice has been (and probably will be) ignored by many science administrators and evaluators, who have even worse menaces up their sleeve. In several recent evaluations, including one of individual scientists at our institute, the so-called 'Journal Impact' method has been used[3–5]. The method is based on the *SCI Journal Citation Reports*, in which each scientific journal registered in the SCI database is assigned an 'Impact Factor' representing the average annual citation rate of articles published in that journal, measured during the first two years following publication[6]. Examples of journals with different impacts are shown in Table I. For evaluation purposes it is assumed that an article published in a certain journal is likely to receive the number of citations corresponding to the journal's average, hence each article is simply awarded the Journal Impact value. This is the ultimate convenience: there is no need for evaluation committees to read articles, and even no need to enumerate citations; just record the journal addresses of the publications, and the quality of that author's research is revealed. A scientist who publishes in *N. Engl. J. Med.* is five times as good as one who publishes in *Biochem. J.* and 14 times as good as one who publishes in *Cancer Lett.*!

Table I. Journal Impact factors (from Ref. 6).

Journal	Journal Impact factor
Cell	21.0
N. Engl. J. Med.	19.3
Nature	15.0
Science	14.3
J. Exp. Med.	11.1
Proc. Natl Acad. Sci. USA	9.4
J. Cell Biol.	8.5
J. Biol. Chem.	6.4
Lab. Invest.	5.6
Cancer Res.	4.1
Biochem. J.	3.8
Biochim. Biophys. Acta	2.4
Cancer Lett.	1.4
Experientia	1.0
Cancer Biochem. Biophys.	0.6
Caryologia	0.1

In addition to the basic flaws of citation analysis, the Journal Impact method contributes a fundamental logical error: inversion of the causality upon deduction from a correlation. 'Stones cannot fly, mother Nille cannot fly, hence mother Nille is a stone' (Ref. 7). A journal has a high impact because it contains many highly cited articles, but an article does not automatically become highly cited because it is published in a high-impact journal. This is best seen when comparing articles sent by the same author to different journals. Figure 1 shows the correlation between journal impact and real article impact (citation rate) for all articles published by a single author over a 17-year period. There is in fact no correlation, i.e. the articles receive, on average, the same number of citations regardless of whether they are published in high-impact or low-impact journals. The mean impact of the articles published by this author is 7.0, but he has submitted them to journals of lower impact, and would have received an average score of only 3.1 by the Journal Impact method. On the basis of an evaluation by the latter method, this scientist might have missed a grant or a position by sending his papers to the 'wrong' address.

The Journal Impact method is inherently biased by the fact that different scientific disciplines follow different citation practices. The systematic error thus introduced may be of considerable magnitude, cf. the several-fold difference in mean journal impact between the different fields of research shown in Table II. Such differences are partly due to different average numbers of references per article within each field (many references giving a high mean impact); partly to differences in the percentage of recent references, which form the statistical basis of Journal

Table II. Mean impact of journals listed within different fields of research in Ref. 6.

Field of research	Mean impact ± standard error[a]
Gynaecology	1.03 ± 0.12 (29)
Biophysics	1.75 ± 0.31 (32)
Haematology	2.14 ± 0.41 (25)
Biochemistry	2.86 ± 0.38 (138)

[a] Number of journals in parentheses.

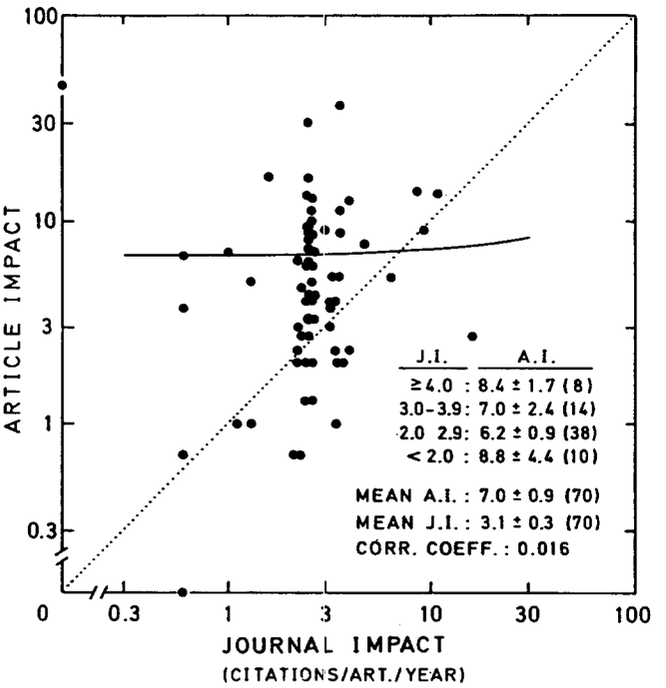

Fig. 1. The relationship between Journal Impact and real article impact (citation rate) for all articles published by one author over a 17-year period, showing that there is no correlation.

Impact factors[8]. Although the bias is strongly in the favour of biochemists, we will hopefully have enough professional pride to refrain from using this as an argument for the adoption of a notoriously misleading evaluation method.

Widespread implementation of the Journal Impact method for scientific evaluation would, in the long run, encourage authors to publish in those journals which presently have a high impact, to the detriment of low-impact journals (why don't publishers protest?) and probably to the detriment of some scientific disciplines better served by specialty journals. However, a word of warning should be issued: to correct for the above-mentioned field differences, some evaluation studies have chosen to measure scientific quality by 'Relative Impact', i.e. the ratio between the real article impact and the corresponding journal impact, the argument being that an impact *above* the journal average indicates supremacy within the corresponding scientific field[4,5,9]. If this method of evaluation becomes fashionable, it will be profitable to publish in low-impact journals rather than in high-impact journals!

The citation-based evaluation methods are clearly founded on erroneous assumptions, and should be discontinued immediately. As responsible scientists we should insist on the same quality standards for scientific evaluation as we require of the scientific work itself. The only acceptable method at present seems to be the peer review, in which published papers are read and evaluated by experts within the same field. Attempts to improve evaluation procedures should be directed towards the standardization of evaluation criteria and conclusion formats, and towards international cooperation to compensate for the lack of objective expertise in small countries. Such efforts would certainly be more worthwhile than the introduction of quick and easy methods which reduce the evaluation workload at the expense of quality and fairness.

References

1 MacRoberts, M. H. and MacRoberts, B. R. (1989) *Trends Biochem. Sci.* 14, 8–12
2 Cole, S. (1989) *Trends Biochem. Sci.* 14, 9–13
3 Luukkonen-Gronow, T. and Suutarinen, P. (1988) *Bibliometric Analysis of Nordic Cancer Research: A Report on Study Data*, FPR-Publication No. 8, Nordic Council of Ministers, Copenhagen
4 Schubert, A., Zsindely, S. and Braun, T. (1985) *Scientometrics* 7, 155–163
5 Vinkler, P. (1986) *Scientometrics* 10, 157–177
6 Garfield, E. (ed.) (1987) *SCI Journal Citation Reports*, Institute for Scientific Information
7 Holberg, L. (1730) *Erasmus Montanus eller Rasmus Berg*, in *Lærde Tidender* (Copenhagen), June 29th
8 Moed, H. F., Burger, W. J. M., Frankfort, J. G. and Van Raan, A. F. J. (1985) *Scientometrics* 8, 177–203
9 Moed, H. F., Burger, W. J. M., Frankfort, J. G. and Van Raan, A. F. J. (1985) *Scientometrics* 8, 149–159

PER O. SEGLEN

Institute for Cancer Research,
The Norwegian Radium Hospital,
0310 Oslo 3, Norway.

The bias of citations

The use of Citation Analysis as a tool to evaluate science is unfair, erroneous and dangerous. It leads to wrong ideas about what science is. It will have perverse effects, as any system designed to evaluate human endeavour modifies the behaviour of the tested population to suit the criterion measured: students prepare themselves for the type of questions that can be expected in multiple choice exams. Unfortunately, Science Citation analysis is easy, provides numbers, and is being lobbied for by administrators and by armchair scientists who derive salary and power from it.

A citation index is not a measure of quality or importance but a measure of recognition. Recognition in science, as in art or other activities, reflects quality but also publicity, power and fashion. The best player at Wimbledon wins by a definite score, the fastest runner wins the race. Unfortunately there is no such objective, absolute criterion for the best scientist.

Computer people are well aware that whatever the apparent precision or fancy representation of computer output, the output is only worth what the input was worth: 'garbage in garbage out'.

Let us start from the input, i.e. the data used in Citation Analysis. The basic assumption of such an analysis is that scientists writing papers by and large apportion their citations proportionally to the importance of the works cited in the methodological or conceptual support of their papers. As pointed out by Cole and MacRoberts, scientists writing a paper do not generally set themselves up as judges of the literature but in a very mundane way try to support as best and as easily as they can their own results and the importance of their own present and past contributions. Moreover, citation may be, consciously or unconsciously, oriented in various ways. I will describe several types of biases:

Self-citation bias: biases towards one's own work, without any malicious intent.

In-house bias: bias for work that one is well acquainted with, because of proximity, friendship, etc.

Journal biases: if you send an article to a journal be sure not to offend the referee by ignoring his work! Why am I cited more frequently by the *same* authors in journals in which I am an