AN EVALUATION OF IMPACT FACTORS AND LEAGUE TABLES

Trust is now gone even in academic medicine and the era of the league table has arrived. All professionals are now familiar with the use of league tables to bully those who are perceived to be under performing, or to allocate additional funding or less stringent cuts to those who are thought to be doing better than average. However, the basis of such rankings is often challenged, either on the basis that the individual elements of the ratings are highly subjective or that the cost of producing a well documented list of ratings outweighs its potential value. One such example is the Research Assessment Exercise (RAE), which rated the quality of research activity in every university department and research institute in the UK and will determine up to one-third of the funding of individual universities for the next five years. However, when the cost of this exercise is computed there will be those who will urge the use of more automated methods in future based by and large on, say, grant income and an analysis of the value of the output by using citation analysis of the published publications of individual scientists in the institute. Indeed, some universities in the UK are already using such techniques to evaluate individual scientists requesting promotion or to allocate of funds to groups within an institute.

Evaluation of individual scientists and the compilation of ratings for departments based on citation analysis depend almost entirely on data gathered by the Institute for Scientific Information (ISI), which is now a private profit orientated company. Quite naturally, the ISI offers to carry out such evaluations for a fee or to sell the appropriate software to the institutes or funding agencies to allow them to do the work themselves.

Since the 1960s the ISI has scanned the reference lists in the major journals and collated the citations to previously published work. The resulting database forms the basis of the modern literature search which enables research workers to quickly and efficiently identify other scientists working on a particular topic using either key words or the names of scientists they know to have contributed to the field. Equally, one can list all the papers that have cited an earlier key paper and determine the novelty of one's own work or that of others in a particular area. Used in this way, the database is extraordinarily valuable to research workers entering a new field, established researchers and, of course, referees and editors of journals. This service is even more valuable when the individual citations are linked electronically to files containing copies of the papers just as they appear in the journal.

However, the ISI has used the same database in a number of other ways, initially to rate journals and more recently individual scientists. This approach is based on the belief that the more important papers are cited most frequently and that the value of a journal can be judged by the frequency with which the papers in it are cited. Thus, the ISI publishes an annual rating known as the impact factor (IMF) for journals which gives a measure of the frequency with which the papers within it are cited. Clearly, the next inevitable step was the rating of individual scientists according to the IMF of the journals in which they published or, more properly, by the citation of their best papers during (for example) the last five years.

The impact factor is an arithmetical measure of the frequency at which the 'average' article in the journal has been cited in a particular year. It is calculated by dividing the 'number of current citations to articles published in the 'journal in the two previous years' by 'the total number of articles published in the journal during the two previous years'. Although much has been written about the deficiencies and limitations of IMF as an index of research excellence, particularly across different specialities, and the ISI clearly indicates the need for discretion when using their published numbers, several institutes have already used IMF and compiled league tables to decide which scientist should contribute to the RAE exercise. Some of the problems are summarised in Table 1.

It is often asked whether statistics complied over two years are representative of long-term citation. Since the average half-life of a published paper with respect to citations is often in excess of six years, it can be assumed that the two-year IMF is a reasonable predictor of the long-term IMF. On the other hand, the modality (or highest likelihood) of citation for a published paper in a particular journal will be associated with the mean of overall citations in the journal only if the distribution of individual citations is normal. Most studies show that the IMF is heavily weighted by a small number of papers with an enormous number of citations, and many of the other publications in the same journal have either zero, or one or two citations. Indeed, several studies have indicated that the distribution is often markedly skewed, with about 15% of the papers receiving 50% of the citations, necessitating log, or even double log, transformation for normalisation. Thus, the geometric mean might best represent the highest likelihood of the IMF in individual publications correlating with the journal IMF. In other words, the variability of the number of citations of a paper

TABLE 1
Known problems with IMF.

- Citation frequency may not be a valid indicator of scientific quality.
- All citations are equal and can include citations such as 'by misinterpretation of their own data, Smith *et al*. 1984' or 'In 2001, we confirmed the findings of Smith *et al*., 1984, etc.'.
- The output data is at the mercy of the input data which undoubtedly contains misspelt names and typing errors in the list of citations which may then be quoted by other authors.
- Impact factors are not suitable to compare different fields or specialities within broad fields, e.g. 'Clinical Pharmacology versus Pharmaceutics' (see Table 2).
- Impact factor of review based journals is higher in general (see Table 2).
- Citations to any type of article (including letters, editorials, communications, meeting abstracts) are used to arrive at a number for total citations, but this is then divided by the number of normal articles and review articles only.
- The IMF for journals from a number of countries, including Russia and China, is reduced by a publication lag further compounded by a delay caused by translation as the paper is entered into the citation records.
- Impact factor is heavily influenced by self-citation and the national bias of North American scientists to cite each other.
- Only a small proportion (perhaps 15%) of highly cited articles determine journal IMF. Thus, journal IMFs are not representative of the average individual article.
- Journal IMFs do not affect the citation of articles published in that journal.
- The journal IMF is not valid for the assessment of the quality of individual papers.
- The journal IMF is not valid for the assessment of the quality of individual scientists.
- The average IMF of the papers of an individual scientist does not necessarily agree with peer assessment.
- As a simple rule, the IMF is unlikely to be a valid for the assessment of the quality of groups of scientists who produce less than 100 papers over two years.

in any one journal is in reality very large and, as such, is a poor predictor of the number of citations a paper accepted by one particular journal will receive compared with a similar paper published in another journal with roughly the same IMF, or indeed an IMF within the same log unit. In a few instances where researchers in this area have deliberately sent groups of almost identical papers to different journals with IMFs ranging between 0·5 and 8, the ratio of individual citations has remained the same within each journal. However, in the journals with the highest IMF the difference between papers was diminished.

Against this background it is difficult to justify the crude use of IMF to assess individual scientists or research groups. However, more and more the consequences of

TABLE 2
Institute for Scientific Information journal citation reports (2000 JCR Science Edition).

Rank	Abbreviated journal title	Impact factor	Cited half-life
1	*N Engl J Med*	29·51	7·1
2	*Nat Med*	27·91	3·3
3	*Nature*	25·81	6·8
4	*Science*	23·87	6·1
5	*Adv Cancer Res*	21·68	7
6	*JAMA*	15·40	6·4
7	*P Natl Acad Sci USA*	10·79	6·4
8	*Lancet*	10·23	6·9
9	*Circ Res*	9·19	7
10	*J Neurosci*	8·50	4·7
11	*Cancer Res*	8·46	6·2
12	*Diabetes*	7·72	5·5
13	*Brain*	7·30	8·1
14	*Brit Med J*	5·33	6·9
15	*Clin Pharmacol Ther*	5·28	8·5
16	*J Physiol-London*	4·46	>10·0
17	*J Pharmacol Exp Ther*	3·45	7
18	*Brit J Clin Pharmacol*	2·15	7·8

such practices are only too clear and research workers are under pressure from heads of departments and institutes to send their papers to journals with the highest IMF. Of course, research workers are good learners and self-citation, often in ways that are difficult to detect, is now the order of the day. Clearly, pressure to publish in journals with high IMFs will distort the dissemination of knowledge. This will be particularly true of small groups of specialists reading a highly selected list of journals (often nationally based). However, looking at Table 2, even the larger specialties in the UK, such as clinical pharmacology, have a difficult choice to make and must more and more be tempted to send their papers to North American journals. In the end this approach will lead to a decline of the standing of European journals as an increasing number of submissions are made to North American journals with higher IMFs. This could have a devastating effect on many scientific societies and communities linked to specialist journals.

NOTE FROM THE EDITOR

Specific citations were not supplied with this guest editorial, as it was considered that doing so would undermine the arguments presented.

For those interested in further reading on this subject, readers are directed to the following sources:

ISI website – www.isnet.com

Garfield Library website – www.garfieldlibrary.upenn.edu/impact factor.html