The impact-factors debate: the ISI's uses and limits

Towards a critical, informative, accurate and policy-relevant bibliometrics.

Sir — Your Opinion article "Errors in Citation Statistics" (*Nature* **415,** 101; 2002) identified how journal impact factors compiled by the Institute for Scientific Information (ISI) are sometimes included as variables in mathematical formulae and directly influence funding decisions by individual research departments. Such use is inappropriate and counterproductive. In addition, the understandably negative reactions of the scientific community towards this type of use mask the great potential of bibliometric methods.

You also uncovered a particular type of inaccuracy in the ISI's citation rate of a paper by a consortium. This is the tip of an iceberg: we have detected errors skewing results among papers, authors, journals and countries, as well as other sources of error, in a large study we have carried out (details available from H. F. M. at moed@cwts.leidenuniv.nl). Scientists subjected to bibliometric assessment and policy officials using it should be aware of these limitations and problems, so that they can properly evaluate and use it (see Box below).

Bibliometric indicators reflect scientific impact, not quality, and provide useful supplementary tools in the evaluation of academic research, provided that they have a sufficiently high level of sophistication; that their pitfalls are taken into account; and that they are combined with more qualitative types of information. Their application can stimulate useful discussion among scientists and research managers about publication strategies and research directions; help peer-reviewers to make quality judgements; and enable policy officials and science administrators to raise critical questions about many aspects of scientific activity and to provide insight for policy (funding) decisions.

Individual scientists may wish to assess their publication strategies, or examine the impact of their work on related fields. Managers of research departments may wish to compare the performance of their department with those of competitors and assess their collaboration strategies. A review committee may have the difficult task of assessing a country's performance in a particular field compared with that of others. The dean of a medical faculty may wish to examine whether it can qualify as a 'research' faculty. A science minister may wish to assess procedures for submitting proposals for research projects to a national research council. Journal publishers or editors may wish to assess the position of their journals, or find suitable reviewers for submitted manuscripts.

For all these needs, context-specific bibliometric indicators were developed as supplementary tools. Yet an indicator that is useful in one context may be inappropriate in another. For instance, in a field in which international journals are dominant channels of written communication, journal impact factors are useful measures if calculated accurately. But such measures have no value in assessing individual scientists or research departments. There can be no direct relationship between statistics such as journal impact factors and policy decisions. Such statistics provide indications and category classifications rather than precise measurements. They need to be adjusted and fine-tuned in close interaction with users and with the scientists who are being evaluated, which may require a long development process. Other types of information should also be taken into account.

In our institute's huge analysis of more than 20 million cited references matched to 8 million target articles extracted from the *Science Citation Index* (*SCI*) and related ISI citation indexes, we found that when data are derived from 'simple' or 'standard' citation-matching procedures, citation statistics at the level of individuals, research groups, journals and countries are strongly affected by sloppy referencing, editorial characteristics of scientific journals, referencing conventions in scholarly subfields, language problems, author-identification problems, unfamiliarity with foreign author names and ISI data-capturing conventions. The overall number of discrepant cited references is about 7% of the number of citations obtained in a simple matching procedure similar to that applied by the ISI in establishing citation links in the Web of Science and calculating statistics for its newsletter *Science Watch*. Typical examples of strongly affected entities are 'consortium' papers; journals with dual volume-numbering systems or combined volumes; journals published in different versions applying different article-numbering systems; and authors from non-English-speaking countries.

This 7% of lost citations skews the distribution of discrepant citations, making some statistics highly inaccurate. For instance, a group of scientists collaborating in a consortium may lose all their joint impact; authors from China or Spain may lose 13% and 8% of their overall citations, respectively; journals publishing combined volumes, such as *Applied Surface Science* and *Physica B*, lose 15–20%. When Spanish or Chinese authors publish the main part of their output in these two journals, the percentage of 'lost' citations can easily rise to 25–30%.

Although the ISI is a monopoly supplier

Crucial questions for producers and users of bibliometric statistics

1. Which version of the SCI was used?
The various versions such as the CD-ROM and the Web of Science have different journal coverage, and statistics may differ between versions.

2. How were publication data collected?
Did the scientists subjected to the analysis verify the data? Which variations of author or departmental names were taken into account? Omission of even one highly cited article can substantially distort the results.

3. What percentage of the total publication output is covered by the SCI and included in the dataset analysed?
As a rule of thumb, if this is below 60%, the picture provided by *SCI*-based statistics may be incomplete.

4. How were cited references matched to target articles?
Did it take into account variations in author names, or discrepancies due to editorial characteristics of articles or journals? Simple matchkeys may yield highly inaccurate statistics.

5. Do the indicators take into account differences in citation and publication characteristics among scientific fields?
ISI journal impact factors do not take these differences into account and therefore have a limited value.

6. What is the policy question to be answered or problem to be solved?
Indicators are context-dependent and need fine-tuning. Those that are useful in one context may be inappropriate in another.

7. What quality factors are used by evaluators and what is their relative weight?
Evaluators must make this explicit, and not hide behind bibliometric indicators.

8. Do the procedures allow for comments by the scientists subjected to the analysis?
Knowing both sides is indispensable for a proper interpretation of bibliometric statistics.

of the *SCI* and Web of Science, it is not a monopoly supplier of bibliometric statistics derived from these bibliographic information products.

Bibliographic and bibliometric use are two distinct types of use of scientific information, each with its own set of operational and quality criteria. The ISI's information products are primarily developed for bibliographic use. When conducted properly, bibliometrics can unravel relationships that were previously unknown, and put new issues on the political agenda. It can be informative in providing condensed overviews of publication and citation frequencies, and accurate if proper data-collection procedures are applied.

Anyone confronted with bibliometric statistics derived from the *SCI*, intended to be applied at the level of individuals or research groups or departments, should know the answers to the questions summarised in the Box.

These minimum criteria are crucial for assessment of the accuracy, validity and usefulness of bibliometric statistics.

Henk F. Moed
Centre for Science and Technology Studies (CWTS), Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands.

Statistics hide impact of non-English journals

Sir — Your opinion article (*Nature* **415,** 101; 2002) provides timely cautions about the errors of *Science Citation Index* (*SCI*) data compiled by the Institute for Scientific Information (ISI). As Chinese scientific publishers, editors and researchers, we wish to point out that these errors are far more serious for journals published in non-English-speaking countries. The ISI's coverage of scientific journals from these countries is far too limited, and the livelihood of many decent journals has been an unintended casualty.

The ISI misses the fact that non-English-language journals often have alternative names. The *Chinese Journal of Geophysics — Chinese Edition* is also called *Diqiu Wuli Xuebao, Acta Geophysica Sinica* and *Chinese J. Geophys.* for historical reasons. In 2000, this journal was cited more than 260 times, yet the ISI's *Journal Citation Reports 2001* gives it a total citation score of 13.

The ISI fails to note that many English-language editions of journals published in China have Chinese-edition counterparts, with different contents and domestic and ISSN registration numbers, and should be considered as different journals. Among China's 12 English-language journals indexed by SCI, 8 also have Chinese editions. In 1998, for example, 91 of the 147 cited items for papers published during 1996 in the *Chinese Science Bulletin*, a prominent science journal, were incorrect: 52 of the 91 errors were attributions to the bulletin, although these citations were actually to the Chinese edition.

In addition to your reminder to people to use citation statistics prudently, we suggest the ISI should pay more attention to journals in non-English-speaking countries. Even though the ISI should not be held responsible for problems in doing science in developing countries, it can certainly be more accurate in its analysis of scientific achievements in places such as China, and hence help to promote international scientific communication.

Shengli Ren*, Guang'an Zu*, Hong-fei Wang†
**Department of Publication, National Natural Science Foundation of China, 83 Shuangqing Road, Haidian District, Beijing 100085, China*
†Institute of Chemistry, Chinese Academy of Sciences, No 2, 1st North Street, Zhong Guang Cun, Haidian District, Beijing 100080, China

Strange results mean it's worth checking ISI data

Sir — Like the editors of *Nature* (see *Nature* **415,** 101; 2002), we too realize that it is in our best interests to carry out editorial checks of data from the Institute of Scientific Information (ISI). In 1998, *The Lancet's* impact factor, as calculated by ISI, dropped to 11.79 from a previous stable value of about 17 for the previous four years. In 1997, *The Lancet* had decided to divide letters into Correspondence (not counted in ISI's denominator) and Research Letters (peer-reviewed and containing original data, hence coded by the ISI as citable for the denominator). This division increased our number of citable items and thus we attributed most of the drop to this change, an assumption confirmed by the ISI.

Recently, we noted that the number of citable items listed for 2000 was higher (821) than informal calculation would suggest. After hand-coding each issue, we found that 684 items should form the denominator.

Meanwhile, out of editorial interest, we looked at citation data for 1999 in a file purchased from the ISI (based on papers published in 1997 and 1998). As *Nature* did, we also found examples of unexpectedly low numbers of citations for large trials that had only a group name in the author byline compared with those that had at least one named author. The International Stroke Trial (*Lancet* **349,** 1569–1581; 1997), for example, was listed as having one citation in 1999. The ISI confirmed to us that group names are a potential landmine for citation accuracy and said it was in the process of developing a new program to address this issue.

Kathleen D. Hopkins, Laragh Gollogly, Sarah Ogden, Richard Horton
The Lancet, *84 Theobald's Road, London WC1X 8RR, UK*

See also the News Feature "The counting house" on pages 726–729 of this issue.

Habilitation not just alive in France, but growing

Sir — In your News Feature (*Nature* **415,** 257; 2002), you suggest that the *Habilitation* postdoctoral thesis is "unique to German-speaking countries". Sadly, not so. *L'habilitation à diriger des recherches* is alive and kicking and France, and is an obligation for anyone who supervises a PhD student. Although not usually as large as its German equivalent (mine was only 15,000 words long), the rules governing its size are effectively determined by each faculty. There is a tendency for it to creep up to the size of the old *thèse d'état*, usually more than 100,000 words long, which it was designed to replace in 1988.

Matthew Cobb
Laboratoire d'Ecologie, Université Paris-6, 75005 Paris, France

Getting space camera back on track soon

Sir — The headline on your News in Brief "Equipment failure derails space projects" (*Nature* **414,** 835; 2001) does a disservice to at least one of the projects discussed, the NASA–European Space Agency Cassini mission to Saturn. Your article accurately states that Cassini engineers are making progress in fixing haze on images that is thought to result from contamination on the camera's optics or detector. It is, however, incorrect to say that Cassini has been derailed by this problem.

Franklin O'Donnell
Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California 91109, USA

Errata In the Correspondence "How scientists can take the initiative in schools" by Mo Afzal (*Nature* **415,** 364; 2002), the citation *Nature* **414,** 1; 2001 should have read *Nature* **414,** 673; 2001.

In the Correspondence "False samples are not the same as blind controls" by L. S. Mills (*Nature* **415,** 471, 2002), reference 5 was incorrect. The correct reference in full is: M. K. Schwartz *et al. Nature* **415,** 520–522 (2002).