nature INDIA

doi:10.1038/nindia.2009.133; Published online 15 May 2009

Commentary

Is citation a good criterion?

Gajendra Pal Singh Raghava

Recently, Thomson Reuters honoured eight Indian scientists for excellence in research. The awards recognise their influential contribution to the global research community.

Since this was the first time these awards were given in India, there was significant curiosity on several grounds — the awards, the criteria of selection, the awardees and their work.

Many viewed the award with skepticism wondering if it was a serious effort to recognise leading Indian scientists or just another brand building exercise by the institutors of the award (Thomson Reuters) to further their cause in India.

A little online **research** revealed that Thomson Reuters had used 'research front methodology' as the yardstick to assess the level of influence of individuals in the specific scientific field.

The author is seen here (1st row, 4th from left) with other Thomson Reuters India awardees, the organisers & C. N. R. Rao (1st row, 2nd from right), who presented the awards.

Research front methodology looks at patterns of intense communication between scientists and is compiled using citation analysis of a list of highly cited papers, defined as the top one per cent of papers in each of the 22 disciplines.

Thomson Reuters reviews how these highly cited works are related and determines how often these papers have been jointly cited (frequency of co-citation of the two highly cited papers) through Essential Science Indicators (ESI), and this is called 'research front analysis'. This tool can assist in identifying areas where important work is being done as also where the scientific community is focusing its attention.

A new technique

In simple words 'research front methodology' is a new technique used to index researchers based on citation of their published papers, which is different from existing techniques like g-index or h-index.

One of the major challenges in science is to measure performance of an individual or country using some objective measures. One of the simplest ways to do it is counting the number of papers published. The main drawback of this measure is that the quality or standard of journals is not considered and quality varies from junk to highly reputed journals (Nature, Science and Cell). Thus simply counting the number of papers is useless.

Now, the question is how to rank journals in order to overcome this problem. Here the impact factor (IF) of journals comes into the picture. The IF is calculated based on citations of papers published in a journal. It means, performance of an individual is counted based on papers published in high IF journals. The IF criterion has also met a lot of criticism. IF is not uniform among disciplines/subject, in some fields even top journals have low IFs. IF of a journal might fluctuate over the years. IF is based on average citation. All the papers of journal are not equally cited, only a few per cent are. This means even if your paper is highly cited it will not get credit if it is published in low IF journals. The reverse is also true – even if your paper is not cited by anybody, you get the credit if it is published in a high IF journal. Another problem is variation in IF of a journal every year. It means scientist performance will vary every year with variation in IF of a journal, where one has published the papers.

In order to measure performance and scientific impact of a scientist, h-index was introduced by Jorge E. Hirsch, a physicist at UCSD. Here, a set of articles of a scientist is ranked in decreasing order of the number of citations that they receive, h-index is the number of maximum papers where each paper gets more or equal citations to this number. The index can also be applied to measure the productivity and impact of a group of scientists, such as a department or university or country. Though h-index is a powerful technique to measure performance, it has a number of limitations: i) it depends on number of publications, ii) is affected by age of the scientist and, iii) all highly cited papers get equal weight.

Another parameter called g-index was introduced based on the average of highly cited papers (g-index is more or equal to h-index); two scientists having equal h-index may have different g-index.

All performance-measuring indices including 'research front methodology' recently introduced by Thomson Reuters are based on citations received on articles published by a scientist. There are a number of sources for getting citation information like web sciences, Google scholar and Scopus.

Citation = quality?

This raises a fundamental question whether citation of a paper reflects its scientific quality.

In reality this in not true. Following are few major issues against this hypothesis.

* Most awards including the Nobel are not based on citation of papers. For example, Oliver Howe Lowry, whose paper (Lowry's method for measuring amount of protein in solution) got the highest citation in history, never got a Nobel Prize. The argument was that it was not serious science, instead it was a method. Whether the argument was justified or not is a matter of debate. Another case in point is of Whitesides, G. M. (h-index 147) and Karplus, M. (h-index 135) who are among the topmost living chemists in term of h-index but have not got a Nobel, whereas a number of scientist with much lower h-index have (e.g Curl, R. F, h-index only ~ 52).

* Most review papers get high citation in comparison to original research papers. It means one should write more review papers than original papers in order to get more citations.

* Citation is affected by the age of scientist in research as citation depends on the age of paper. This means young researchers have lower chances to compete with well-established scientists. This also means senior scientist will have high h-index or g-index than junior scientists. Fortunately, research front methodology considers recent publications in a given period.

* When counting a citation, we don't consider where the paper has been cited. It is possible that a paper is cited in the top reputed journals whereas another paper is cited in journals that have no IF. The paper cited in reputed journals has more impact on science than a paper cited in lower journals. This is a chicken-egg problem: the journal's IF depends on citation and citation depends on which journal the paper is cited.

* Self citation is another important issue. It is not limited to an individual journal but applies to a group of journals and countries.

* The depth understanding of scientific problems versus application or scientific methods is another important issue. Most scientific methods or tools commonly used by scientists are highly cited. In contrast, papers which address serious or complex issues get low citations as they are only used by scientists working closely in the field.

* Citation also depends on visibility and advertisement of a paper. A paper published in an open access journal, which is freely available to scientific community, will get high citations. Also, citation depends on how the paper is advertised on internet via blogs, lists and groups.

* Citation of a paper depends on usability of work. A method available to public in the form of service (software or web server) will get much more attention and citation than a simple theoretical method. This has logic because other scientists will only take interest if your work if it is useful to them.

* One criticism of citation is credit in case of multiple authors. Even if one is not the main author, every user gets equal credit. This means scientists who work in collaboration with other groups will have high citations.

However, despite a number of criticisms of citation-based evaluation, it is thus far the most evolved criterion for measuring performance of a scientist or group or nation. Interestingly, Thomson Reuters is predicting Nobel prize winners with reasonably high accuracy based on citation records of scientists.

The author, among this year's Thomson Reuters awardees, is the Head of the Bioinformatics Centre at Institute of Microbial Technology, Chandigarh, India.

Nature India EISSN 1755-3180