5-1-2009

Measuring up: how does the h-index correlate with peer assessments

Research Trends Editorial Board

Recommended Citation

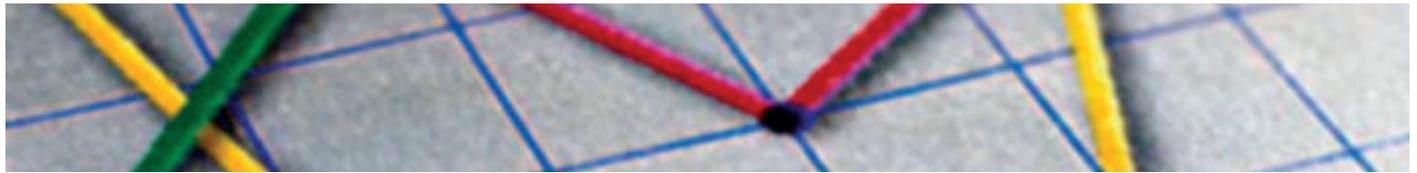## Research trends

Measuring up: how does the *h*-index correlate with peer assessments?

Since it was first proposed in 2005, Hirsch's h-index *(1)* has made a considerable impact on both bibliometricians and the wider scientific community by offering an additional yardstick for assessing individual researchers' scholarly output and influence. Hirsch's original paper has been cited more than 280 times in journals, conference proceedings and book series in 14 languages from fields as diverse as medicine and mathematics to engineering and economics (data from Scopus).

The h-index is defined as the number of an individual researcher's articles that have received the same number (or more) of citations since publication. It is easily derived from any comprehensive list of an author's papers by ranking them in descending order of citations received and then identifying the rank position at which the number of citations is not less than the ranked value. Since it combines measures of productivity (the upper limit of the h-index for a given author is the total number of papers published) and a proxy for quality (citations received), it has become an attractive all-in-one metric for comparing researchers.

The h-index, and the numerous variants that have proliferated since 2005, can only be used to compare researchers within the same research field; this is true of all metrics that do not account for the publication and citation practices of the various research fields.

Is the *h*-index a match for peer assessment?
An important and interesting question when evaluating individuals is how well the results of bibliometric assessment compare with peer assessment.

For many years, Lutz Bornmann and Hans-Dieter Daniel, at the Swiss Federal Institute of Technology in Zurich and the University of Zurich respectively, have been investigating the review processes used by funding institutions.

Explaining their findings, Bornmann says: "In two investigations *(3, 4)*, we have shown that for individual scientists the h-index correlates well with the number of publications and the number of citations that these publications have attracted. This is hardly surprising given that the h-index was proposed to do exactly that."

In three studies *(2, 3, 4)*, they also examined the relationship between the h-index and peer judgments of research performance. "In these studies, we have shown that the average h-index values of accepted applicants for biomedicine research grants are statistically significantly higher than for rejected applicants."

Impact versus quantity
However, the h-index has certain disadvantages, including a bias towards older researchers and a failure to place emphasis on highly cited papers. This has led to the development of numerous variants of the h-index. The m-quotient, for example, is computed by dividing the h-index by the number of years that the scientist has been active since the first published paper. Unlike the h-index, the m-quotient avoids a bias towards more senior scientists with longer careers and more publications.

Another variant, the a-index, indicates the average number of citations of publications in the Hirsch core (publications with ≥h citations). In contrast to the h-index, which corresponds to the number of citations for the publication with the fewest citations in the Hirsch

Lutz Bornmann

Hans-Dieter Daniel

Best-practice: getting the most out of the *h*-index and variants

Use several indicators to measure research performance: the publication set of a scientist, journal, research group or scientific facility should always be described using a multitude of indicators, such as the numbers of publications with zero citations, highly-cited papers and papers for which the scientist is first or last author. Non-publication indicators, such as awards, grant funding and speaking engagements could also be used.

To measure the quality of scientific output using h-index variants, it is sufficient to use just two variants: one that measures productivity and one that measures impact (e.g. the h-index and a-index) *(5)*.

If the h-index is used to evaluate research performance, the fact that it is dependent upon the length of an academic career and the field of study in which the papers are published and cited should always be taken into account. The index should only be used to compare researchers of a similar age and within the same field of study.

core, the a-index is meant to give more weight to highly cited papers.

Bornmann says: "The results of our study *(5)* show that the *h*-index and its variants are, in effect, two types of indices: one type describes the most productive core of a scientist's output and the *number* of papers in that core; the other type depicts the *impact* of those papers in the core."

Using indices wisely

Bornmann and Daniel believe that while their studies *(2, 3, 4)* provide an initial confirmation of the *h*-index's validity, more time and research is required before it can be used in practice to assess scientific work.

"As a basic principle, it is always prudent to use several indicators to measure research performance," says Bornmann. "The publication set of a scientist, journal, research group or scientific facility should always be described using a multitude of indicators, such as the numbers of publications with zero citations, highly-cited papers and papers for which the scientist is first or last author."

Bibliometric indicators can and should be used to support peer review, especially where efficiencies are sought. Current research clearly supports the hypothesis that such indicators can approximate the results of peer review, and many research institutes and research councils are already using indices to support their assessments. Informed peer review currently is the state of the art of research evaluation.

Useful links:
The *h*-index: Hirsch's original 2005 paper

References:
(1) Hirsch, J.E. (2005) "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, pp. 16569–16572.
(2) Bornmann, L. and Daniel, H.-D. (2005) "Does the *h*-index for ranking of scientists really work?", *Scientometrics*, Vol. 65, pp. 391–392.
(3) Bornmann, L. and Daniel, H.-D. (2007) "Convergent validation of peer review decisions using the *h*-index. Extent of and reasons for type I and type II errors", *Journal of Informetrics*, Vol. 1, pp. 204–213.
(4) Bornmann, L.; Wallon, G. and Ledin, A. (2008) "Is the *h*-index related to (standard) bibliometric measures and to the assessments by peers? An investigation of the *h*-index by using molecular life sciences data", *Research Evaluation*, Vol. 17, pp. 149–156.
(5) Bornmann, L.; Mutz, R. and Daniel, H.-D. (2008) "Are there better indices for evaluation purposes than the *h*-index? A comparison of nine different variants of the *h*-index using data from biomedicine", *Journal of the American Society for Information Science and Technology*, Vol. 59, pp. 830–837.

Further reading:
(6) Bornmann, L. and Daniel, H.-D. (2009) "The state of *h*-index research", EMBO Reports, Vol. 10, pp. 2–6.
(7) Bornmann, L and Daniel, H-D (2008). "Selecting manuscripts for a high-impact journal through peer review: a citation analysis of communications that were accepted by *Angewandte Chemie International Edition*, or rejected but published elsewhere", *Journal of the American Society for Information Science and Technology*, Vol. 59, issue 11, pp. 1841–1852.