



H-index: an index to quantify the impact of scientific research

In a previous editorial I discussed the impact factor (IF), an index frequently used to quantify the relative importance of a journal within its field, with journals with higher IFs deemed to be more important than those with lower ones. I also stressed that the IF is often misused as a quality indicator for articles published by a given author. Now I wish to discuss an index which has been developed with the specific aim of quantifying the impact of the scientific research of a particular individual or, in simpler words, to describe how good a researcher is: the h-index.

The h-index attempts to measure both the scientific productivity and the apparent scientific impact of a scientist. The index is based on the set of the scientist's most cited scientific articles and the number of citations that they have received in other scientists' publications. The index was introduced by Jorge E. Hirsh, as a tool for determining theoretical physicists' relative quality¹ and it is also referred to as the 'Hirsch index' or 'Hirsch number'.

The h-index is intended to address the main disadvantages of other bibliometric indicators, such as total number of articles or total number of citations. Total number of articles does not account for the quality of scientific publications, while total number of citations can be disproportionately affected by participation in a single publication of major influence. The h-index is calculated on the distribution of citations received by a given researcher's publications. In other words, an author, having an h-index of 15, has published 15 articles each of which has been cited by others at least 15 times². The h-index reflects both the number of publications and the number of citations per publication.

The h-index is much less affected by methodological papers proposing successful new techniques or methods that can be highly cited. Only scientists working in the same field can be properly compared with the h-index, since citation rates differ widely among different research fields. Limitations of the h-index are³:

- It depends on the total number of publications. It grows as citations accumulate over the years and thus it depends on the 'academic age' of a researcher, the older scholar having the higher index (if his/her papers are cited). Scientists with a short career are at a disadvantage, regardless of the importance of their publications.
- It is heavily affected by the accuracy of the citation databases from which it is calculated. Erroneous citations will not properly match to the correct author or paper, often resulting in a higher h-index for mistakenly attributed authors. This is more common for authors having very common family names.
- It does not consider the context of citations. An article can be cited in the introduction, otherwise having no other significance or when a citation is made in a negative context or to fraudulent papers.
- It does not take into account self-citations. A researcher working in the same field for a long time will likely refer to his or her previous publications, with growing self-citations affecting the h-index in a self-referenced way.
- It does not account for confounding factors such as 'gratuitous authorship', a common practice in some research cultures, and the favourable

citation bias associated with review articles. In addition, two equally capable researchers could agree to share authorship on all of their papers, thus increasing each of their h-indices.

- It does not account for single successful publications. Two authors may have the same h-index, for example 20, but one has 10 papers that have been cited more than 100 times and the other has none.

The h-index can be calculated using free Internet databases, such as Google Scholar or subscription-based databases. Those who will attempt to calculate their own h-indices will soon find out that each database is likely to produce a different h-index, because of different coverage in each database. The calculation may not be so simple as it looks since all

the papers belonging to the right author need to be carefully identified, eliminating those from authors having similar or even identical names. It may take a while, so have fun.

Yours truly,
Marco Esposito
Editor-in-Chief

■ References

1. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A* 2005;102:16569-16572.
2. McDonald K. Physicist proposes new way to rank scientific output. Available at: <http://www.physorg.com/news7971.html>. Accessed 15 March 2010.
3. Wendl MC. H-index: however ranked, citations need context. *Nature* 2007;449:403.

