

HOW TO EVALUATE SCIENTIFIC PRODUCTION

Como avaliar produção científica

ALBERTO AZOUBEL ANTUNES

ABSTRACT

Objective: To conduct a critical analysis of the two main bibliometric indexes used by science: the impact factor and the H index. **Method:** Research was conducted on PubMed using the keywords: impact factor, Bibliometrics and H index. **Results:** The citations of articles tend to follow a curve in which the articles published in a given year increase sharply to a peak occurring between two and six years after its publication. From this peak citations decline over time. **Conclusion:** The analysis of the scientific merit should not be based on only in bibliometric measure, but in the association of various parameters. The impact factor and the H index is mainly based on the number of citations of scientific papers, and this parameter, although important, should not be used alone, nor overvalued in the evaluation of teaching merit.

Key Words: Impact factor. Bibliometrics. H index.

INTRODUCTION

The teaching performance in research involves three key steps: the production of knowledge, its appropriate dissemination, and finally its application. The quantitative and qualitative evaluation of this process is of great importance, since it allows to identify the most productive centers and who need to redesign to improve its production. The science that is dedicated to this evaluation is bibliometrics. It constitutes a quantitative and statistical technique that measures levels of production and dissemination of knowledge, follow the development of several scientific areas and authorship standards, publication, and the use of research results.

The objective of this study was to perform a critical analysis of the two main bibliometric indexes used by science: the impact factor (IF) and the H index

METHODS

It was held a Pubmed search using the following keywords: impact factor, bibliometrics, H index.

RESULTS

The IF is measured reflecting the average number of citations of scientific articles published in a given journal. It is often used to evaluate the importance of a given journal in its field, and those with higher IF are considered more important than those with less. This factor was created by Eugene Garfield, founder of the Institute for Scientific Information (ISI), now part of Thomson Reuters Corporation. Since 1972 the IF is calculated annually for journals indexed in ISI, and then published in the Journal of Citation Reports (JCR), also from Thomson Reuters.

It is calculated by dividing the number of citations that a journal receives in the two preceding years by the number of articles published in the same years. In another way, if A=number of times that the articles published in 2007 and 2008 for certain journal were cited by indexed journals during 2009 and B=total number of "citable items" published in 2007 and 2008 of this same journal (usually articles, reviews, conference abstracts or notes, not counting editorials or letters to the editor), it can be said that the $IF_{2009} = A / B$.

The 2009 impact factors are published in 2010 and, for this, they cannot be calculated until all 2009 publications

have been received by the indexing agency. New journals receive their respective IF only after two years of indexing. The counting of the IF of yearbooks and irregular publications is usually affected. This factor relates to a specific period of time; then it can be calculated for any desired period, and the JCR includes IF in five years. It also shows journals listings for IF and, if desired, either area or discipline. Among the medical journals with great IF stand out the New England Journal of Medicine (IF: 54,420), Nature (IF: 42,351), The Lancet (IF: 39,207), Science (IF: 31477), among others.

Citations of articles published in a given year increased sharply to a peak occurring between two and six years after its publication. From this peak citations decline over time.

The citation curve of any journal can be described by the relative size of the curve (in terms of area under the line), the extent to which the peak of the curve is close to the origin, and the rate of decline of the curve. These characteristics form the basis of the ISI indicators impact factor, immediacy index and cited half-life. Of the three measures described above, the impact factor is the most commonly used and also most misunderstood.

The H index refers to the higher "h" number of papers that a researcher have that, at least, had the same "h" citation number in each one. For example, a researcher with H 30 index is one that published at least 30 scientific papers that have been cited in at least 30 other journals. The weighting excludes some papers with very low citations. Also disregards highly cited articles if they are isolated examples. This index is indicator proposed by physicist Jorge Hirsch in 2005 to measure both the productivity and the impact of researcher, based on his most cited articles. Its application has spread and the concept is also applied to measure the productivity and impact of research groups, universities, scientific journals and even nations.

DISCUSSION

The value of IF is influenced by sociological and statistical factors. Among the sociological include: the thematic area of the journal, the type of the journal, and the average number of authors per article (which varies according to the theme). Statistical factors include: the size of the journal (number of papers/year) and the period of citations.

In relation to the theme, the basic and fundamental areas tend to have higher IF average than specialized or applied areas. The variation comes to be such, that the best

journal in one determined area may have FI lower than the worst journal of another area. The influence of the number of authors can also be marked, for journals that have only three or less authors per article tend to have a lower FI. One of the reasons that can explain this phenomenon lies in the fact that more authors known by the scientific community increases the chances of being cited by their peers, who have just becoming aware of the article in question. The type of journal can influence the citations by the types of article that it publishes. Journals with the highest number of review articles tend to have higher IF, as these articles are most read by the scientific community.

With regard to statistical factors stands out the size of the journals. An analysis of 4000 journals revealed that the IF variation of the ones that publish more than 150 articles per year is less than 20%. On the other hand, the variation of journals publishing less than 35 articles per year usually it is greater than 40%. The time (or window) period influences FI variation rates. Periodic evaluated on a five-year window usually have less variation evaluated in just two years. Various other features may influence the FI (Figure 1).

- ❖ Fator numerador x denominador;
- ❖ Baixa correlação: n de citações do artigo e do periódico;
- ❖ Privilegia áreas com vida média de citações curtas
- ❖ Favorece língua inglesa;
- ❖ Propensão dos autores de países desenvolvidos de citar trabalhos de seu próprio idioma/país; nos países em desenvolvimento a tendência é inversa;
- ❖ Pode acelerar de forma inadequada o processo de amadurecimento científico

FIGURE 1 – *Impact factor limitations*

In general, what is taken into account in the numerator is not always in the denominator. This fact tends to increase the IF average of journals; the citation index of articles is the one that determines, not the contrary; is observed low correlation of the number of citations of individual articles with the IF of the journal; the indicator gives priority to areas that have short life papers (life sciences and exact sciences) and other disadvantages; as mentioned, review articles receive more citations than original articles and some publishers tend to favor this type of article in its periodic; there is clear predominance of journals in English, despite recent efforts to boost journals that portray the local science in some areas through its inclusion in the JCR.

The isolated use of IF in academic evaluation can be highly destructive, according to the signatories of the DORA - San Francisco Declaration on Research Assessment - a document endorsed by more than 150 scientists and 75 scientific organizations in December 2012. According to it, this practice can prevent periodic publishing articles in areas or subjects less cited, in addition to overload journals with high impact with often inadequate submissions. But the most harmful consequence for science is to prevent the natural progress of research that, in the search for new approaches, can lead to relatively long periods without generating publications. Researchers should be able to "take advantage" of this period without publication and citation without being penalized for it. What can also

occur in some areas is the acceleration of a maturation process, with researchers increasingly trying to publish in higher impact journals. In other cases, however, it is not a natural evolution. "You cannot expect that sociology will present the same internationalization process as physics. In parallel, there must be mechanisms for looking a national journal, published in Portuguese, and say: this journal is good. Will not be the citation and the H index that will give this parameter".

It should surprise the fact that the use of an indicator becomes eligible someone by the fact that he published in a top IF journal, leaving the impression that it is more important to know where the article is published than read it. DORA stresses the need to evaluate research on its own merits and not by the journal in which it is published.

Several criticisms have been made regarding the use of IF. In addition to the debate on the real use of the citation metric, most of the criticism is related to the real validity of the IF (this includes the self-citation), its handling and its possible misuse. Another aspect criticized is the fact that journals that publish only review articles and original papers and review articles have higher IF than journals that publish only original articles. Finally, factors such as the number of journals by knowledge area, the variation in the number of references per article in each area, or regionalism of some areas and journals should be discussed. Thus, many factors must be considered when interpreting the value of the IF of a given journal and use them on evaluations of scientists and institutions.

In relationship to H index, it can be said that its main advantages are: 1) can be easily obtained by anyone with access to databases such as Web of Science, and is easy to understand; 2) can characterize the scientific output of a researcher objectively, especially in areas where there is consolidated culture publication in refereed journals, and can be useful in making decisions about promotions, allocation of funds and allocation of premiums; 3) combine quantity and impact of research on a single indicator; 4) has better performance than the other individual indicators, such as IF, number of articles, number of citations, citations per paper and number of highly cited papers, to assess the scientific productivity of the researcher.

Among its main disadvantages stand out: 1) does not serve to compare researchers of different areas, because the volume of citations varies according to the size of each community of researchers; 2) it can also be manipulated by means of self-citations; 3) gives the books the same weight of the original articles, making it hard to compare researchers from areas where there is a culture of publishing the research results in books, such as the humanities; 4) does not consider the context of the citation, does not distinguish between article by one researcher or a small group and article with hundreds of authors whose individual participation is difficult to assess; 5) the impact of a publication is not only measured by citations, but also for many other things such as their contribution to technological innovations or for the formulation of public policies, for example.

Finally, other inconsistencies are noted when evaluating the H index. The Italian Mauro Degli Esposti, professor at the University of Manchester in the UK, recently compiled a list of researchers from all areas with H index above 100, based on Google Scholar data. In its ranking, with nearly 200 names, appear very few researchers in human and social applied sciences - for example the Nobel economist Joseph Stiglitz (H 130 index) and linguist Noam Chomsky (index H 123) -, and prevalence in the highest strata the scientists from medicine and biochemistry; furthermore,

there is no direct correlation between the Nobel winners and the top of the list. In the Top 30, there are only four Nobel and one winner of the Fields Medal, the main honor of the young mathematicians.

CONCLUSIONS

The analysis of scientific merit should not be based on only in bibliometric measures, but in a pool of measurements. The IF and the H index are mainly based on the number of citations of scientific papers, and this parameter, although important, should not be used alone, nor overvalued in the evaluation of teaching merit.

RESUMO

Objetivo: Realizar análise crítica dos dois principais índices bibliométricos utilizados pela ciência: o fator de impacto e o índice H. **Método:** Foi realizada pesquisa no Pubmed utilizando as palavras-chave: impact factor, bibliometrics, H index. Resultados: As citações dos artigos tendem a seguir uma curva em que os artigos publicados em determinado ano aumentam agudamente para um pico que ocorre entre dois e seis anos após sua publicação. A partir deste pico, as citações declinam ao longo do tempo. **Conclusão:** A análise do mérito científico não deve ser baseada em apenas uma medida bibliométrica, mas em associação de medidas. O FI e o índice H se baseiam fundamentalmente no número de citações dos artigos científicos, e este parâmetro, apesar de importante, não deve ser utilizado sozinho e tampouco supervalorizado na avaliação do mérito docente.

Descritores – Fator de impacto. Bibliometria. Índice H.

REFERENCES

1. Garfield E. The history and meaning of the journal impact factor. JAMA 2006;295:90-93.
2. McVeigh ME, Mann SJ. Journal impact factor denominator: Defining citable items. JAMA 2009;302:1107-1109.
3. Mabe M and Amin M. Impact factors: Use and Abuse. Perspectives in publishing 2007.
4. Os limites do índice H. Pesquisa FAPESP; maio de 2013; 35-39.

Received on: 19/02/2015

Accepted for publication: 12/09/2015

Conflict of interest: none

Source of funding: none

Address for correspondence:

Alberto Azoubel Antunes

antunesuro@uol.com.br