

Maurício Lima Barreto^{I,II}

Erika Aragão^{III,III}

Luis Eugenio Portela Fernandes de Sousa^{I,II}

Táris Maria Santana^{II}

Rita Barradas Barata^{IV}

Differences between h-index measures from different bibliographic sources and search engines

ABSTRACT

OBJECTIVE: To analyze the use of the h-index as a measure of the bibliometric impact of Brazilian researchers' scientific publications.

METHODS: The scientific production of Brazilian CNPq 1-A researchers in the areas of public health, immunology and medicine were compared. The mean h-index of the groups of researchers in each area were estimated and nonparametric Kruskal Wallis test and multiple comparisons Behrens-Fisher test were used to compare the differences.

RESULTS: The h-index means were higher in the area of Immunology than in Public Health and Medicine when the Web of Science base was used. However, this difference disappears when the comparison is made using Scopus or Google Scholar.

CONCLUSIONS: The emergence of Google Scholar brings a new level to discussions on the measure of the bibliometric impact of scientific publications. Areas with strong professional components, in which knowledge is produced and must also be published in the native language, *vis-a-vis* its dissemination to the international community, necessarily have a standard of scientific publications and citations different from areas exclusively or predominantly academic and they are best captured by Google Scholar.

DESCRIPTORS: Public Health. Scientific Publication Indicators. Bibliometric Indicators. Databases, Bibliographic. Bibliometrics.

^I Instituto de Saúde Coletiva. Universidade Federal da Bahia. Salvador, BA, Brasil

^{II} Instituto Nacional de Ciência e Tecnologia. Ciência, Inovação e Tecnologia em Saúde Salvador, BA, Brasil

^{III} Centro de Pesquisa Gonçalo Muniz. Fundação Oswaldo Cruz. Salvador, BA, Brasil

^{IV} Departamento de Medicina Social. Faculdade de Ciências Médicas da Santa Casa de São Paulo. São Paulo, SP, Brasil

Correspondence:

Maurício Lima Barreto
Instituto de Saúde Coletiva – UFBA
Rua Basílio da Cunha, s/nº Canela
40110-040 Salvador, BA, Brasil
E-mail: mauricio@ufba.br

Received: 9/1/2012
Approved: 2/21/2013

INTRODUCTION

Growth in scientific production, its importance in economic and social development and the consequent consolidation of science as a public policy object in the second half of the 20th century brought with it the need to develop indicators capable of measuring and evaluating the performance of complex scientific activities in general, and of its components – researchers and institutions – in particular. In spite of their recognized limitations, bibliometric indicators are the most widely used to evaluate scientific activity and its influence and impact.¹³ Bibliometric measurements capable of measuring and qualifying scientific productivity are developed to report the performance of researchers, groups and research institutions and to guide the promotion of scientists, fostering research and personnel training. Bibliometrics is a vast field of empirical study and one of the bases of scientometrics.

Bibliometric measurements have evolved over time. Initially, they were limited to counting the number of publications. However, in a short time, the number of publications became increasingly less relevant in qualifying the productivity of a researcher if it were not related to some measure of quality, expressed by peer recognition. This quality was translated into bibliometrics through the number of citations obtained in scientific publications. However, using the gross number of citations as a measure of the influence of a publication had its limitations and was not always a reflection of quality.¹³ A series of indices have been suggested for substituting this method, based on citations. The h-index is the most popular.

This index was developed by a physicist interested in producing a measure which, based on citations, reduces the shortcomings related to simply counting them and overcomes the problems of the denominators used in calculating the impact factor. Hirsh^{8,9} (2005, 2007) suggested that the h-index was better than other indices used up until now – total number of articles, total number of citations, mean number of citations, number of ‘significant’ publications –, as it combines the number of citations with the number of citations of commonly cited articles.

The h-index became renowned due to the possibility of using one single measure, which is calculated in a particularly simple way, to characterize the impact of a researcher’s scientific output. It is calculated based on the descending order of the number of citations of each piece of work by the author (or research group, journal, institution), the h-index being defined as the point at which the number of citations correspond to the number of the order. A researcher who has published 50 articles, of which 22 received 22 or more citations, would have an h-index = 22. It is a robust index, as it combines quantity of scientific production (number of

publications) and aspects of their quality or relevance (citations).⁸ As it has become one of the most commonly used indicators to evaluate scientific production, it also has become the object of serious debate on aspects related to bibliometric measures in science.⁴

There is transparency on h-index variability compared between different scientific areas. Areas with more prolific numbers of publications registered higher h-indices. There is variability when different bibliographic bases or search engines are used to derive the indices.^{12,14} This is because these bases differ with regards to the coverage of their bibliographic or citation records.¹ For the areas of Social and Human sciences, some databases are less representative as fewer books, reports or conference proceedings are indexed.¹⁵ Therefore, the choice of database used to calculate the h-index directly influences the values found.

Two bibliographic databases stand out for their wide ranging coverage of scientific areas and for counting citations: the ISI Web of Science (WoS), with bibliographic records dating back to 1945; and Scopus, created more recently to compete with the former, with records dating from 1960 and, in a more systematic way, from 1996. Recently, Google Scholar (GS) has gained importance, although it is not a bibliographic database like the two former. It is a search engine which uses algorithms to identify scientific publications and their citations available on the internet. This characteristic means that Google Scholar embraces a greater diversity of bibliographic productions, including books, seminars, lectures and others. In this article, the term bibliographic database will be used to refer to all three sources, including GS.

This article aims to analyze the use of h-index as a measure of bibliographic impact of the scientific output of Brazilian researchers. The perspective is to call attention to the use of databases appropriate to the specifics of each field of knowledge, highlighting the particularities of Public Health.

METHODS

Three areas were selected for the analysis: Public Health, Immunology and Medicine, which form part of the so-called life sciences, which includes basic sciences, such as Biology, and applied sciences, such as agricultural and health sciences. Immunology was chosen as it is one of the subareas of biological science in which the greatest impact factors of researchers and the journals used for divulging their output are observed. Medicine was selected because it constitutes a range of the areas of health sciences which have the highest numbers of articles and a high impact factor.

Public Health showed the greatest internal diversity, including researchers in the subareas of epidemiology, social sciences in health and politics and health care management, as well as having a smaller scientific community than Medicine. In spite of this, it stands out with regards to scientific output in Brazil.

Only those who received productivity grant 1-A from the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq – National Council for Scientific and Technological Development) were included in the analysis. The CNPq Productivity in Research grant is aimed at researchers who stand out amongst their peers, valuing scientific output according to normative criteria. Those who receive a productivity grant can be divided into three categories: 2, 1 (D, C, B and A) and senior.^a

In order to apply for a category 2 grant, the researcher should have completed a doctorate at least three years beforehand, or eight years for a category 1 grant. The criteria used to judge grant applications are: (a) the candidate's scientific production; (b) human resources training at a Post-Graduate level; (c) scientific and technological contribution and contribution to innovation; (d) coordination or participation as a lead researcher in research projects; and (e) participation in editorial and scientific management activities as well as administration of institutions and scientific and technological centers of excellence.

To qualify for a senior grant, the researcher should have completed 15 years, which may or may not be consecutive, with a category 1 grant. In category 1, the researcher falls into one of the four subcategories (A, B, C or D), based on their performance in the last ten years and compared with the performance of their peers. For category 2, their productivity is evaluated on one level only, with emphasis on published work and guidance, both referring to the preceding five years.

Category 1-A, the top of the hierarchy, contains researchers who show continued excellence in scientific production, in human resources training and who lead consolidated groups of researchers. The choice of this group restricted comparison to researchers with high levels of productivity and scientific leadership in each of the three areas in question. The sample grouped together researchers at the same stage of their career, as length of time spend in the profession strongly influences the h-index.

The list of researchers was obtained from the CNPq Carlos Chagas platform in April 2011. It included 98 researchers: 20 from Public Health, 59 from Medicine and 19 from Immunology. Access to the WoS and Scopus databases was obtained through the Capes

Journal Portal. GS was accessed using the Publish-or-Perish interface, free access software which organizes searches and calculates h-index.^b

Searches were conducted in each database using the field "name in bibliographic citations" of the selected authors' C.Vs. on the Lattes platform (CNPq). The areas of research and institution of each researcher were checked. Documents found in the bibliographic databases were compared to those referred to in each authors' C.Vs. on the Lattes platform. This check enabled publications by authors with the same name to be excluded, as this could have distorted the results.

The h-indices for each researcher were estimated using the three databases, considering the period covered by each. The total indexed production of each author in the different bases was found. Any distortion with regards to the difference time covered is present in all three areas; therefore, it does not affect comparisons of h-index behavior in the three different areas/subareas in the three databases. This process was carried out blind, i.e., without prior knowledge of the researchers' area, which avoided any bias on the part of the authors of this study.

Means and medians were obtained for each group of researchers by area and by origin of the h-indices estimates. To test the statistical significance of differences between the groups, the Kruskal-Wallis test for two or more groups, equivalent to non-parametric variance analysis, was used to compare the three areas in question using the same database, or the three databases for each area. In those cases in which the result of the Kruskal-Wallis test was significant ($p < 0.05$), the Behrens-Fisher multiple comparisons test was used, which tests the groups two by two, to ascertain in which groups the difference occurred.⁵

RESULTS

The data referring to the Public Health researchers showed a greater range of variation, with more extreme minimum and maximum values (Table 1). In the graphics a, b and c of the Figure, the distribution of the h-index values and their respective medians for the researchers in each area are shown in more detail, generated for the three different sources used in the study (WoS, Scopus and GS).

When GS was used, the researchers from the areas of Public Health and Medicine had h-index medians significantly higher than when using the other two databases (Tables 2 and 3). The three areas did not differ with regards to the h-indices obtained using Scopus and GS, although Immunology had significantly higher

^a Behrens-Fisher test a Ministério de Ciência e Tecnologia. Conselho Nacional de Desenvolvimento Científico e Tecnológico. Resolução Normativa nº 9, de 24 de abril de 2009. Brasília (DF); 2009.

^b Harzing AW. Publish or Perish. Version 3.0.3813. Londres; 2010 [cited 2013 Jan]. Available from: www.harzing.com/pop.htm

Table 1. Mean and median of h-indices for CNPq 1-A researchers in the areas of Public Health, Medicine and Immunology estimated from different sources.

| | Web of Science | Scopus | Google Scholar |
|----------------------|----------------|--------|----------------|
| Public Health | | | |
| Mean | 19.4 | 20.4 | 31.0 |
| Standard Deviation | 14.1 | 10.1 | 14.6 |
| Median | 17 | 19 | 28 |
| Minimum | 4 | 2 | 14 |
| Maximum | 62 | 49 | 68 |
| Medicine | | | |
| Mean | 20.6 | 20.5 | 25.2 |
| Standard Deviation | 7.4 | 6.4 | 10.2 |
| Median | 18 | 20 | 24.5 |
| Minimum | 6 | 10 | 6 |
| Maximum | 39 | 39 | 57 |
| Immunology | | | |
| Mean | 28.4 | 24.0 | 31.4 |
| Standard Deviation | 9.1 | 6.6 | 12.8 |
| Median | 27 | 24 | 27 |
| Minimum | 17 | 12 | 18 |
| Maximum | 50 | 40 | 59 |

medians than the other two areas when using WoS. This difference disappeared when the comparison was made using the Scopus and GS databases.

Among the researchers from Medicine and Public Health, there was a significant increase in the h-indices medians when GS was used compared with the other two databases (23% and 475 higher, respectively).

Table 2. Median of h-indices for CNPq 1-A researchers in the areas of Public Health, Medicine and Immunology estimated from different sources (Web of Science, Scopus and Google Scholar) and (A) p-values from comparing the different sources and (B) p-values from comparing the different areas for each source.

| A | | | B | | |
|----------------------|------|----------------|-----------------------|--------|----------------|
| Source | Area | p ^a | Area | Source | p ^a |
| Public Health | | | Web of Science | | |
| Web of Science | 17.0 | 0.008 | Public Health | 17.0 | 0.001 |
| Scopus | 19.0 | | Medicine | 18.0 | |
| Google Scholar | 28.0 | | Immunology | 27.0 | |
| Medicine | | | Scopus | | |
| Web of Science | M | 0.001 | Public Health | 19.0 | 0.157 |
| Scopus | 20.0 | | Medicine | 20.0 | |
| Google Scholar | 24.5 | | Immunology | 24.0 | |
| Immunology | | | Google Scholar | | |
| Web of Science | 27.0 | 0.145 | Public Health | 28.0 | 0.120 |
| Scopus | 24.0 | | Medicine | 24.5 | |
| Google Scholar | 27.0 | | Immunology | 27.0 | |

^a Kruskal-Wallis test

DISCUSSION

Using GS generates higher h-indices for public Health researchers, followed by those in Medicine, but not for Immunology, compared with h-indices generated by WoS and Scopus. Immunology had significantly higher medians than the other two areas in WoS, probably due to it being a basic science, the articles of which are traditionally published in English and much cited among peers. According to Journal Citation Reports (JCR, da ISI Web of Knowledge), in 2011 there were 139 journals indexed for Immunology and 234 for Public Health, considering the social sciences collection. In spite of this, the citation indices are higher for Immunology.

Results of this nature are relevant, as deciding which base to use for the h-index calculation has implications on the ranking of the researchers and the academic areas.

Researchers in Immunology had similar h-index median values calculated in the three different databases, whereas there were significant differences in those of researchers in Medicine and Public Health estimated using GS compared with WoS and Scopus. This difference was greater than 50% in the case of Public Health. The area of health care, in general, has a greater number of journals published in Brazil which are not indexed or have only recently been indexed in the Scopus and WoS databases.

Using a different approach, Pereira & Bronhara¹⁶ (2011) estimated the h-indices for all active Brazilian lecturers in Post-Graduate Public Health programs in 2009. They used WoS and found a national mean h-index of 3.1, with 29.8% of the lecturers obtaining h-indices of zero.

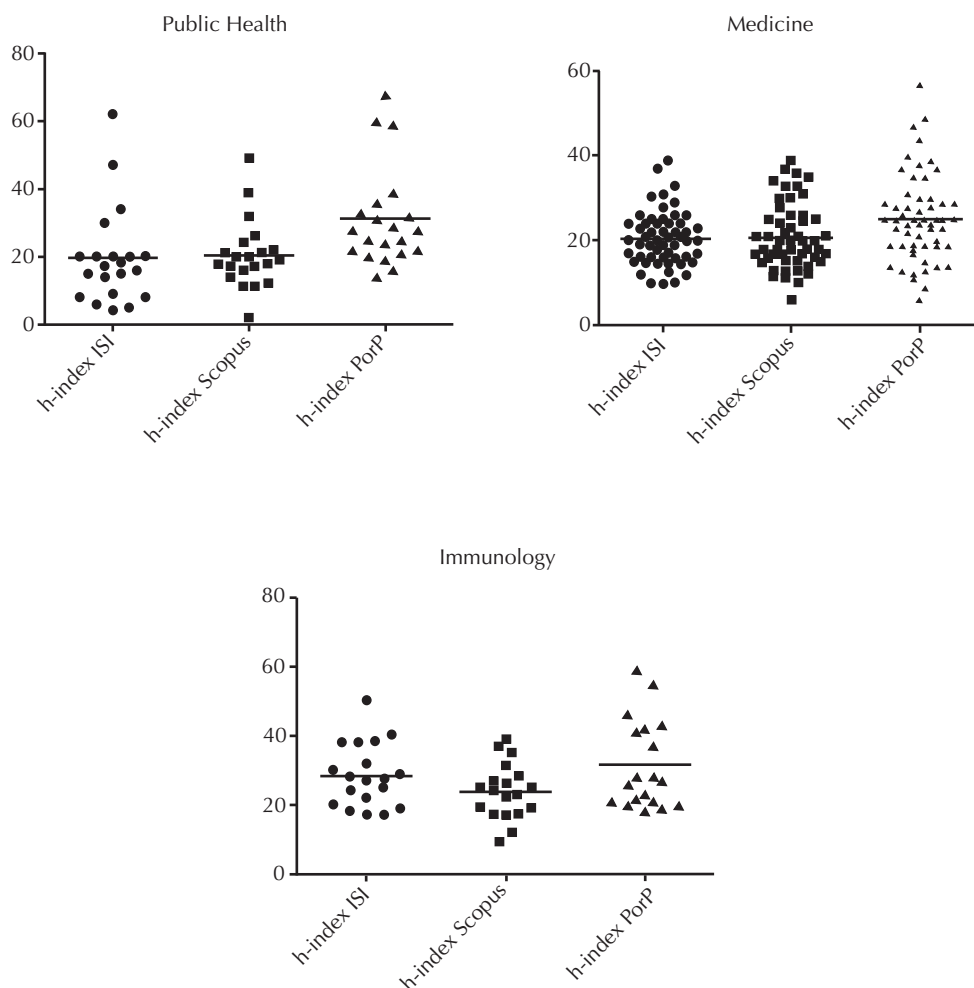


Figure. Estimated h-indices and medians in the Web of Science, Scopus and Google Scholar (calculated by Publish or Perish) databases for 1-A researchers from the National Council for Research and Technological Development (CNPq) in the areas of Public Health, Medicine and Immunology.

Using GS instead of WoS produces different results. The majority of 1-A researchers in Public Health are epidemiologists, with a publication pattern closer to that of natural or exact sciences, such as Immunology; therefore, they are well represented in the WoS. In public Health, there are lecturers of the subareas of social sciences in health care and politics and health care management, whose publishing and citation profiles are closer to those of the social sciences in general, which has a lower coverage of indexation and registered citation in WoS than in GS. This occurs because a significant part of the publication in social science is in the form of books or other types of documents, captured by GS, but not indexed in Scopus or WoS.

The emergence of GS and interfaces which maximize its use has brought a whole new level to discussions of measures of bibliographic impact of scientific

publications. Areas with strong professional components, in which knowledge produced is also published, as it should be, in the native language, *vis-à-vis* its dissemination to the international scientific community, have a different pattern of publication and citation to those areas which are exclusively or predominantly academic.

The differences found in the h-indices from the three sources are related to the particular characteristics of these sources. WoS belongs to Thomson Reuters, the third largest publishing company in the world, which charges for access. It is the most traditional of the bases, and the most commonly used, including in research institutions in Brazil. Scopus, developed by Elsevier, the second largest publishing company in the world, is private and charges for access. In spite of being more recent, it is a strong competitor with WoS. It has a wider coverage of scientific journals,

Table 3. Results of multiple comparison tests: comparison between sources for each area and comparison between areas for each source.

| Comparison between sources for each area | p ^a | Comparison between areas for each source | p ^a |
|---|----------------|--|----------------|
| Public Health | | Web of Science | |
| Scopus <i>versus</i> Web of Science | 0.603 | Public Health <i>versus</i> Immunology | 0.004 |
| Scopus <i>versus</i> Google Scholar | 0.003 | Public Health <i>versus</i> Medicine | 0.404 |
| Web of Science <i>versus</i> Google Scholar | 0.002 | Immunology <i>versus</i> Medicine | 0.001 |
| Medicine | | Scopus | |
| Scopus <i>versus</i> Web of Science | 0.996 | Public Health <i>versus</i> Immunology | 0.240 |
| Scopus <i>versus</i> Google Scholar | 0.013 | Public Health <i>versus</i> Medicine | 0.907 |
| Web of Science <i>versus</i> Google Scholar | 0.010 | Immunology <i>versus</i> Medicine | 0.178 |
| Immunology | | Google Scholar | |
| Scopus <i>versus</i> Web of Science | 0.267 | Public Health <i>versus</i> Immunology | 0.999 |
| Scopus <i>versus</i> Google Scholar | 0.158 | Public Health <i>versus</i> Medicine | 0.284 |
| Web of Science <i>versus</i> Google Scholar | 0.878 | Immunology <i>versus</i> Medicine | 0.234 |

^a Behrens-Fisher test

including those not in English and published outside of the North American and Western Europe axis. According to information available in the respective portals, in 2012 the Scopus database had 19,500 journals, more than 240 of which were Brazilian. In WoS, of the more than 12 thousand journals indexed, around 240 were Brazilian. GS, in turn, is part of the most popular search engine on the internet.

One of the first authors to compare these three sources highlighted differences between them which need to be better explained.¹ Whereas Bar-Ilan¹ found both higher and lower variations among h-indices estimated by GS compared with the other two sources,¹ in this study, GS systematically generated higher results, with statistically significant differences in the cases of Public Health and Medicine.

These divergences may be due to the fact that, in the beginning, the indices were estimated directly from GS, without the possibility of excluding citations not referring to the articles, or by authors with the same names. Therefore, the creation of *Publish-or-Perish* and, more recently, a new interface developed by Google itself (the *my citations* command in GS) has improved the system, reducing inconsistencies. The current differences in results between h-indices generated in GS compared with the other two databases may better reflect the real differences between the numbers of publications and citations.

The researcher who developed the *Publish-or-Perish* software called attention to GS's superiority in estimating h-indices, especially for researchers in applied and social and human sciences, whose scientific journals are not well covered in the other two databases used.⁷ An advantage of GS is that it does not depend on closed commercial databases. As it indexes references

and citations available on the internet, GS is open and allows access to a large database which is not indexed in Scopus or WoS.

Search engines also have disadvantages. There is a greater degree of 'rubbish' in the data obtained, i.e., the inclusion of publications and citations of non-scientific articles, which brings limitations and means that more care is needed when using it and constructing indices based on it, which means that GS has both supporters and detractors.^{6,10} The fact that estimating the h-index depends on the number of publications which is a small percentage of the total publications of an active researcher means that it is easier to verify the publications and respective citations which are entered into the index calculation, excluding incorrect mentions.

Knowing the advantages and drawbacks of each of the sources used allows a more productive use of such bibliometric indicators, better exploiting the potential of each. This may avoid their use as a form of academic control or the creation of (false) hierarchies of researchers and research institutions.

Variation in h-index according to the bibliographic source or the search engine used is not an inherent disadvantage to this measure. However, there are drawbacks in the index itself which have been highlighted by various authors. There are two types of criticism of the use of this indicator: one, of a more general character, is related to the use of indices based on citations as a measure of scientific impact; the other is related to its specificities.

More general criticism states that using citations may be affected by various factors – social, political or geographic – and contests the relationship between 'popularity' for generating a high number of citations and the

transparent expression of effective scientific quality.¹³ From a more radical perspective, some argue that the use of scientific measures, especially bibliometrics, is part of a greater project aiming to impose 'quantified control' on academic activities.³

More specific criticism of the h-index highlights the fact that it is dependent on time. It is cumulative, and is related to the number of citations, but also the number of publications. An author with ten publications with thousands of citations will never have a h-index higher than 10. This aspect is important, and Hirsch⁹ (2007) himself, in the original article, stipulated that the index would serve for evaluating researchers at the same stage in their careers. The h-index is useful for making comparisons between the more productive scientists, who generally have been active in their field for a greater length of time, which justifies the choice of CNPq 1-A researchers.² Another disadvantage of

the h-index refers to the fact it can be manipulated by self-citation or other mechanisms.

Another relevant aspect that should be mentioned is the variation in h-index between scientific areas. A comparison of h-indices of members of the ten scientific areas of the *Academia de Brasileira de Ciências* (Brazilian Academy of Sciences) shows that higher h-index means were calculated using WoS in the areas of Biomedicine, Health and Chemistry (23, 20 and 19, respectively), lower means were found in Earth Sciences, Engineering and Mathematics (9, 8 and 7, respectively) and means of practically zero in the Human Sciences (1).¹¹

Using the appropriate database for each field of knowledge is critical. This enables a more robust use of the h-index for reporting performance of researchers, groups and research institutions and of promoting scientists, fostering research and training personnel.

REFERENCES

1. Bar-Ilan J. Which h-index? A comparison of WoS, Scopus and Google Scholar. *Sciencimetrics*. 2008;74(2):257-71. DOI: <http://dx.doi.org/10.1007/s11192-008-0216-y>
2. Bornmann L, Hans-Dieter D. What Do We Know About the h Index? *J Am Soc Inf Sci Technol*. 2007;58(9):1381-5. DOI: <http://dx.doi.org/10.1002/asi.20609>
3. Burrows R. Living with the h-index? Metric assemblages in the contemporary academy. *Sociol Rev*. 2012;60(2):355-72. DOI: <http://dx.doi.org/10.1111/j.1467-954X.2012.02077.x>
4. Cronin B, Meho L I. Using the h-index to rank influential information scientists. *J Am Soc Inf Sci Technol*. 2006;57(9):1275-8. DOI: <http://dx.doi.org/10.1002/asi.20354>
5. Dalgaard P. *Introductory Statistics with R*. 2 ed. New York: Springer Science; 2008.
6. Harzing AWK, Van-der-Wal R. Google Scholar: the democratization of citation analysis? *Ethics Sci Environ Polit*. 2008;1(1):61-73. DOI: <http://dx.doi.org/10.3354/esep00076>
7. Harzing AWK, Van-der-Wal R. A Google Scholar H-Index for journals: an alternative metric to measure journal impact in economics & business? *J Am Soc Inf Sci Technol*. 2009;60(1):41-6. DOI: <http://dx.doi.org/10.1002/asi.20953>
8. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA*. 2005;102(46):16569-72. DOI: <http://dx.doi.org/10.1073/pnas.0507655102>
9. Hirsch JE. Does the h index have predictive power? *Proc Natl Acad Sci USA*. 2007;104(49):19193-8. DOI: <http://dx.doi.org/10.1073/pnas.0707962104>
10. Jacso P. Metadata mega mess in Google Scholar. *Online Inf Rev*. 2009;34(1):175-91. DOI: <http://dx.doi.org/10.1108/14684521011024191>
11. Kellner AW, Ponciano LC. H-index in the Brazilian Academy of Sciences: comments and concerns. *An Acad Bras Cienc*. 2008;80(4):771-81. DOI: <http://dx.doi.org/10.1590/S0001-37652008000400016>
12. Lacasse JR, Hodge DR, Bean KF. Evaluating the productivity of social work scholars using the h-index. *Res Soc Work Pract*. 2011;21(5):599-607. DOI: <http://dx.doi.org/10.1177/1049731511405069>
13. Lindsey D. Using citation counts as a measure of quality in science measuring what's measurable rather than what's valid. *Scientometrics*. 1989;15(3-4):189-203. DOI: <http://dx.doi.org/10.1007/BF02017198>
14. Norris M, Oppenheim C. Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *J Informetr*. 2007;1(2):161-9. DOI: <http://dx.doi.org/10.1016/j.joi.2006.12.001>
15. Ouimet M, Bedard PO, Gelineau F. Are the h-index and some of its alternatives discriminatory of epistemological beliefs and methodological preferences of faculty members? The case of social scientists in Quebec. *Scientometrics*. 2011;88(1):91-106. DOI: <http://dx.doi.org/10.1007/s11192-011-0364-3>
16. Pereira JCR, Bronhara B. H-index of Collective Health professors in Brazil. *Rev Saude Publica*. 2011;45(3):599-606. DOI: <http://dx.doi.org/10.1590/S0034-89102011005000027>