Making Use of H-index: the Shape of Science at the University of Sarajevo

Haris Memisevic[1], Irzada Taljic[1], Amra Macak Hadziomerovic[2]

[1]University of Sarajevo, Faculty of Educational Sciences, Sarajevo, Bosnia and Herzegovina

[2]University of Sarajevo, Faculty of Health Studies, Sarajevo, Bosnia and Herzegovina

Corresponding author: Haris Memisevic, PhD. University of Sarajevo, Faculty of Educational Sciences. 71000 Sarajevo, Bosnia and Herzegovina. Phone: +38761178123. Fax: +38733207963. ORCID ID: http://www. orcid.org/0000-0001-7340-3618. E-mail: hmemisevic@gmail.com

doi: 10.5455/aim.2017.25.187-190

ACTA INFORM MED. 2017 SEP; 25(3): 1187-190

Received: Jul 27, 2017 • Accepted: Sep 05, 2017

ABSTRACT

Introduction: Quantifying science and scientific contribution has become one of the main tasks in evaluating researchers and their impact. How do we value research and science in Bosnia and Herzegovina (BIH)? Scientific community has mostly agreed upon that one of the best ways to value researchers is through their h-index value. However, there are many databases and services from which h-index can be retrieved. **Aim:** To describe different databases and services such as Google Scholar, Web of Science, Scopus and Researchgate in evaluating the researcher. An additional aim of this paper is to present "the shape" of science at the University of Sarajevo and to examine what are the best predictors of h-index. **Materials and methods:** We analyzed the data from 100 Google Scholar Profiles of researchers from University of Sarajevo. **Results:** The study showed some benefits and shortcomings of mentioned databases and services. Most researchers in the sample were from natural sciences, in particular from the field of medicine. The mean value of h-index in relation to the researcher's gender was not statistically significant. We conclude the article with some ideas on how to improve the visibility of researchers from BIH.

Keywords: h-index, databases, Google Scholar, University of Sarajevo.

1. INTRODUCTION

Publish or perish. Most researchers are reminded of this hamletian dilemma every time a promotion at the department or a national project application is supposed to take place. Researchers are aware that without valuable scientific output the chances for tenure, promotion and "academic fame" are significantly reduced. Thus, the need to publish in journals that are scientifically visible is becoming a must for a scientist. Although, the practice of quantifying researcher's influence has a long tradition in developed countries, it is increasingly being used in countries on a scientific periphery such as Bosnia and Herzegovina (BIH). Scientific periphery is the term used by Marusic and Marusic (1) for countries formed after the fall of communism but also many other countries that lag behind in research production. The task of quantifying and valuing research output has become a serious endeavor being imposed by funding agencies, promotion committees and employers (2). Measuring and quantifying academic accomplishments has become a "fact of scientific life" (3). Currently there is no universally accepted, international standard procedure for evaluating researchers based on their scientific output (4).

However, there are some methods for measuring scientific impact such as the number of citations that a publication receives (5). One of the most cited measures, although not without flaws, in measuring scientific output is h-index. Hirsch (6) proposed a single number that will be useful in measuring scientific output of a researcher. H-index, as proposed by Hirsch, is the number of publications that have that many number of citations. So, we can see that this metrics is influenced by both, researcher's productivity and the impact that research had on his/her peers. The h-index has achieved significant acceptance as a measure of individual research achievement and it has several advantages as it combines scientific productivity with scientific echo and is not susceptible to extreme values (7). It is important to note that h-index is not without its critics and numerous other indexes have been proposed in-

stead such h (bar) index (8). However, for the purposes of this study we will stick to the h-index. Studies have shown that researchers with higher h-index are generally more likely to obtain grants and fellowships (9).

Scientific community in BIH has slowly begun to accept h index as an indicator of researcher's output. However, which h index will be used remains a dilemma. For example, for promotion purposes should the promotion committee look at h index from Google Scholar, Researchgate, Web of Science or Scopus?

Google Scholar and Researchgate are free-of-charge services that look at researchers output and can be used in analyzing citation index of an author. Google Scholar is probably the largest scientific database that combines content from several large databases unavailable to the public web, plus academic web documents from the Google search engine (10). Google Scholar covers a large number of journals and other publications. Researchers have the option to create their Google Scholar Profile, which has all the papers that the researcher authored, which can be automatically retrieved by Google or can be manually added by the researcher. The papers can be manually added and show to the public research that is not covered by Google search engines. However, the citations cannot be manually added and they are retrieved by Google. So, the Google Scholar automatically counts all the citations it can find for the author and creates their h-index. However, although very useful, the accuracy of the Google Scholar services is often inconsistent (11).

On the other hand, Researchgate (RG) is a social network for researchers in which they can list their publications and interact with their peers (12). Benefits of RG is that it is quite easy to access and use and has the potential to become the library of the 21st century (13). RG creates a list of citation based on the papers that are uploaded on its site. Researchers registered on RG can also add their papers manually or RG can find papers belonging to them. Just like in Google Scholar, RG automatically finds citations for the papers that are uploaded on its platform. However, many papers citing someone's work have not been uploaded to the platform and thus the number of potential citation is probably reduced.

Let us now point to the potential academic misuses of the Google Scholar and Researchgate services. We already mentioned that one can manually add papers to Google Scholar Profile. Potential misuse of this option is when the authors add papers that are not theirs. For example, in the first 120 ranked researchers affiliated with the University of Sarajevo, profiles of 8 authors (7%) contain papers that they did not write or co-authored but are in their Scholar Profile. Very often these papers are highly cited and Google adds these citations to the author's Scholar Profile. It is not clear whether the authors do this inadvertently or deliberately but in such a way their h-index is artificially increased. This is one of the potential misuses of Google Scholar service.

The similar misuses can happen in RG as well. The paper can be added that does not belong to the author and the impact (or RG score) of the author will be increased. Or the author can decide to share his/her work on RG that has not been published before. It is perfectly legitimate to do this, but the author can use this option dishonestly and in that way increase their number of citations and h-index.

Let us now turn to the commercially available databases such as Web of Science and Scopus. Web of Science covers journals published from 1900 to present and Scopus covers journals covering the period from 1966 to present (11). Although they are only commercially available, they do not suffer from some of the shortcomings mentioned earlier for RG and Google Scholar. Authors cannot add papers manually to WOS and Scopus nor can they manipulate the numbers of citations and h-index as they can in Scholar and RG. Therefore, WOS and Scopus are more accurate but there can be a long time lag between publication of the paper and its visibility on WOS and Scopus databases.

Discussion of which of the h-indexes can be used in evaluation of the researcher probably left us with more dilemmas than definitive answers but it also pointed to certain areas in which RG and Google Scholar can be improved. Is it better to use Google Scholar h-index or WOS h-index? The answer is: It depends. It would probably be the best to evaluate as many of these indices as possible, and if large discrepancies exist, that can be a reason for more in-depth analysis.

The second part of this paper deals with the "shape" of scientific output of the researchers from University of Sarajevo, BiH. University of Sarajevo is a state-funded University and is, by large margin, the best University in Bosnia and Herzegovina. According to Webometrics, University of Sarajevo is ranked on 1881st place in the world rank, and second and third best Universities from BiH are ranked on 3655 and 3660 positions respectively (14).

2. AIM

The specific aims in this study are:

- To examine gender distribution of 100 researchers from University of Sarajevo;
- To investigate what are the scientific fields that the researchers are coming from;
- To determine what are the best predictors of h-index.

3. MATERIALS AND METHODS

For the purposes of this analysis, we analyzed the scientific output of 100 researchers from University of Sarajevo ranked between 11th and 110th position on the Google Scholar. The researchers are ranked according to their number of citations. We did not present or analyzed the scientific output of the first 10 ranked researchers separately, as they present significant outliers. So, we picked researchers who had less than 1000 citations in their Google Scholar Profile on the date of 24.07.2017., and the first researcher who had less than 1000 citations was ranked 11th on Google Scholar list.

4. RESULTS

In relation to the gender, among the researchers from 11th and 110th position, there were 51 females (51%) and 49 males (49%). As a note let us mention that among the first 10 ranked researchers, according to Google Scholar Profiles for University of Sarajevo, there were 9 males and 1 female.

In relation to the total number of citations, there were no statistically significant differences between males and females. The numbers are presented in Figure 1.

The mean number of citations for females was 255.6 (SD=137.9) and for males it was 292.1 (SD=183.6). According

to the t-test results, there were no statistically significant differences in the mean numbers of citations for males and females (t(1,99)=1.1; p=.27) in the sample.

Next, we wanted to examine from which scientific fields these researchers are coming from. These results are shown in Table 1.

Field	Count	%
biology	5	5
chemistry	6	6
electronics and IT	15	15
genetics	7	7
mathematics	4	4
medicine	30	30
other	7	7
pharmacy	8	8
physics	6	6
social_sci	12	12
Total	100	100

Table 1. Scientific field of the researchers from University of Sarajevo ranked between 11th and 110th place on Google Scholar

As can be seen from the table, majority of researchers come from natural sciences, and only 12% come from social sciences. Most researchers are in the field of medicine, followed by merged category of electronics and IT.

The final goal of this paper was to examine what are the best predictors of h-index. As potential predictors, we put following variables in the model: gender, scientific field, h10 index, total number of citations, number of citations in 2016 and number of citations in 2017. The results are shown in Table 2.

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3,420	0,254	13,45	<,0001*
total_cites	0,001	0,001	0,83	0,4078
field[biology]	-0,864	0,419	-2,06	0,0421*
field[chemistry]	0,107	0,379	0,28	0,7792
field[electronics_IT]	0,642	0,260	2,46	0,0157*
field[genetics]	-0,441	0,357	-1,23	0,2207
field[math]	-0,272	0,455	-0,60	0,5508
field[medicine]	0,175	0,199	0,88	0,3820
field[other]	0,525	0,394	1,33	0,1863
field[pharmacy]	-0,013	0,343	-0,04	0,9698
field[physics]	0,055	0,433	0,13	0,8999
gender[f]	0,019	0,108	0,18	0,8604
h10_index	0,631	0,038	16,74	<,0001*
n_2016	-0,007	0,011	-0,62	0,5387
n_2017	-0,001	0,013	-0,06	0,9503

Table 2. Regression analysis predicting the h-index of the researcher

As can be seen from the table, the significant predictors of h-index are h10-index and the scientific fields of biology and electronics (IT sector). To put the above results in the perspective, if the researchers are from the field of biology, they can expect to have h-index smaller than researchers from other field (all other parameters being equal). In the same line, researchers from the field of electronics (IT) have higher indexes than researchers in other fields (keeping all other parameters equal). For example, a male researcher in biology with an h10 index of 5 is predicted to have a mean h-index of 5.7, while a male researcher in electronics (IT) with an h10 index of 5 is, on average, is predicted to have a mean h-index of 7.2.

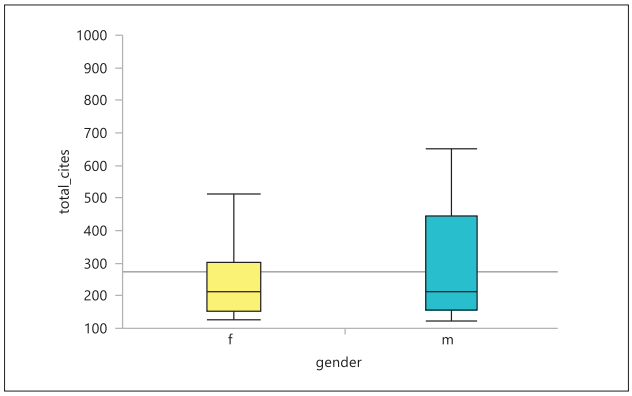

Figure 1. Mean number of citations of males and females

5. DISCUSSION

We described several options for using h index, including two, free-of-charge, services and two commercially available databases. It is obvious that there is still room for improvement of these services, especially in Google Scholar and RG services. The major objection for these services is their inability to recognize and prevent potentially dishonest behavior of the researchers in terms of authorship of publications. We could see that at the University of Sarajevo, there were 7% of researchers among the first 120 ranked researchers who had (deliberately or inadvertently) publications at their Google Scholar Profile that do not belong to them in the sense of authorship. This is probably the case only among the top 200 ranked researchers, and not among the weaker positioned researchers. Therefore, due to its accuracy, Web of Science and Scopus should be the primary source of evaluating someone's scientific impact. Google Scholar should definitely complement other bases such as WOS and Scopus in evaluating researcher's performance. On the other hand, RG should be limited to serve as a social network between researchers but not as a means to evaluate somebody's scientific contribution. RG allows researchers to pose questions and attract many scientists who will help them in their scientific efforts and dilemmas. However, RG scores can be artificially inflated and thus should not be used in evaluating someone's scientific impact. Some current research has already pointed to the issue of ghost academic reputation in RG scores (15).

Regarding the shape of science at the University of Sarajevo, it is a very positive trend that there is no gap in the scientific achievement in relation to the researcher's gender, at least for the researchers who are not in the top 10 list. In relation to the scientific field, we could see that majority of these 100 researchers are from the field of natural sciences, in particular from the field of medicine. As there were very few researchers from some scientific disciplines, they were all merged in a single category of Social sciences, including humanities as well. Generally, the case of uncited documents in more prevalent in certain disciplines more than in other. For example, engineering and social sciences disciplines tend to be under cited (16).

The last research question in this study dealt with predictors of h-index. It is interesting that the total number of citations was not a significant predictor of h index. It was revealed that h10 index, that is number of publications that have more than 10 citations is almost perfectly correlated with h index

and thus could even be used as a proxy for h index. Gender did not have a predictive role in determining the h-index but some scientific disciplines (biology and electronics IT) were significant contributors to h index. We should immediately note the limitation of the sample size for this regression analysis. Future studies with larger sample sizes should confirm or refute the results of this regression. Another limitation is that we did not analyze the role of other potentially important demographic predictors of h-index such as age, family status etc. Of course, due to the short nature of this paper we did not analyze other, potentially equally important, measures of scientific impact such as g-index, impact factors, eigenfactors etc.

University of Sarajevo, as the largest and best rated University in BiH has only 10 scientists who have more than 1000 citations. Situation at other BiH Universities is even bleaker. Universities of Banja Luka, Tuzla, Mostar and Zenica do not have registered researchers who have more than 1000 citations. As a comparison, we will take University of Split, which, on the same date (24.07.2017), had 21 researchers with more than 1000 citations. It is obvious that state authorities in BiH, together with academic community have to do much more for the growth of science, scientific impact and visibility of universities in Bosnia and Herzegovina. One of the ways that would certainly increase visibility is for individual researchers to use social media to the maximum extent possible (17). Sustainable economic growth is heavily dependent on science and generation of knowledge (18). Investing in research universities is considered an investment in the central institutions of 21st century knowledge economies (19). Thus, the authorities in BIH need to understand this and start to invest more financial resources in science and support to scientists in BiH. Only by investments in science can BiH catch up with its European neighbors.

6. CONCLUSION

H-indexes retrieved from databases and services such as Researchgate, Google Scholar, Web of Science and Scopus can be used to assess researcher's scientific impact. Web of Science and Scopus, as commercially available databases, are more accurate than Google Scholar and Researchgate. However, Google Scholar and Researchgate can be used as a complement to provide a more detailed insight into researcher's scientific output. All academic staff at the state-funded universities should have Google Scholar Profile and Researchgate profile. This in turn will lead to better visibility of researchers from BIH and will have a positive impact on the ratings of Universities from BIH.

- All authors contributed equally in preparation of this manuscript. All authors approved the final version of this paper.
- The authors report no conflict of interest.

REFERENCES

1. Marusic A, Marusic M. Small scientific journals from small countries: breaking from a vicious circle of inadequacy. Croat Med J. 1999 Dec; 40: 508-14.
2. Kelly CD, Jennions MD. The h index and career assessment by numbers. Trends Ecol Evol. 2006 Apr 30; 21(4): 167-70. doi: 10.1016/j.tree.2006.01.005.
3. Lane J. Let's make science metrics more scientific. Nature. 2010 Mar 25; 464(7288): 488-9. doi: 10.1038/464488a.
4. Memisevic H. Izbor časopisa za objavljivanje naučnih radova. Učenje i nastava. 2016; 2(1): 181-90.
5. Bollen J, Van de Sompel H, Hagberg A, Chute R. A principal component analysis of 39 scientific impact measures. PloS one. 2009 Jun 29; 4(6): e6022. doi: 10.1371/journal.pone.0006022.
6. Hirsch J E. An index to quantify an individual's scientific research output. 2005; [Accessed 23/07/17]. Available at: https://arxiv.org/abs/physics/0508025
7. Masic I, Begic E. Scientometric Dilemma: Is H-index Adequate for Scientific Validity of Academic's Work?. Acta Inform Med. 2016 Jul 16; 24(4): 228-32. doi: 10.5455/aim.2016.24.228-232.
8. Hirsch JE. An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. Scientometrics. 2010 Dec 1; 85(3): 741-54. doi: 10.1007/s11192-010-0193-9.
9. Bornmann L, Daniel HD. Does the h-index for ranking of scientists really work?. Scientometrics. 2005 Dec 8; 65(3): 391-2. doi: 10.1007/s11192-005-0281-4.
10. Aguillo IF. Is Google Scholar useful for bibliometrics? A webometric analysis. Scientometrics. 2012 May 1; 91(2): 343-51. doi: 10.1007/s11192-011-0582-8.
11. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. The FASEB journal. 2008 Feb 1; 22(2): 338-42. doi: 10.1096/fj.07-9492LSF.
12. Thelwall M, Kousha K. ResearchGate: Disseminating, communicating, and measuring Scholarship? J Assoc Inf Sci Technol. 2015 May 1; 66(5): 876-89. doi: 10.1002/asi.23236.
13. Masic I, Begic E. Biomedical Scientific and Professional Social Networks in the Service of the Development of Modern Scientific Publishing. Acta Inform Med. 2016 Dec; 24(6): 409-412. doi: 10.5455/aim.2016.24.409-412.
14. Webometrics. Ranking web of Universities. 2017. [Accessed 23/07/17]. Available at: http://www.webometrics.info/en/Ranking_Europe/Central_Eastern_Europe?page=2
15. Orduna-Malea E, Martín-Martín A, Thelwall M, López-Cózar ED. Do ResearchGate Scores create ghost academic reputations?. Scientometrics. 2017 Jul 1: 1-8. doi: 10.1007/s11192-017-2396-9.
16. Hamilton DP. Research papers: who's uncited now?. Science. 1991 Jan 4; 251(4989): 25-6. doi: 10.1126/science.1986409.
17. Trajkovski VE. The role of article level metrics in scientific publishing. Journal of Special Education and Rehabilitation. 2016 Mar 28; 17(1-2): 5-15. doi: 10.19057/jser.2016.0.
18. King DA. The scientific impact of nations. Nature. 2004 Jul 15; 430(6997): 311-6. doi:10.1038/430311a.
19. Altbach P G, Salmi J, Wang Q H, Liu N C, Postiglione G A,...& Materu P. The road to academic excellence: The making of world-class research universities. World Bank. 2011.