Review of: "Measuring researchers' success more fairly: going beyond the H-index"

Lilian Nassi-Calo[1]

1 Pan American Health Organization (PAHO)

Potential competing interests: No potential competing interests to declare.

The author makes an important contribution to the complex topic of research and researchers assessment. By increasing the number of variables that make up a potential new indicator related to the H index, the author points out that the resulting metric would be a fairer metric to researchers assessment, although, as he observes, not a perfect one, because this would be not possible.

I would like to call the attention to some aspects of research evaluation that have not been discussed in the article. For instance, the strong limitations to the Impact Factor – which is often mentioned by the author and used in the calculations of the new metric – pointed out in the San Francisco Declaration on Research Assessment (DORA) [1], proposed in 2012 by the American Society of Cell Biology that have been signed, so far, by over 23,000 individuals from 160 countries. DORA recommends that citation-based metrics such as the Impact Factor are not used to evaluate researchers for hiring, promotion, or in decisions to concede grants for research projects. A similar initiative, the Leiden Manifesto [2], was launched in 2014 in the 19th International Conference on Indicators in Science and Technology, in Leiden, the Netherlands, guiding the use of metrics to evaluate science in Europe. The document has been translated into 25 languages, adopted by institutions, and recognized by publishers around the world.

Since the new metric is intended to be a fairer way to assess researchers, it should consider that impact indexes based on citations, such as SJR, Cite Score or the like, will bear the same limitations as does the IF, pointed out by DORA and the Leiden Manifesto.

The equation proposed in the article is quite efficient to account for the variables that have not been considered by Hirsch in his formulation of the H index and aim at compensate its limitations. However, many authors [3] support the principle that it is quite surprising that even knowing the discrepancies and flaws shown by different indexes from various databases and rankings, the scientific community – and that include funding agencies and research institutions who decide which candidates to hire or promote – still rely on them to inform their decisions.

It is worth mentioning the recent Agreement on Reforming Research Assessment [4] by the European Research Area (ERA), European University Association, Science Europe, and the European Commission. The agreement relies on the support of more than 350 public and private associations from over 40 countries. Signatories to the agreement will commit to a shared vision that research evaluation, researchers and research organizations recognize the diverse outcomes, practices and activities that maximize research quality and impact. This requires grounding the assessment primarily in

qualitative judgment, for which peer review is central, supported by the *responsible use of quantitative indicators*. A short video made available by ERA details the reasons and the overhaul of the research evaluation process, with the viewpoints of interviewed researchers.

To conclude, I would like to quote Olof Hallonsten [5] "*the current ubiquity of performance evaluation in science for the most part is pointless and counterproductive, and that this state of science policy is in dire need of reevaluation in order to secure science's continued productivity and contribution to social and technological innovation […] it is time to stop evaluating it with metrics that obviously fail to make justice to its success, and most of all time to stop governing it on basis of what these metrics show. Either Lord Kelvin (or Peter Drucker, or whoever really said it) was wrong in stating that 'if you can't measure it, you can't improve it', or science does not need improving, or alternative and more accurate means of science evaluation need to be developed. Or maybe all three. Serious debate, built on historical evidence, sociological insight, and logic, should ensue to handle this important matter*".

[1] DORA (Internet). *The San Francisco Declaration on Research Assessment* (viewed 22 March 2023). Available from: https://sfdora.org/

[2] Hicks, D., Wouters, P., Waltman, L. et al. (2015) Bibliometrics: The Leiden Manifesto for research metrics.*Nature* 520:429–431. https://doi.org/10.1038/520429a

[3] Neylon, C. (2022) Stop misusing data when hiring academics .*Nature*. 607:637 https://doi.org/10.1038/d41586-022-02038-7

[4] Agreement on reforming research assessment [online]. European University Association. 2022 [viewed 22 March 2023]. Available from: https://www.eua.eu/downloads/news/2022_07_19_rra_agreement_final.pdf

[4] Hallonsten, O. (2021). Stop evaluating science: A historical-sociological argument. *Soc Sci Inform*. 60:7-26. https://doi.org/10.1177/0539018421992204