Journal of Biomedical Informatics 75 (2017) 128

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Guest Editorial

The ranking of scientists based on citations

In this issue of JBI, Zerem discusses the role of the scientific impact factor, and in particular how it can be used to assess the impact of the contributions of scientists [1]. This has long been controversial when touted as the best approach for assessing the work of individuals. He also criticizes the H-index, and proposes a new score, which he calls the "Z-score," or Zerem score.

The scientific impact factors for journals are important, as they offer an objective way to assess the impact of publications. This and similar metrics are widely used, even more in Europe than in the U.S., for a variety of purposes, including assessing the scientific contributions of departments, centers and universities. In some places such as the UK funds are even distributed to universities based in large part on citation indices.

For authors, the main indices are the H index [2] and the PageRank index [3]. Zerem brings up some of the limitations of the H index, including that there is a long lag before articles are cited, and the problem of new articles with very large numbers of authors (sometimes 50 or 100) who cannot possibly have all made substantive contributions to an individual article. A further issue is that in some domains—like informatics—even the top journals have relatively low citation indices, so that an author who publishes mainly in that area will be relatively disadvantaged. Yet another is that meta-analyses get arguably more bibliometric credit than they deserve; they are a cited a lot, but do not represent new research.

Zerem proposes a new approach in which the first author gets 100% credit for an article if they are not the corresponding author, the corresponding author gets 50% and the other authors share 50%. If the first author is the corresponding author, they get 100% credit, while the other authors share 100%. While this has some advantages, in some ways it is overly simplistic. A second author often contributes a great deal yet would get no credit for that. Often, the first two authors will include an asterisk saying that they have agreed to be "co-first-authors". In our institution, we generally consider second authorship to be as good as senior authorship position, or corresponding authorship. Often there are situations in which the first and second author contributes equally. In addition, even third and fourth authors contribute a lot, and overall authorship order is often (but not always) in descending order of contribution, except for the senior author. The corresponding author may be selected as a matter of convenience—they may simply be the author on the paper with a stable address. But it would be hard to develop a formula which addresses all these nuances. One approach would be to ask the authors to assign some sort of "contribution score" to each author at the time of submission. But this could be contentious—even determining authorship order can be a challenge.

In Zerem's proposal, there is also a calculation involving the journal's impact factor, multiplied by 10, and the total number of citations, divided by 1000; the net impact of this is hard to ascertain, though it is intended to deal with the issue of evaluation of more recent articles.

An important underlying issue is that any scheme should ideally encourage collaboration, as most important discoveries today are made by teams, many of which are diverse and include contributors from quite different fields. If only two authors per paper got much credit for it, that could discourage such collaborations.

Overall, this topic deserves more attention and better metrics could help. As Zerem notes, formulas like this are being given additional weight by a variety of entities, including grant-makers, promotion committees and the like. This is probably particularly important in disciplines like informatics, which are multidisciplinary, and also have even their best journals not getting cited frequently. As robust as it is, the H index has a number of shortcomings. I think Zerem's proposal also has pros and cons, and would be surprised to see it supplant the H index. But we do need better measures, and this proposal will likely stimulate useful debate. Despite their limitations, because of their convenience, such indices are likely to be used even more in the future.

References

[1] E. Zerem, The ranking of scientists based on scientific publications assessment, J. Biomed. Inf. (2017) [this issue].
[2] R. Van Noorden, Metrics: a profusion of measures, Nature 465 (7300) (2010) 864–866, http://dx.doi.org/10.1038/465864a.
[3] U. Senanayake, M. Piraveenan, A. Zmaya, The Pagerank-Index: going beyond citation counts in quantifying the scientific impact of researchers (eCollection 2015), PLoS One 10 (8) (2015) e0134794, http://dx.doi.org/10.1371/journal.pone.0134794.

David W. Bates, MD, MSc*
The Center for Patient Safety Research and Practice, Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, USA
Clinical & Quality Analysis, Partners HealthCare, Somerville, MA, USA
Harvard Medical School, Boston, MA, USA
E-mail address: dbates@bwh.harvard.edu

* Address: Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, 1620 Tremont Street, 75 Francis Street, Boston, MA 02115, USA.

http://dx.doi.org/10.1016/j.jbi.2017.10.012

Available online 25 October 2017