Journal of Biomedical Informatics 79 (2018) 145–146

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Letter to the Editor

The ranking of scientists

In a recent issue of JBI, Zerem criticizes the widely adopted H-index and proposes a new score (called the "Z-score") to measure the impact of scientists [1]. While we agree with the author on the limitations of the H-index, we believe the proposed Z-score has its own weaknesses and introduces new problems.

The "Z-score" oversimplifies the attribution of author contributions. According to this metric, authors get a certain amount of credit based on the order of authorship, i.e., the first author gets 100% credit if he is not the corresponding author; the corresponding author gets 50% and the other authors share 50%. If the first author is the corresponding author, he gets 100% credit, while the other authors share 100%. An obvious flaw in this formula occurs when there are only two authors; with the first author also being the corresponding author, the second author would get the same credit of 100%. However, as the number of authors increases, the formula increasingly emphasizes the contributions of the first author and the corresponding author and de-emphasizes the contributions of other authors, without distinguishing contributions among them. This may further exacerbate the already contentious issue of authorship and discourage investigators who are neither first nor last authors from contributing their best efforts to collaborative research projects. Modern science, especially biomedical science, is increasingly dependent on interdisciplinary collaborations. We can foresee some collaborators would want to chop up a study into "minimum publishable units" so they can be first authors too. This runs the risk of obscuring the overall significance and hindering the effective communication of findings of important studies to the scientific community and the public.

A good bibliometric would have to recognize the differences among diverse fields, specialties, research types and publication types of scientific publications. As such differences are well appreciated by many investigators, they are extremely difficult to address in constructing a bibliometric. For example, in certain super-specialized surgical fields, the numbers of practitioners or investigators are small. Research articles in these fields accordingly tend to have low citation numbers, but this does not take away the value of such publications. In this case, the lack of citations does not equal the lack of impact.

Similarly, in academic medical centers, there are large numbers of clinical publications with medical students as first authors. For most of such articles, the design and investigational interventions are all conducted by attending physicians and other investigators, whereas the roles of the first-author medical students are mainly collating data and drafting the paper, under the supervision of attending physicians. This is not to diminish the work of the first-author medical students, but rather to highlight an important aspect of modern medical education that is difficult to appreciate in bibliometrics. Zerem criticized the H-index when applied to publications of multicenter clinical trials, stating "a "scientist" may have H-index score over 20 without having actually written a single significant article." While this is true for some clinical studies, we have opposing observations on certain other multicenter clinical trials, especially interdisciplinary ones. In such trials, the design, data acquisition, analysis, and results interpretation involve intensive efforts of more than just a few investigators. Assigning equal contribution to all authors except for the first and the corresponding authors is not fair to those investigators making considerable intellectual contributions to the studies, and is an overly simplistic approach that creates new problems.

A good bibliometric would also appreciate that citations to individual articles do not carry the same weight. The importance of citations can range from simply acknowledging awareness of previous work to laying foundations for the study to be presented. Not all of them can be counted as positive impact or contributions to the field. Unfortunately, we do not have a reliable yet easy-to-implement mechanism to differentiate them. This is the biggest well-known but rarely discussed deficiency in citation analyses. We identified at least the following types of citations (Tables 1 and 2). A good ranking model should be able to evaluate the publication according to different types of citations, not just the number of times they are cited. With the advances in natural language processing, we see potential in automating contextual analysis and citation type classification [2], so that we can better separate positive, negative, neutral and insignificant citations.

There are also many other measures of scientific impact that are not necessarily reflected in citations, such as downloads of shared tools and datasets, integration of evidence into practice or policy, and public communication via the news. Author- or editor-driven efforts to publicize an article may be ethically problematic, leading to positive citation bias, often by sensationalizing the study to make it newsworthy. Some journals even promote their impact factors by citation coercion. As such, when creating a quantitative measure of the impact of scientists, no assumption should be made that "it is generally accepted that the IF (WoS) and the total number of citations of articles published in the journal, are the most relevant parameters of the journal's significance." We argue that many people would disagree with this.

We propose that quality assurance processes, which are well-established for many disciplines, may also be applicable for assessing publications and scientists. Possibly we can learn a great deal from existing quality assurance methods. Quality, by definition, is "fitness for use". Last but not least, we argue that scientific metrics needs to go beyond a score but should utilize a more nuanced evidence-appraisal model that takes into account multifaceted impact quantification and operates in a context-specific model.

We argue that evidence appraisal – the critical appraisal of evidence by informed stakeholders – is much needed for the scientific literature but still under-developed from a systems perspective [3]. It needs infrastructure investment and cultural changes, so that people will not just focus on short-term benefits at the cost of long-term sustainability. Career advancement and grant procurement should not be so heavily based on flawed prior

https://doi.org/10.1016/j.jbi.2018.02.006

Available online 15 February 2018

Table 1
Classification of citations by types of the citing publications.

Types of publication that makes the citation	The purpose of the citation
Review & survey article	One or more of the following: – To cite the awareness or the significance of the problem – To highlight the contributions to related research areas
Methodology article	Serve as baseline for the newly proposed methods
Application article	Provide methodology details

Table 2
Classification of citations by locations in the citing publication (original research article).

Section of the paper where the citation is made	The purpose of the citation
Introduction or background	– To acknowledge awareness of the published work – To describe significance of the problem – To describe the limitations of previous studies – To contrast differences in design or methods with previous studies – To provide the baseline for new improvement
Methods	To use an existing method
Results	To give reference ranges, etc.
Discussion	To use as support material in discussing the achievements or caveats of the new study

measures. We need better metrics that gauge the true scientific impact of a publication, but that also appreciate that different types of publication may require different evaluating metrics, and accept that they may not be comparable.

References

[1] E. Zerem, The ranking of scientists based on scientific publications assessment, J. Biomed. Inform. 75 (2017) 107–109.
[2] J. Xu, Y. Zhang, Y. Wu, J. Wang, X. Dong, H. Xu, Citation sentiment analysis in clinical trial papers, AMIA Annu. Symp. Proc. 2015 (2015) 1334–1341.
[3] A. Goldstein, E. Venker, C. Weng, Evidence appraisal: a scoping review, conceptual framework, and research agenda, J. Am. Med. Inform. Assoc. 24 (6) (2017) 1192–1203.

Chunhua Weng
Department of Biomedical Informatics, Columbia University, New York, NY USA
E-mail address: chunhua@columbia.edu

Andrew Goldstein
Department of Biomedical Informatics, Columbia University, New York, NY USA
Department of Medicine, New York University Medical Center, New York, NY USA

Chi Yuan
Department of Biomedical Informatics, Columbia University, New York, NY USA
Department of Computer Science, Nanjing University of Science and Technology, Nanjing, China

Zhiping Zhou
Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, NY USA