Getting off the Impact Factor: an antidote

Aleksey V. Belikov
Correspondence to: belikov.research@gmail.com

Articles published in high Impact Factor (IF) journals receive higher visibility and attached prestige, regardless of their actual scientific merit. This results in unfair gain in subsequent number of citations to these articles, which also further increases the journals' IFs. While there is a gradual move in assessment of researchers from IFs towards individual performance metrics, such as the h-index, those metrics are calculated on the basis of citation counts, and hence are still affected by the described phenomena. Naturally, this leads to increased submission and rejection rates in high IF journals, considerably delaying publication of manuscripts and wasting researchers' time. Additionally, an article's visibility increases with time, as it accumulates citations, thus severely disadvantaging new articles and early-career researchers. Here I propose a simple method for evaluation of individual articles and researchers that compensates for the effects of the journal IF and article age. In essence, the number of citations to an article is divided by the median number of citations to the articles published in the same journal in the same year. This ratio indicates the performance of an article relative to its closest competitors, is free from journal and age bias, and thus reflects an article's scientific merit. An author's (or institution's) index is calculated as the sum of these article scores. Widespread adoption of this index, especially by decision-making authorities, will refocus scientists from besieging elitist journals to actually doing research.

It is widely known that journal impact factors[1] (IFs) do not reflect the value, and even the impact of individual articles[2-9]. The distribution of citations to articles in a journal is usually wide and skewed, with long upper tail[10]. Moreover, such distributions largely overlap between journals with different IFs[11], indicating that editorial selection and peer review are not very efficient in predicting subsequent article impact. This statement is supported by a long list of initially rejected Nobel prize-winning articles[12]. Hence, the whole IF or prestige ranking of journals is questionable. Nevertheless, articles published in high IF journals receive higher visibility and are considered more prestigious. This would have been an innocent vanity exercise, would not IFs be used to evaluate publications and researchers, strongly affecting how grants and positions are assigned. While there is an ongoing move in assessment of articles and researchers from IFs towards individual performance metrics[13,14], such as the h-index[15] and others[16-18], those metrics are nevertheless calculated on the basis of citation counts, and a high IF of a journal automatically increases citations to any paper published there[19,20], by creating high visibility and perceived prestige. Another problem is that older articles have more time to accumulate citations, and the more an article is cited, the more it is likely to be cited again[21,22], making it difficult to compare articles of different age and researchers at different career stages.

Here I propose to normalize citations to an article by the median number of citations to articles published in the same journal in the same year (i.e., normalize by the Median Neighbors' Citedness, MNC), thus obtaining the Neighbor Article Score. The median is preferred to the mean because of the skewed nature of citation distribution[10]. Thus, an article will be

compared against its closest competitors, neutralizing any advantages in citation rates given by high IF/MNC "elitist" journals and by an earlier appearance. What should remain is the true value of an article for the scientific community. Articles that are cited better than their competitors from the same journal and year will get a Neighbor Article Score greater than 1, whereas those which perform poorer will score from 0 to 1. *The index for evaluation of researchers (or institutions) is then simply a sum of Neighbor Article Scores.* All these indices can increase and decrease with time, and should be regularly updated.

It has to be noted that similar ideas have been proposed before. In case of the Relative Citation Rate, proposed by Schubert & Braun, "the reference standard to which the citation count of each paper is matched is the *mean* citation rate per publication of the journal in which the paper in question was published"[23]. However, the mean is not an adequate measure of a skewed distribution, and approximately 70% of articles in a journal have citation rates below the mean[11]. Moreover, it is not clear for which year this rate has to be calculated. The Percentile Rank Index, proposed by Pudovkin & Garfield, "indicates the citation rank of the author's individual papers among the papers published in the same year and source (journal or multi- authored monograph or book.)"[24]. Although this index clearly has the same goal as the Neighbor Article Score, it appears methodologically inferior. First, differences between citation numbers in percentiles are much smaller than in ratios, e.g., 10 most cited papers in a journal with 1000 publications per year will all belong to 100th percentile but can have Neighbor Article Scores of 5, 7, 11, 14, 17, 23, 28, 37, 42 and 63. Even in the article by Pudovkin & Garfield[24] it can be seen in Table 4 that, for the h-core selection of 29 articles of an author, the Percentile Rank Index (PRI) varies from 63 to 100 (1.6 fold), whereas the citation ratio c_f varies from 0.9 to 14.3 (16 fold). Small percentile differences make their summation in order to calculate an author's index impractical. The problem is "solved" by the Author's Superiority Index, which "is determined by the number of the author's papers with a Percentile Rank Index at or above a specified value (99, 95, or 75)"[24]. However, this approach necessitates the use of an arbitrary threshold, which leads to uncertainty and subjectivity, conceals differences between articles, and simply ignores the majority of publications. Overall, despite the commendable previous efforts, the Neighbor Article Score appears to be the first reasonable implementation of a journal-normalized citation index.

It is particularly interesting to hypothesize how the behavior of scientists during selection of a journal for manuscript submission will change upon adoption of the Neighbor Article Score. With the current system, scientists aim for the highest IF[25], because they have a lot to gain in case of (unlikely) acceptance, but seemingly little to lose in case of (likely) rejection. However, considerable amount of time is lost in searching for the appropriate publication venues, studying and applying journal-specific formatting guidelines, writing journal-specific cover letters, waiting for the initial decisions, and less often, waiting for the reviewers' comments, performing revision experiments, writing point-by-point responses, and waiting for the final decisions, often only to get the manuscript rejected. This time is multiplied proportionally to the difference between the IF of a journal where a scientist starts his submission journey and the IF of a journal where the paper actually belongs. Most of this time could be instead spent on the actual research.

With the evaluation criteria proposed here, scientists would be reluctant to

send anything but their very best work to high IF/MNC journals. This is because papers are compared against each other *within* a journal, and mediocre papers, even if accidentally accepted, will likely accumulate fewer citations relative to the median article in that journal, resulting in a Neighbor Article Score below 1. Amongst other benefits, this will relieve the burden on editors of high IF journals, reduce the number of immediate rejections, and increase the number of papers sent for proper evaluation by peer review. Ironically, it will also lead to further increase in the already high IFs/MNCs of prestigious journals, as fewer mediocre papers will be published there. Moreover, it will then become even less desirable to send mediocre papers there, because of the increased MNC. The only motivational factor to send manuscripts to high IF/MNC journals that would remain is subjectively perceived prestige in case of publication, but it would be counterweighted by the risk of receiving a lower Neighbor Article Score than in less prestigious journals.

Similarly, scientists would not be overly motivated to submit good manuscripts to very low IF/MNC journals, despite the high probability of them receiving more citations than the median article in that journal, because of low perceived prestige and the risk of an article being overlooked. Most importantly, if many good articles will be submitted to a journal with a low IF/MNC, its IF/MNC will rise, so those articles will lose the initial advantage in the Neighbor Article Score. Eventually, articles will be submitted to journals that are most appropriate for them, and sustainable steady-state equilibrium will be achieved, benefitting all parties involved. Adoption of the Neighbor Article Score for research evaluation will not lead to elimination of journal IFs or prestige differences. On the opposite, it might increase them, but also make them more meaningful. However, it will eliminate unhealthy obsession with high IFs, shorten the path of articles to the most appropriate publication venue, and thus remove the unnecessary load on scientists and editors. Most importantly, it will reduce considerable IF-mediated bias in evaluation of research and researchers. What remains for all of us is to bring this proposal to life.

Competing interests

The author is an active researcher and hence is affected by current research evaluation practices.

Grant information

The author received no specific funding for this work.

References

1 Garfield, E. Citation analysis as a tool in journal evaluation. *Science* **178**, 471-479 (1972).
2 Adam, D. The counting house. *Nature* **415**, 726-729, doi:10.1038/415726a (2002).
3 Not-so-deep impact. *Nature* **435**, 1003-1004, doi:10.1038/4351003a (2005).
4 Beware the impact factor. *Nature materials* **12**, 89, doi:10.1038/nmat3566 (2013).
5 Alberts, B. Impact factor distortions. *Science* **340**, 787, doi:10.1126/science.1240319 (2013).
6 Schekman, R. & Patterson, M. Reforming research assessment. *eLife* **2**, e00855, doi:10.7554/eLife.00855 (2013).
7 On impact. *Nature methods* **12**, 693 (2015).
8 Pulverer, B. Dora the Brave. *EMBO J* **34**, 1601-1602, doi:10.15252/embj.201570010 (2015).
9 Verma, I. M. Impact, not impact factor. *Proc Natl Acad Sci U S A*

112, 7875-7876, doi:10.1073/pnas.1509912112 (2015).
10 Seglen, P. O. The skewness of science. *J Am Soc Inform Sci* **43**, 628-638, doi:10.1002/(SICI)1097-4571(199210)43:9<628::AID-ASI5>3.0.CO;2-0 (1992).
11 Lariviere, V. *et al.* A simple proposal for the publication of journal citation distributions. *bioRxiv* (2016).
12 Campanario, J. M. Rejecting and resisting Nobel class discoveries: accounts by Nobel Laureates. *Scientometrics* **81**, 549-565, doi:10.1007/s11192-008-2141-5 (2009).
13 Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols, I. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* **520**, 429-431, doi:10.1038/520429a (2015).
14 *San Francisco Declaration on Research Assessment*, <http://www.ascb.org/dora/>
15 Hirsch, J. E. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A* **102**, 16569-16572, doi:10.1073/pnas.0507655102 (2005).
16 Van Noorden, R. Metrics: A profusion of measures. *Nature* **465**, 864-866, doi:10.1038/465864a (2010).
17 Belikov, A. V. & Belikov, V. V. A citation-based, author- and age-normalized, logarithmic index for evaluation of individual researchers independently of publication counts. *F1000Research* **4**, 884, doi:10.12688/f1000research.7070.1 (2015).
18 Belikov, A. V. & Belikov, V. V. A formula to estimate a researcher's impact by prioritizing highly cited publications. *bioRxiv* (2016).
19 Perneger, T. V. Citation analysis of identical consensus statements revealed journal-related bias. *J Clin Epidemiol* **63**, 660-664, doi:10.1016/j.jclinepi.2009.09.012.
20 Larivière, V. & Gingras, Y. The impact factor's Matthew Effect: A natural experiment in bibliometrics. *J Am Soc Inform Sci* **61**, 424-427, doi:10.1002/asi.21232 (2010).
21 Merton, R. K. The Matthew Effect in Science: The reward and communication systems of science are considered. *Science* **159**, 56-63, doi:10.1126/science.159.3810.56 (1968).
22 Price, D. D. S. A general theory of bibliometric and other cumulative advantage processes. *J Am Soc for Inform Sci* **27**, 292-306, doi:10.1002/asi.4630270505 (1976).
23 Schubert, A. & Braun, T. Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics* **9**, 281-291, doi:10.1007/BF02017249 (1986).
24 Pudovkin, A. I. & Garfield, E. Percentile Rank and Author Superiority Indexes for Evaluating Individual Journal Articles and the Author's Overall Citation Performance. *Collnet J Scientometrics Inform Management* **3**, 3-10, doi:10.1080/09737766.2009.10700871 (2009).
25 Calcagno, V. *et al.* Flows of research manuscripts among scientific journals reveal hidden submission patterns. *Science* **338**, 1065-1069, doi:10.1126/science.1227833 (2012).