Editorial

Dora the Brave

Bernd Pulverer

The San Francisco Declaration on Research Assessment (DORA) points out that using the Journal Impact Factor as a proxy measure for the value or quality of specific research and individual scientists leads to biased research assessment. How can we resist misusing metrics?

If you notice any particularly fidgety journal editors this month don’t worry—this is merely a symptom of the imminent release of the next round of the dreaded, dreadful Journal Impact Factors (JIFs). Editors are concerned, because the JIF directly impacts their journal, as it influences if researchers choose it to publish their research. JIF has a number of flaws, but one entirely outside an editor’s control is noise: a few citations to a single paper can displace a journal in the IF rank list pecking order. Indeed, the JIF would appear to be elaborated to the astonishing significance of three decimal places precisely to minimize the number of ties in journal ranking tables—even if this is at the expense of statistical significance (see ASCB post “A False Sense of Precision”).

Matters are worse for journals just below an arbitrary IF threshold set by research assessment policies. A few years ago, when this journal dipped below 10, its editors were on occasion invited back by senior faculty to discuss submission of their work once the JIF had returned to a level deemed relevant by their institution. The only immunity to such JIF excesses appears to be to sport a well-recognized journal name in lieu of perceived JIF deficiencies. Indeed, the remarkable influence of brand recognition is borne testament by the rapid proliferation of journal families around a number of well-recognized names.

As always, there will be winners and losers in this year’s JIF league tables—but do these numbers reflect real differences in the quality and interest of the science published in the affected journals?

The power of JIF

Journal editors are understandably concerned about the stranglehold of JIF over their journals, but a far bigger concern is its influence on research itself. The JIF has reached such dominance that it influences the publication strategies of journals, hiring at institutions and even how researchers cite; worse, it steers the research itself. Since JIF does not measure the absolute value of research, it can side-line smaller research communities, while over-emphasizing fashionable research.

The use of journal name as a guarantor for research deserving of institutional or funder support preceded that of JIF, but both derive from the same need to predict the quality and importance of the research. The JIF is one of a number of attempts to provide a quantitative, universal metric that promises a quality judgment on the over 25,000 journals and the over 2 million papers published annually in the biosciences. The initial raison d’être for JIF was to aid librarians, who now assess their holdings based on more diverse information including web access. The unabated influence of JIF on science lies elsewhere: its overuse in research assessment. To be sure, the JIF is not per se more flawed than other metrics, and it can and has in fact served as a first step to move countries mired in publication volume-based assessment and cronyism to more rational policies. The JIF’s continuing influence may be down to the fact that the resulting journal rankings are—apart from a number of notable inconsistencies—generally in line with the performance that scientists intuitively expect of the journals they know well. On average, JIF and journal name will correlate with the quality and interest of research published in a given journal. The problem arises if a specific JIF value or journal name is a precondition to place a grant or faculty position—at that point the tail is wagging the dog.

The trouble with JIF

JIF is patently ill-suited for the assessment of individuals, as it in no way predicts citations to a specific paper in a journal. However, there are also inherent flaws in the JIF that limit its utility for the assessment of journal performance. Apart from the misleading extension of the JIF’s significance to three decimal places, and the binning of journals into questionable subject-based league tables, one particularly troublesome aspect is that it is based on a mean (JIF = last year’s citations to all papers published in the preceding two years/citable papers published in those years). Given the skewed nature of the citation profiles of scientific journals, presentation of the median, which indicates that a paper has a probability of 50% or higher of getting that number of citations, would be more appropriate. A journal with a low median may still sport a high impact factor based on high citing outliers. Fig 1 illustrates typical citation distributions—here for this journal. It is immediately obvious that the distribution is not normal—journals publish papers with a wide range of citations; 20% papers in a journal can account for 80% of its total citations. Thus, the median and mean are quite divergent for many journals—illustrated in Fig 2 for this journal. DORA
Editorial

Bernd Pulverer

DORA arose from the realization that while solutions to the JIF dominance are not trivial, it is essential for the scientific community to point to the problem with one voice. A central aim was to move from a tendency to blame others for the problem to a public realization that all the key stakeholders in the research ecosystem are equally beholden to the JIF and could effect change by adopting a number of measures in parallel.

The declaration was launched exactly two years ago and has drawn widespread attention. Some things have improved, such as the separation of reviews and primary research papers in many research assessment forms. EMBO is a funder of researchers, and supports conferences; these activities are subject to competitive selection by peers, and EMBO has taken steps to ensure the evaluation steers clear of relying on metrics, by encouraging the declaration of key references, the provision of a description of major contributions, and instructions to evaluation committees not to rely on metrics.

Journals, researchers, institutions and notably funders need to work together toward a post-JIF research assessment. No single link in the research ecosystem is able to break the spell alone. A scientific journal selects for papers that match its scope and quality criteria—it is problematic if research assessment is delegated to journals. Most of us are both assessors and assessed. In times of constrained research funding, it is hard to expect the assessed to make a stand—but certainly as assessors, we can and should make a change. Scientists are represented at every level of the system and can therefore change it. Start today by evaluating grants, colleagues and papers beyond the reach of any single metric. Before you do that, why not sign up to DORA.