

Five years post-DORA: promoting best practices for research assessment

Sandra L. Schmid*

Department of Cell Biology, UT Southwestern Medical Center, Dallas, TX 75390

ABSTRACT The San Francisco Declaration on Research Assessment (DORA) was penned 5 years ago to articulate best practices for how we communicate and judge our scientific contributions. In particular, it adamantly declared that Journal Impact Factor (JIF) should never be used as a surrogate measure of the quality of individual research contributions, or for hiring, promotion, or funding decisions. Since then, a heightened awareness of the damaging practice of using JIFs as a proxy for the quality of individual papers, and to assess an individual's or institution's accomplishments has led to changes in policy and the design and application of best practices to more accurately assess the quality and impact of our research. Herein I summarize the considerable progress made and remaining challenges that must be met to ensure a fair and meritocratic approach to research assessment and the advancement of research.

Monitoring Editor

Doug Kellogg
University of California,
Santa Cruz

Received: Aug 29, 2017

Accepted: Sep 7, 2017

The San Francisco Declaration on Research Assessment (DORA) movement began 5 years ago next month when a small group of scientists and journal editors convened at the American Society for Cell Biology (ASCB) meeting in San Francisco in December 2012. As journal editors, they were disturbed by the scientific community's obsession with the "Journal Impact Factor" (JIF), a deeply flawed measure of scientific value, which was disproportionately influencing where papers were submitted. More importantly, as scientists they were deeply concerned about the unintended consequences of this obsession on how we communicated our findings, on the career advancement of our young scientists, and on how we (as individuals, academic institutions, funding agencies, etc.) assessed the value of our scientific contributions. Emerging from this meeting and a flurry of emails over the next few months was the DORA (www.ascb.org/wp-content/uploads/2017/07/sfdora.pdf).

Although originally motivated by the misuse of JIF metrics, the greater purpose of DORA was to stimulate discussion around ways to fairly and accurately assess the quality and impact of our scientific contributions. To this end, DORA provided a list of best-practice

recommendations for funding agencies, institutions, publishers, the organizations that supply metrics, and researchers (DORA, 2012). First and foremost was the general recommendation, "Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions." And DORA's assertion, to all stakeholders, that "the scientific content of a paper is much more important than publication metrics or the identity of the journal in which it was published."

DORA, which has subsequently been signed by more than 12,000 individuals and ~400 organizations and institutions, has raised community awareness, started important discussions, and changed policies. But, ultimately, community members themselves, as mentors, department chairs, review panel members, science advocates, policy makers, students, and postdoctoral fellows—we who are the judged and the judges—are responsible for determining how we assess the value of our and each other's scientific contributions. Thus we must continue to question our methods, advocate for sound decision making, protest malpractice, and commit our efforts toward ensuring the integrity and accuracy of research assessment.

JIF: A FLAWED METRIC OF QUALITY AND CONTRIBUTION

The JIF was developed as a tool for academic librarians, who could use this information to make decisions as to which journals to house and pay for with limited space and funds. JIF simply reports the average number of citations received per paper published in that journal during the 2 preceding years. While probably a reasonable estimate of the frequency with which a specific journal is read and

DOI:10.1091/mbc.E17-08-0534

*Address correspondence to: Sandra L. Schmid (sandra.schmid@utsouthwestern.edu).

Abbreviations used: ASCB, American Society for Cell Biology; DORA, San Francisco Declaration on Research Assessment; JIF, Journal Impact Factor; NIH, National Institutes of Health.

© 2017 Schmid. This article is distributed by The American Society for Cell Biology under license from the author(s). Two months after publication it is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>). "ASCB," "The American Society for Cell Biology," and "Molecular Biology of the Cell" are registered trademarks of The American Society for Cell Biology.

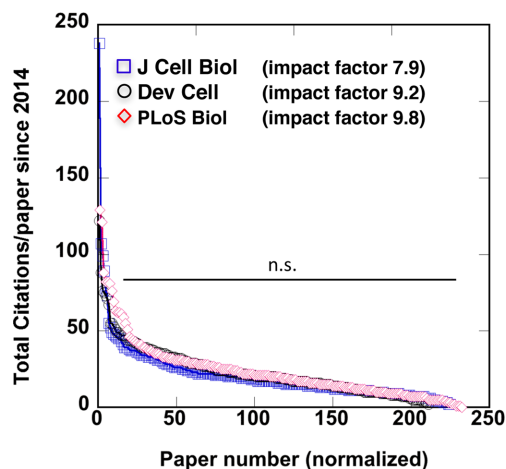


FIGURE 1: JIFs are not a statistically relevant representation of the citation rates of individual papers. The total numbers of citations of papers published in 2014 in *Journal of Cell Biology* ($n = 225$), *Developmental Cell* ($n = 211$), and *PLoS Biology* ($n = 186$) are exponentially distributed (data obtained from Web of Science). Applying the appropriate Kolmogorov–Smirnov (KS) statistical tests yields an average p value >0.1 . Thus the citation rates of individual papers in these journals are not significantly different.

cited, it was never intended as a proxy for the quality of the individual contributions presented in more or less “popular” journals.

As scientists, we know that the mean of a highly skewed data set is statistically flawed, and often meaningless. Indeed, the distribution of citation rates for individual papers in a journal is extremely broad and exponential, as rigorously established by Larivière *et al.* (2016). Consequently differences in citation rates of individual papers in journals with perceived differences in impact factors are, in fact, insignificant (Figure 1). Even for journals with astronomical impact factors, the citations of individual papers are exponentially distributed. For example, the average number of citations of the top 10% and bottom 10% of papers published in *Nature* in 2014 is 341 ± 175 and 18 ± 5 , respectively: a ~ 20 -fold difference! Nonetheless, the all-too-prevalent use of JIF as a surrogate measure of the quality and field-specific impact of individual contributions has led scientists to direct their research toward a subset of “brand-name” journals and their spin-offs, often with unintended and detrimental consequences. A few of the most damaging consequences are discussed below.

The singular pursuit of JIF can delay and sometimes skew communication of research

The perceived requirement to publish in a small subset of the highest-impact-factor journals has led to intense and nonproductive competition. Editors and referees of these journals demand increasingly more information per paper, which now seems limitless due to the combined effects of multi-panel figures that are readable only after magnification on a computer screen, and the advent of online supplemental materials (Schmid, 2011b). Together these factors can delay communication of important discoveries. It’s worthwhile to remember that the molecular biology revolution was launched with a one-page, one-figure letter to *Nature* (Watson and Crick, 1953) in which the authors state, “*It has not escaped our notice that the specific pairing [now colloquially called “Watson-Crick” pairs] we have postulated, immediately suggests a possible copying mechanism for the genetic material.*” Might the Watson and Cricks of today have had to conduct innumerable further experiments to validate this assertion before publishing (Vale, 2015)?

This same, unhealthy competition can lead to the overinterpretation of findings or, on rare occasions, fraud. The perception that each publication should be “definitive” results in polished submissions, which can downplay important details or potential incongruities that could spur new discoveries. Moreover the “appearance” of a complete and definitive story discourages others from the conducting the follow-up and augmentative studies necessary to verify findings.

Unnecessary and unproductive delays in career progression

The consequences of using JIF as a proxy for the value of an individual contribution can be devastating to the career progression of young scientists. The perception, and for too many countries and institutes the sad reality, is that publication in name-brand, high-impact journals is a prerequisite for obtaining an academic position. This is true despite clear evidence that the number of first author papers, regardless of the journal they are published in, is as or more predictive of future success than JIF (van Dijk *et al.*, 2014; von Bartheld *et al.*, 2015). Nonetheless, young scientists and their mentors feel compelled to wait to “have a complete story” before submitting their publications. Worse, they waste inordinate amounts of time and resources on pursuing incremental avenues of experimentation, relegated to “supplemental materials,” at the whim of well-meaning referees and editors. Packaging of a complete PhD thesis or a 4-year postdoctoral study in a single paper is detrimental to a young scientist’s career. Instead, the evidence suggests that young (and old) scientists benefit from the “cumulative impact” of a series of papers that open up and develop new areas of research or solve complex problems (Schmid, 2011a). Moreover, students and postdocs learn invaluable skills by finishing projects and writing papers. Seeing their work in print provides needed motivation and essential training for future success.

THE “IMPACT” OF DORA

First and foremost, the publicity around DORA and the volume and prestige of its signatories catalyzed a dialogue within the international community. Several accompanying commentaries helped to raise awareness of the perils of JIF-driven decisions as to where to publish, who to hire, or who to fund (Alberts, 2013; Eisen *et al.*, 2013; Johnston, 2013; Misteli, 2013; Pulverer, 2013). As a result, researchers, funding organizations, and academic institutions are developing new, more effective means of assessing the quality of an individual’s research contributions and more rapid and efficient ways to communicate our findings. Whether directly or indirectly attributable to this heightened awareness, several positive changes have occurred over the past 5 years, which are listed below:

- Funding organizations in Europe (EMBO, Wellcome Trust, others), the United States (National Institutes of Health [NIH], National Science Foundation), and around the world (Australia, Canada, and others) have instituted, strengthened, and/or made more explicit their guidelines to curtail the use of JIFs and to allow researchers to articulate the significance of their own work, through selected and annotated bibliographies. I serve on several review panels, for NIH, the Wellcome Trust, and the Jane Coffin Childs Advisory committee, and in making their funding decisions, these groups focus exclusively on the scientific content of the papers, not where they have been published.
- Scientific societies, such as EMBO and ASCB, are using JIF-independent mechanisms to evaluate potential awardees at the junior and senior levels.

- Researchers are developing new article-level and field-independent metrics that provide a more accurate and transparent approach to research assessment (Hutchins *et al.*, 2016; Santangelo, 2017).
- Many journals are providing article-level and field-specific metrics accessible to authors and their assessors alike. Importantly, statistics are available on the immediate attention a publication receives, providing timely data for young investigators seeking positions or promotion.
- Many journals, including *Molecular Biology of the Cell*, *Science*, *PLoS*, *eLife*, and all American Society for Microbiology journals, are distancing themselves from their JIFs, by no longer displaying these metrics on their websites.
- Nobel laureates (Schekman, 2013; Nobel Prize, 2017) and bloggers are speaking out against JIF, encouraging scientists, as articulated by Bruce Beutler, “to publish as high as is practical, don’t waste a lot of time on repeated attempts to get in the top tier.”
- DORA’s request of Thomson Reuters, the developer of JIF, to make their data more available and transparent has, in part, been answered (van Noorden, 2014), as the Web of Science (currently administered by Clarivate Analytics), now allows data to be directly downloaded from their site to Excel spreadsheets for independent analysis (see Figure 1).
- The launch and success of bioRxiv, and other preprint servers, together with their acceptance of this practice by almost all journals, is enabling more rapid and efficient communication of results. Indeed, I have relied on the availability of bioRxiv preprints to make positive decisions in hiring (as a department chairman), in funding (as a grant referee), and as an external referee for tenure decisions.
- International forums are taking place to discuss research assessment and the utility and impact of bibliometrics (von Humbolt, 2014).
- Studies are being performed to obtain hard data on the best indicators of future success (van Dijk *et al.*, 2014; von Bartheld *et al.*, 2015).

EFFECTING CHANGE AT THE DEPARTMENTAL LEVEL

As departmental, institutional and organization leaders, we need to develop and disseminate best practices for assessing scientific contributions for most effective hiring and promotion decisions. While the move toward article-level metrics is laudable, these too have their limitations, especially with regard to evaluating early-career researchers and their contributions. In this regard, in my department at UT Southwestern Medical Center we have implemented several changes to how we identify and recruit junior faculty members (Schmid, 2013). Rather than quickly perusing CVs, in which where one publishes and their research pedigree predominates, I ask candidates in their brief cover letter to clearly articulate the value of their past contributions and their future vision. With access to the AcademicJobs website, the entire faculty is able to participate and view applications. If even one member of the faculty expresses an interest in interviewing the candidate, then he or she is invited for a Skype interview. During these 30-minute interviews, attended by three or four interested faculty members, we ask two questions, “Where will your research program be in 5 years” and “How can UT Southwestern help you get there?” Candidates are informed of these questions ahead of time, allowing them to gather their thoughts and present their best sides. This approach has allowed us

to expand the number of candidates we typically interview from four to as many as 30! Eliminating the need for consensus and closed-door committees has also increased the diversity of candidates we interview. The answers to these questions inform us as to which candidates are the best match and most likely to succeed in the specific environment provided by our department and institution. On the basis of these interviews, the entire faculty meets to narrow down the field. We invite three or four candidates for in-person interviews and then make our final offers. In preparing for their visits I ask each candidate to choose three representative papers, which I circulate among my faculty.

Since applying this approach, we have been extremely successful in identifying and recruiting fantastic faculty, ready and eager to integrate into the UT Southwestern community and launch their research programs. Have the successful candidates published in high impact, “brand-name” journals? Most have. However, in my 17 years of experience as a department chair at The Scripps Research Institute and UT Southwestern, this is not the best predictor of success. Instead, in agreement with recent studies (van Dijk *et al.*, 2014; von Bartheld *et al.*, 2015), the best predictor of future success is the candidate’s ability to ask and articulate an important question and to doggedly pursue the answer. In addition to their research statement, seminar, and chalk talk, the best evidence for these key attributes is a series of papers, published both as a graduate student and as a postdoc, each presenting a solid advance, published in rigorously peer-reviewed, field-specific journals (*Journal of Cell Biology*, *Molecular Biology of the Cell*, *Journal of Cell Science*, *Traffic*, etc.). This commitment to a deep understanding of fundamental and important biology frequently leads to the serendipitous discoveries that we see published in the popular science magazines. I see these publications as a by-product of good science, never the primary objective.

MORE WORK TO BE DONE

There is still much to be done, and as Mark Johnson paraphrased, “We have met the enemy and it is us” (Johnston, 2013). As academic leaders, we need to rebel against the practices of some countries that provide monetary awards for publications based on the impact factor of the journal (Franzoni *et al.*, 2011; Quan *et al.*, 2017) and others that arbitrarily ascribe zero value to research published in lower-impact journals. We hope that continued education and awareness of the limitations and even hazards of these practices will eventually lead to their discontinuation. The perception that leading academic institutions throughout the world continue to place undue emphasis on where papers are published rather than on the overall track record of accomplishment of a young investigator (i.e., the cumulative strengths of their contributions) continues its harmful prevalence. We can erase this perception only if we actually change our practices.

As principal investigators and mentors, we need young scientists to recognize the importance of communicating their valuable research in a timely manner, and building a reputation for solid achievement over time, rather than waiting for an often-elusive, singular “home run” achievement. Academic positions are highly competitive. This is a fact. We need to help our trainees to realistically evaluate their competitiveness for these positions and to identify career options that best suit their unique skill sets and qualifications. For those committed to academic research, your success will depend equally on having demonstrable expertise in your chosen area, a track record of completed projects, and well-developed and innovative plans for the future. Have faith that as our community’s commitment to the principles and practice of DORA grows, your talents will be recognized and your careers successful.

ACKNOWLEDGMENTS

Thanks to Mark Patterson, David Drubin, Doug Kellogg, and Erika Shugart for their comments and helpful suggestions and to Khouloud Jaqaman for help with the statistical analysis. Thanks also to the ASCB, which continues to advocate for cell biology, cell biologists, and their careers.

REFERENCES

- Alberts B (2013). Impact factor distortions. *Science* 340, 787.
- DORA (2012). Declaration on Research Assessment. Retrieved from www.ascb.org/dora/.
- Eisen JA, Maccallum CJ, Neylon C (2013). Expert failure: re-evaluating research assessment. *PLoS Biol* 11, e1001677.
- Franzoni C, Scellato G, Stephan P (2011). Science policy. Changing incentives to publish. *Science* 333, 702–703.
- Hutchins BI, Yuan X, Anderson JM, Santangelo GM (2016). Relative Citation Ratio (RCR): a new metric that uses citation rates to measure influence at the article level. *PLoS Biol* 14, e1002541.
- Johnston M (2013). We have met the enemy, and it is us. *Genetics* 194, 791–792.
- Larivière V, Kiermer V, MacCallum CJ, McNutt M, Patterson M, Pulverer B, Swaminathan S, Taylor S, Curry S (2016). A simple proposal for the publication of journal citation distributions. Retrieved from <http://dx.doi.org/10.1101/062109>.
- Misteli T (2013). Eliminating the impact of the Impact Factor. *J Cell Biol* 201, 651–652.
- Nobel Prize (2017). The research counts, not the journal! Nobel laureates speak out against impact factors. Retrieved from https://twitter.com/NobelPrize/status/877804020345262080?ref_src=twsrc%5Etfw&ref_url=http%3A%2F%2Ftheconversation.com%2Fwhy-i-disagree-with-nobel-laureates-when-it-comes-to-career-advice-for-scientists-80079.
- Pulverer B (2013). Impact fact—or fiction? *EMBO J* 32, 1651–1652.
- Quan W, Chen B, Shu F (2017). Publish or impoverish: an investigation of the monetary reward system of science in China (1999–2016). arXiv:1707.01162v1.
- Santangelo GM (2017). Article-level assessment of influence and translation in biomedical research. *Mol Biol Cell* 28, 1401–1408.
- Schekman R (2013). How journals like Nature, Cell and Science are damaging science. Retrieved from www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science.
- Schmid SL (2011a). Lessons from baseball. *ASCB Newsletter: President's Column* 34(June), 3–5. Retrieved from www.ascb.org/wp-content/uploads/2015/12/june3.pdf.
- Schmid SL (2011b). Selective pressures on the evolving scientists. *ASCB Newsletter: President's Column* 34(Mar), 3–5. Retrieved from www.ascb.org/wp-content/uploads/2015/12/mar2.pdf.
- Schmid SL (2013). Beyond CVs and impact factors: an employer's manifesto. Retrieved from www.sciencemag.org/careers/2013/09/beyond-cvs-and-impact-factors-employers-manifesto.
- Vale RD (2015). Accelerating scientific publication in biology. *Proc Natl Acad Sci USA* 112, 13439–13446.
- van Dijk D, Manor O, Carey LB (2014). Publication metrics and success on the academic job market. *Curr Biol* 24, R516–R517.
- van Noorden R (2014). Transparency promised for vilified impact factor. *Nature News*, doi:10.1038/nature.2014.15642.
- von Bartheld CS, Housmanfar R, Candido A (2015). Prediction of junior faculty success in biomedical research: comparison of metrics and effects of mentoring programs. *PeerJ*, doi:10.7717/peerj.1262.
- von Humboldt (2014). Beyond bibliometrics—identifying the best. Retrieved from www.humboldt-foundation.de/pls/web/docs/F-1941001081/duzSpecial_AvH_M04_2015_WEB.pdf.
- Watson JD, Crick FH (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737–738.