# 10 Impact Factor ≠ Impact: Lessons Learned from Research Evaluation

**Luca D. Bertzbach**

Imagine, the most recommended biomedical research paper of the *Faculty of 1000* would describe a fancy CRISPR-mediated technology leading to non-human primates that are fully resistant to multiple hereditary cancers. Unsurprisingly, the publishing journal has a stratosphere-high impact factor, an SJR score of 18.52 and an Eigenfactor® score of 0.332... Furthermore, the senior author's research metrics suggest he is nothing short of a genius with an h-index of 77, an RG score of 49.2, an i10-index of 427, >500 connections on LinkedIn, and about half a million followers on Twitter. Their future research gets mega funding and others build their laboratories based on this outstanding work. However, after years of research and millions invested, it later turns out that a huge proportion of the non-human primate paper was fundamentally flawed. *Quo vadis*, research evaluation?

People have tried to evaluate research and researchers for decades and it is widely agreed that this evaluation should aim at (i) assessing the relevance, efficiency and effectiveness of research projects and (ii) evaluate the scientific contribution of individual researchers. Here, the evaluation of individual researchers optimally serves as a means of assessing their ability to conduct research or to fund their future work—based on their past performance. However, this has always been extremely challenging for those who evaluate research, especially in terms of evaluating potential impact. In part, this is because it gets very complicated to evaluate "real world impact", i.e., impact beyond academia [1]. Moreover, evaluation is often biased since evaluation of researchers is mostly limited by the number of published papers and by the number of citations received [2]. Decades after Science published the concept of journal impact factors [3], this basis of evaluation remains a commonly used proxy for the quality of a scientific article and of a researcher's scientific output. It very often still dictates the choice of where to submit a paper (with prestige of the respective publisher or journal playing a very important role). Until now, those metrics strongly influence a researcher's or funding agency's decision whether or not research can be considered as sound and solid or questionable and dodgy. While it is tempting to boil down complex statistics and evaluations into a single index number, the use of easily quantifiable indicators such as the journal impact factors, the h-index, or similar metrics is too simple and can lead to false assumptions.

*"It is common, and encouraged by many journals, for research to be judged by the impact factor of the journal that publishes it. But as a journal's score is an average, it says little about the quality of any individual piece of research."*

*Randy W. Shekman*

What does the future hold…? Maybe we implement an intelligent combination of advanced metrics and rely on qualitative measures in order to decrease the sexiness of the so-called "smallest publishable unit" and withstand the "publish or perish"-downward spiral. Something like the h-index divided by age or by career-years and only compared to researchers who work in the respective field could be an idea. Researchers should choose their journals wisely according to cost and speed of the publication process and visibility of their work rather than journal impact factors. To date, the majority of online open access journals unfortunately cannot compete with established publishers, mainly because of questions regarding the perceived quality and acceptance of online open access publications [4]. Moreover, as one of the loopholes of academic publishing, predatory journals pose a serious threat to science integrity [5]. These issues nicely illustrate the enormous challenge of unbiased research evaluation and unequivocally emphasize the need for action.

To answer the question of how research and researchers should be evaluated and rewarded, the system requires refinement. A break with publishing traditions, goal-oriented student and post-doc education towards better science communication, a more transparent peer-review process and concepts like the "publish first, curate second" approach [6] proposed by online open access advocates will undoubtedly be baby steps in the right direction. In addition, a focus on team science (spanning boarders and different specialties), the potential of a researcher or a group of researchers to attract funding, and funding of exciting and challenging interdisciplinary efforts that improves common knowledge in the long run should be the premise for funding agencies.

But… what's actually next? One of the biggest publishers could soon launch the *Journal of p > 0.05* to broaden its scope. Doug Zongker's paper "Chicken Chicken Chicken: Chicken Chicken" [7] could finally get retracted (for what reason exactly…?). A group of scientists could pull off a cunning hoax that makes it into news feeds worldwide as yet another example of 'real' fake research in peer-reviewed journals. Way to go, "research evaluation".

## References

1.    Penfield, T.; Baker, M.J.; Scoble, R.; Wykes, M.C. Assessment, evaluations, and definitions of research impact: A review. *Res. Eval.* **2014**, *23*, 21–32.

2. Nature Editorial. Not-so-deep impact. *Nature* **2005**, *435*, 1003–1004.
3. Garfield, E. Citation Analysis as a Tool in Journal Evaluation: Journals can be ranked by frequency and impact of citations for science policy studies. *Science* **1972**, *178*, 471–479.
4. Nature Research. Author Insights 2014. *Figshare* **2015**.
5. Beall, J. Predatory publishers are corrupting open access. *Nature* **2012**, *489*, 179.
6. Stern, B.M.; O'Shea, E.K. A proposal for the future of scientific publishing in the life sciences. *PLoS Biol.* **2019**, *17*, e3000116.
7. Zongker, D. Chicken Chicken Chicken: Chicken Chicken. *Ann. Improbable Res.* **2006**.