Editorial

Are Impact Factors corrupting truth and utility in biomedical research?

What motivates governments to spend taxpayer money on biomedical research? A pragmatic answer would be to improve human/animal health and to support biotech/pharma. Clearly, the research output from such expenditures should be both true (repeatable, robust, reliable) and useful (helpful, applicable, growing knowledge). Most biomedical research funding is competitively awarded as two- to five-year grants or fellowships, with success rates often ≈20% or less, depending upon the grant mechanism. The efforts of ≈80% of applicants (mostly taxpayer-funded scientists) are thus arguably wasted, and therefore expensive in terms of funds, human resources and opportunity costs [1]. A key performance indicator (KPI) or criterion for award of research funding is journal publication output, whose quality is often judged by the Impact Factor (IF) of the journals within which the applicant(s) have published. A timely paper in a high IF journal (e.g., Nature, Cell, Science), can mean the potential difference between continued funding or searching for a new position. Over-reliance on IFs as KPIs has been deemed inappropriate for a range of reasons including; (i) the processes used by the Thompson Corporation to calculate IFs are neither transparent nor openly available (although it should be stated that this represents intellectual property of a for-profit company, not something usually made public), and was actually originally devised as a metric to help librarians determine which journals to purchase; (ii) very bad publications can end up in high IF journals [2], whereas publications with profound benefits for health or science may not (inter alia Gardasil [3] and DNA sequencing [4]); (iii) journal IFs are an average with enormous and growing variances, with the most cited papers increasingly coming from journals with lower IFs [5]; (iv) IFs can be “gamed” by editorial policy [6]; (v) IFs depend on the field, such that bigger fields have journals with higher IFs; (vi) the IF is more a measure of popularity (i.e., how many researchers cite papers in the journal), rather than a measure of truth, utility, or value to the community [7–9].

Such considerations resulted in the “San Francisco Declaration on Research Assessment,” which promoted 18 recommendations and concluded, “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” [7]. The Australian Medical Research Council similarly stated, “Journal IF is not a sound basis upon which to judge the impact of individual papers,” with funding applications no longer including IFs [9]. However, it has become an article of faith that the best science ends up in the best journals, and the best journals have high IFs. This subjective correlation is likely true in the broadest sense, but is often treated as objective and near absolute, with remarkably little effort expended formally to validate IFs as a sound metric of truth, value, and/or utility. Despite the many compelling arguments against their inappropriate use, IFs retain their powerful influence on biomedical science, academia, and research funding decisions. In the simplest setting, grant review panel members with large numbers of 100 to 120-page grant applications to assess in a week may, out of ease, add up IFs rather than carefully evaluate the science.

If a medical practitioner’s pay was determined by counting the number of patients he/she saw (without reference to outcome), one might expect the quality of medical care to deteriorate, irrespective of how much regulatory pressure is placed upon the profession to practice good medicine. The over-use of IFs and associated metrics has, in our view, created analogous problems for biomedical research. Arguably, the IF has become a driver against truth and utility, and has corrupted the motivation for undertaking government-funded biomedical research from “addressing government/industry/community needs and/or growing knowledge” to “publishing in high IF journals.” One study by Amgen reported on the ability to repeat the work published in 53 peer-reviewed papers, and could only reproduce the data in 6 (11%) of these papers [10]. Three were from high IF journals, and three from lower IF journals, suggesting IF is not a reliable arbiter of scientific reproducibility. Even more depressing, the reproducible work was cited on average less than non-reproducible work [10]. Citation frequencies (and associated metrics, e.g., H factor) are thus also not guarantors of truth and utility, with non-reproducible papers spawning further research (presumably misguided and largely taxpayer funded) [10]. A similar undertaking by Bayer revealed that the relevant published data was completely in line with in-house findings in only ~20–25% of projects [11]. There is also a surprisingly robust positive correlation between journal retraction frequency and IF (p = 0.0001) [12]. Retraction rates are rising faster than publication rates, with detailed analyses suggesting a reduction in barriers to publication of flawed articles [13]. Although many “data free” opinions might be offered, these studies suggest our current system has growing problems. It is intuitively obvious that scientists on 2–5 year contracts facing unemployment (or even deportation for scientists on temporary visas), will be tempted to push their data as far as possible. Scientists are human, and have children and mortgages (and often large student loans to repay).

Journals also do not really suffer over the short term if they publish work that is non-reproducible (or even ends up being deleterious to human health [2]), they only suffer if, over the longer term, their IF and sales decrease. To maintain their position, they may also be tempted to publish exciting, topical and/or fashionable papers over true and useful ones. Journals are not
accountable to governments or taxpayers on whose money their existence largely depends. Yet, government-funding agencies are all too often influenced by journals IFs in their decision making. “Scientific excellence,” “cutting edge research” and similar proclamations have also simply come to mean publishing in high IF journals. This is not the stance of the journal Vaccine, and recently we have published our explicit intentions regarding publishing “excellent science” [14]. In general, important KPIs such as scientific truth, utility and/or efficient use of tax payer monies to help improve health and/or support biotech/pharma are curiously often secondary considerations for funding agencies, if considered at all.

Another under-appreciated effect of the over-emphasis on IFs as KPIs is that scientists are less likely to patent their work because patents are infrequently used as KPIs by funding agencies, and patenting is time consuming and can delay publication. The system thus contains inherent disincentives to translate potential taxpayer-funded inventions into new products, processes and industries (which have obvious benefits to the community). As new biomedical interventions can take many years (~8–12 years) before they yield returns on investment, few scientists will risk unemployment on the distant chance of patent-related income. Contrary to the mantra of many institutional technology transfer offices, patenting significantly diverts resources from publication of papers. Returns to inventors are also not stipulated by law in most countries (with the exception of Germany), so a change in institutional guidelines (after, say, 10 years) can deprive the inventor of potential returns. There is little doubt that R&D investment correlates with economic growth and is part of the solution to exiting from the economic crisis [15]. Real world applicability also demonstrates the truth and utility of the science. Why then the obsession with journal publication metrics? In our experience, metrics for academic patents (or other translation metrics such as benefit to patients) are poorly developed, under-utilized, and under-valued. Whether research leads to a triadic patent, serious investment, company formation/participation, product development, benefit to the community, etc., is thus often judged as less valuable than a high IF paper. While we do not have the “best” answers to these dilemmas, clearly it is an important issue, and taxpayers should demand that disbursement of limited government funds is appropriate and provides a worthwhile return on their investment. In the meantime, a robust debate is warranted on how best to measure and insure that taxpayer-provided funds are effectively used. As a first step we concur [7] that the value of IFs should be de-emphasized at the journal, academic, and competitive research funding levels. We also propose that a search commence for a metric that measures truth and utility in biomedical research.

Conflict of interest

Prof. A. Suhrbier is employed on a five-year contract, is active in translation (17 patents), and only has two papers in journals with an impact factor above 20.

References


NHMRC Principal Research Fellow
Andreas Suhrbier (BA, MA, PhD)*
Group Leader Inflammation Biology, QIMR Berghofer
Medical Research Institute, Brisbane, Queensland, Australia

Editor-in-Chief, Vaccine
Gregory A. Poland (MD, MACP, FIDSA)
Mary Lowell Leary Professor of Medicine, Mayo, Vaccine Research Group, Mayo Clinic and Foundation, Rochester, MN 55905, USA

*Corresponding author.
E-mail addresses: Andreas.suhrbier@qimrberghofer.edu.au (A. Suhrbier), poland.gregory@mayo.edu (G.A. Poland)

1 Council of 100 Member, Vaccine.

24 September 2013
Available online 31 October 2013