

The misuse of bibliometrics in research assessment infrastructures

Nandita Quaderi*

Editor-in-Chief, Web of Science, Clarivate, London, United Kingdom

Abstract. This paper is based on a presentation from a panel discussion on Innovation, Technology and Infrastructure at the 2023 APE Conference. It provides an overview of the use, misuse and value of bibliometrics in research assessment, examines the responsible use of metrics and the importance of going beyond existing bibliometric indicators of scholarly impact.

Keywords: Research assessment, bibliometrics, research integrity, responsible metrics

1. Introduction

The first model for regular research assessment of a public sector research base was introduced in 1986 in the United Kingdom. As selectivity in the allocation of public funding increased, coupled with the growth of the research base in many regions resulting in researcher demands exceeding available resources, formal research assessment began to assume prominence [1].

However, research assessment based on qualitative expert review is a time-consuming and costly exercise. This has resulted in a need for solutions to reduce this burden, creating increasing demand for quantitative performance indicators to support the allocation of resources with greater efficiency and effectiveness. These indicators typically include the total number of publications, total citation count, Journal Impact Factor (JIF)TM and h-index.

When used alongside expert qualitative assessment, bibliometrics have value – they can add objectivity and counter any biases in peer review. However, they are also powerful, normative tools that can drive changes in behaviors and create pressure to conform to international and inter-disciplinary norms. Indeed, bibliometrics should be used to strengthen, not supplant, qualitative assessment. Unfortunately, bibliometric indicators have been increasingly adopted as a short-cut around qualitative assessment.

2. Perverse incentives and unintended consequences

Evaluation based on the formulaic use of data or reliance on single measures is not good practice. It creates perverse incentives and can result in unintended consequences as the scores or ranks become a

*Corresponding author: Dr. Nandita Quaderi, Editor-in-Chief & VP, Web of Science, Clarivate, United Kingdom.
E-mail: Nandita.Quaderi@clarivate.com.

goal in themselves. In the arena of research assessment, the misuse of metrics incentivizes maximizing the quantity, rather than quality, of publications especially if they are highly cited or published in journals with a high JIF.

New forms of manipulation are emerging as some stakeholders seek an unfair advantage. Manipulation used to be limited to individuals or small sets of conspirators but continues to increase significantly. Fraudulent enterprises have appeared, using ever more sophisticated technology, that exploit the increased pressure to publish and be cited.

New forms of manipulation include:

- Salami-slicing – splitting data from a single study across multiple articles;
- Inappropriate self-citation – adding spurious citations to inflate citation counts, such as at the journal or author-level;
- Citation rings – trading spurious citations between two or more parties;
- Coercive citation – pressure on authors to add spurious citations by peer reviewers or journal editors.

Fraudulent enterprises that have emerged include:

- Paper mills – produce and sell fabricated manuscripts; provide fake peer reviews;
- Predatory publishers – solicit manuscripts and charge a fee to publish without proper editorial oversight or peer review;
- Hijacked journals – title, ISSN and other features of a legitimate journal are copied to create a separate, fraudulent entity;
- Author marketplaces – bring together parties who are willing to pay for co-authorship with parties who are willing to sell co-authorships;
- Citation marketplaces – bring together parties who are willing to cite for payment with parties who are willing to pay for citations.

In the process, research integrity is being compromised and the scholarly record is being polluted.

3. Responsible use of metrics

Metrics can be a valuable part of assessment when used appropriately, however because they are normative and powerful, we must be careful about the norms we establish.

Dr Eugene Garfield, a founding father of bibliometric analysis, who established the Institute for Scientific Information (ISI) – the forerunner of Clarivate™ – and created the JIF, warned against the “promiscuous and careless use of quantitative citation data for sociological evaluations, including personnel and fellowship selection” [2] in 1963.

There have been many recent voices expressing concern about the misuse of bibliometrics and offering recommendations for improvement, from *The Leiden Manifesto* [3] and *The Hong Kong Principles* [4], to *The Declaration of Research Assessment (DORA)* [5] and *The Metric Tide* [6]. These voices present different perspectives but the recurring themes include concerns about metrics supplanting qualitative analysis and about journal-level metrics such as the JIF being used as a proxy for researcher performance.

For example, *The Metric Tide* introduced the concept of ‘responsible metrics’ to frame how quantitative indicators can be used in an appropriate manner in the governance, management and assessment of research. *The Leiden Manifesto* outlines 10 principles, combining qualitative and quantitative measures, to guide research assessment.

4. Continuing misuse of metrics

Given the damage to the integrity of the scholarly record and the multitude of voices expressing concerns, why do we still see the misuse of bibliometrics?

To quote Robert J. Shiller, 2013 Nobel Laureate in Economic Sciences: “People are often more willing to act based on little or no data than to use data that is a challenge to assemble”.

Summary statistics and league tables offer simplicity and so hold an innate appeal. This however leads to the continued misuse of journal-level metrics such as the JIF as a proxy for researcher performance and information getting lost as data about researchers and their institutions are squeezed into a simplified metric such as the h-index or a league table. The *Profiles, Not Metrics* [7] report from the Institute for Scientific Information (ISI)TM describes alternative visualizations that unpack the information that lies beneath four commonly used metrics and rankings to support responsible research management.

5. Upholding research integrity

Currently, there are experiments focussed on returning to purely qualitative review. However, the volume of global research outputs is far too large for this to be feasible at scale so there needs to be a healthy balance between expert qualitative assessment, use of artificial intelligence (AI)/machine learning (ML) and use of metrics.

While AI/ML could reduce the burden of menial tasks, judgement of research quality by experts remains necessary. In addition, care is needed to ensure that existing prejudices are not propagated or amplified by technology solutions. Furthermore, there is a need to see increased adoption of visualizations and data profiles in place of the single point metrics that are currently in use.

Given the pollution of the scholarly record and erosion of trust, indicators of trustworthiness as well as scholarly impact are needed today more than ever. Currently, only science and social journals with high scholarly impact are eligible to receive a JIF [8]. From the 2023 release of the Journal Citation ReportsTM (JCR), the JIF will be expanded to all Web of Science Core CollectionTM journals that have passed our rigorous evaluation process [9]. This means that the JIF will become an indicator of trustworthiness, as well as impact, at the journal level. In addition, we will move to display the JIF from three to one decimal place in the 2023 JCR release to encourage users to consider other indicators and descriptive data in the JCR when comparing journals [10].

There is growing recognition of the importance of more holistic research evaluation, with approaches that go beyond existing indicators of scholarly impact – based on citations – to introduce indicators of real-world impact that influence policy and practice. The recently published, *The Future of Research Evaluation* discussion paper [11], considers the current issues and developments in research evaluation systems. Critically, the paper highlights the need to move “from manifestos to action”. Collective action is vital so the global research community can continue to evolve research assessment, remove perverse incentives and uphold research integrity.

References

- [1] <https://clarivate.com/lp/research-assessment-origins-evolutions-outcomes/>.
- [2] <http://www.garfield.library.upenn.edu/essays/V1p042y1962-73.pdf>.
- [3] <http://www.leidenmanifesto.org/>.

- [4] <https://www.wcrif.org/downloads/main-website/hong-kong-principles/126-hkp-article-the-hong-kong-principles-for-assessing-researchers/file>.
- [5] <https://sfdora.org/>.
- [6] <https://www.ukri.org/wp-content/uploads/2021/12/RE-151221-TheMetricTideFullReport2015.pdf>.
- [7] <https://clarivate.com/webofsciencigroup/campaigns/profiles-not-metrics/>.
- [8] <https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/web-of-science/core-collection/editorial-selection-process/editorial-selection-process/#>.
- [9] <https://clarivate.com/blog/clarivate-announces-changes-to-the-2023-journal-citation-reports-release/>.
- [10] <https://clarivate.com/blog/mapping-the-path-to-future-changes-in-the-journal-citation-reports/>.
- [11] <https://futures.council.science/publications/evaluation>.