



Check for updates

## OPINION ARTICLE

# Building the infrastructure to make science metrics more scientific [version 1; referees: 3 approved with reservations]

Jennifer Lin<sup>1</sup>, Fiona L. Murphy <sup>2</sup>, Mike Taylor<sup>3</sup>, Liz Allen<sup>4,5</sup><sup>1</sup>Department of Product Development, Crossref, Oxford, UK<sup>2</sup>Associate Fellow, Institute for Environmental Analytics, University of Reading, Reading, UK<sup>3</sup>Department of Research Metrics, Digital Science, London, UK<sup>4</sup>Department of Strategic Initiatives, F1000, London, UK<sup>5</sup>Visiting Senior Research Fellow, King's College London, London, UK**v1** First published: 21 Dec 2016, 5:2897 (doi: [10.12688/f1000research.10422.1](https://doi.org/10.12688/f1000research.10422.1))  
Latest published: 07 Apr 2017, 5:2897 (doi: [10.12688/f1000research.10422.2](https://doi.org/10.12688/f1000research.10422.2))

## Abstract



Research leaders, policy makers and science strategists need evidence to support decision-making around research funding investment, policy and strategy. In recent years there has been a rapid expansion in the data sources available that shed light onto aspects of research quality, excellence, use, re-use and attention, and engagement. This is at a time when the modes and routes to share and communicate research findings and data are also changing.

In this opinion piece, we outline a series of considerations and interventions that are needed to ensure that research metric development is accompanied by appropriate scrutiny and governance, to properly support the needs of research assessors and decision-makers, while securing the confidence of the research community. Key among these are: agreed 'gold standards' around datasets and methodologies; full transparency around the calculation and derivation of research-related indicators; and a strategy and roadmap to take the discipline of scientific indicators and research assessment to a more robust and sustainable place.

## Open Peer Review

Referee Status: ? ✓ ?

	Invited Referees		
	1	2	3
<b>version 2</b> published 07 Apr 2017	<b>REVISED</b> ? report	✓ report	
	↑	↑	
<b>version 1</b> published 21 Dec 2016	? report	? report	? report

- 1 **David J. Currie** , University of Ottawa Canada
- 2 **Maria Nedeva**, The University of Manchester UK
- 3 **Ivan Oransky** , The Center for Scientific Integrity USA, **Alison Abris**, Retraction Watch, The Center For Scientific Integrity USA

## Discuss this article

Comments (0)

**Corresponding author:** Fiona L. Murphy ([fiona.murphy@reading.ac.uk](mailto:fiona.murphy@reading.ac.uk))

**How to cite this article:** Lin J, Murphy FL, Taylor M and Allen L. **Building the infrastructure to make science metrics more scientific [version 1; referees: 3 approved with reservations]** *F1000Research* 2016, **5**:2897 (doi: [10.12688/f1000research.10422.1](https://doi.org/10.12688/f1000research.10422.1))

**Copyright:** © 2016 Lin J *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Grant information:** The author(s) declared that no grants were involved in supporting this work.

**Competing interests:** No competing interests were disclosed.

**First published:** 21 Dec 2016, **5**:2897 (doi: [10.12688/f1000research.10422.1](https://doi.org/10.12688/f1000research.10422.1))

## Introduction

It is an exciting and challenging time for research evaluators and strategists; in the post-digital era, technical limitations around what can be used to assess different aspects of research are falling away. The availability of article-based citation metrics and indicators that capture research article reach, attention, and engagement is helping to reduce a reliance on misleading journal-based assumptions of scientific quality and importance. Many researchers now openly share components of their research – often within a research article, but increasingly outwith. For example, databases, datasets, software, and artistic outputs are often now on a range of platforms (e.g. [Figshare](#), [Zenodo](#)) and independently citable (through the use of a digital identifier, such as a [DOI](#)). In addition, many researchers share analysis through non-traditional media (e.g. preprints, blog posts and policy documents).

At their essence, research metrics are designed to shed light on a range of attributes of research to support decision-making around resource allocation and research funding strategy (including tenure, career appointments and grant applications). In addition, metrics today routinely support national research assessment exercises, as exemplified by [REF2014](#) in the UK and [ERC2015](#) in Australia. Despite this, there continues to be limited investment in either research on the quality and validity of the indicators or the governance and stewardship of the data upon which indicators are derived.

Policy experts and researchers have long petitioned to make research metrics more robust, evidence-based and *scientific* ([Lane, 2010](#)) and therefore acceptable to the community they are meant to serve. Recent analyses have also reported on the current limitations of research metrics, calling for more research on, and improvements in, the infrastructure to support science indicators ([Hicks et al., 2015](#); [Wilsdon et al., 2015](#)). The EU also recently issued a consultation to put ‘alternative’ metrics on firmer footing as part of its drive to encourage open science approaches and robust ways to evaluate research ([Amsterdam Call for Action on Open Science, 2016](#)). However, the ‘science’ of research metrics (scientometrics) paradoxically remains an orphan discipline given that more effective and accurate science metrics could make science more effective.

## Building an evidence base for metrics

We are now at a pivotal point of the research indicator story where a political and administrative appetite for research metrics to build and sustain efficient and effective research systems co-exists with a burgeoning in the sources of intelligence about research outputs. What is needed to harness this momentum is cross-sector agreement on the next best steps and actions to make research metrics more robust, transparent and empowered to work for the whole research community.

Several initiatives are underway whose aim is, at least in part, to consider how to improve the evidence base upon which science is evaluated and make science more effective (see for example, the [EU Open Science Policy Platform](#), and the [UK Forum for Responsible Research Metrics](#) [announced in September 2016]). The key ways that such initiatives will be able to make a real difference, is

four-fold. First, ensure active participation from across the whole scientific research community in a broad way to include researchers, institutions and funding agencies, alongside scientific publishers, learned societies and technology platform providers. Second, deliver a roadmap for the key requirements needed to build and assure quality science metrics for the benefit of science. Third, question existing assumptions around how we conduct and reward research, and test out new approaches and ways of working. Fourth, secure access to resources and influence, as well as make actionable decisions.

Against this backdrop, we believe that there are now a number of very practical ingredients that can potentially act as part of a roadmap to ensure the development of robust and fair science indicators that have community support. We outline these below.

## Definitions, descriptions and sources

For research metrics to be understood and used consistently there needs to be agreement around common vocabulary and descriptors of terms. As an example, [CASRAI](#) is building a dictionary of scholarly research output terminology. This dictionary has multiple users, including groups involved in the development of research metrics.

The definitions themselves need to be definitive, openly sourced, managed, curated, versionable and quality assured. Additionally, the data upon which the indicator is best derived need to be identified. One of the challenges around research indicator derivation to date is that many of those in common usage are based upon opaque methodologies and proprietary datasets. This has eroded trust among the user base - many of whom don’t have access to the data - and pragmatically makes it difficult for particular metrics to be reproduced and explained.

## Availability and preservation of Gold Standard (GS) data

An important concern around current research metrics is that they are often compiled and enabled through proprietary databases with locked access to the underlying data. This creates challenges for third parties wanting to replicate a metric, apply it in a different context or produce aggregate datasets from multiple sources. It also leads to mistrust and scepticism among users and those whose research is described ([Wilsdon et al., 2015](#)).

The community needs a reference set – a Gold Standard (GS) dataset – for proper metrics development. A GS dataset would also enable an ongoing appraisal of best practice for a particular metric’s use and application – and potential inter-relationship with other metrics. Currently, a wide array of metrics is available. These make similar claims, but derive from different formulations. If enabled to work by correlating against a GS dataset, analysts can conduct systematic and rigorous testing and benchmarking for these options to surface the ones most useful across different applications. In short, while the open availability of raw metrics data is critical to transparency and to support innovation in metrics development and provisioning, we need a separate reference dataset that ensures the raw data which underlie a specific metric or metrics are properly preserved and audited.

### Towards open standards

In addition to the raw data, required analytical tools also need to be made available for true transparency and reproducibility (and thereby trust in the metrics). This includes products, such as a defined (minimum core) dataset, and open source standards on how the data are derived and defined (perhaps through an intermediary such as [Crossref](#) or by a cross-functional stakeholder group). [The National Information Standards Organization's](#) work in this area can be built upon in future research. Commercial entities might also serve as potential sources where available to the broader community.

### Research on research metrics and scientific indicators

Perhaps most importantly given the stakes involved, we need greater consensus around how science and research-related metrics are best used to support decision making in science. As noted earlier, metrics need to be created to answer specific research evaluation questions. Research on research (science of science) is needed to help answer the important research evaluation questions and determine which metrics are useful and have the potential to provide insight to these research questions. As researchers adopt new ways to share and publish their research at speed, metrics and indicators that track and assess the value, quality and utility of those activities need to keep pace.

We see a valuable role for funders to play in supporting this particular research area. The community working in the field is small and funding can be difficult to allocate even where funding for research evaluation studies is available (such as the [UK's Medical Research Council's report on how science is funded](#)). Focused funding is also needed to train a cadre of researchers to conduct experiments around what works for science and research, and this includes analyses of research assessment and metrics. Additionally, they (along with policy-makers) can contribute use cases and research questions to the researchers developing metrics to ensure that the outputs are practical and meet real needs. Simply by taking additional notice of this field, funders will be making a critical contribution towards highlighting its significance and expediting progress. Having key leverage on the drivers, incentives and value systems of the research ecosystem, they can enable a shift in behaviours and culture.

### Investing in the online & digital infrastructure

As noted in [Wilsdon \*et al.\*, 2015](#), the digital infrastructure underpins not only the research enterprise but also the creation of metrics. Scholarly outputs of all stripes – articles, pre-prints, datasets, software, and peer review reports – need identifiers (such as DOIs) within this networked ecosystem to facilitate the derivation of metrics. This need extends beyond research artefacts: identifiers for researchers ([ORCIDs](#)), funders ([Open Funder Registry](#)), as well as research institutions. For research metrics to be open, trusted and useful, research objects need to be reliably and meaningfully linked to each other, as well as to researchers, institutions and funding agencies to support strategy and decision-making (see for example [Amsterdam Call for Action on Open Science](#), 2016).

### Community memory on metrics development

Currently, research and documentation on metrics is dispersed. As a non-disciplinary grouping, not a single scholarly community or society spans all the relevant groups working on theory, advancing analytics, data quality, visualisation, policy (and economics). No single party takes responsibility for collecting or documenting process, evidence of good or bad practice, or any other significant issues. The value of these resources may not be immediately obvious, but their absence can stunt the progress of metrics utility, innovation, transparency and dependability.

### A path to fulfil these needs

As researchers adopt new ways to share their scholarly contributions at speed, metrics which describe and provide insight into that work need to keep pace. Different metrics are likely to have different value across output types, research fields and in different circumstances. Yet we believe that a coordinated, cross community effort to enhance our knowledge and application of research metrics is both the timely and sensible route to take. By leveraging the capacity of multiple sectors, we can more effectively create the evidence base context needed to develop metrics able to serve a modern, vibrant research enterprise.

Recent and current initiatives to study and report on scientometrics are evidence of the growing urgency of this issue, but do not so far encompass a sufficient range of functions, regions, technologies or the wider community. For this discipline to be able to progress towards its true potential, a global, cross-stakeholder and truly open project and consultation process needs to be devised. Governance and consultation processes will be critical in order to build trust amongst a wide range of users. Mediation between non-profit and commercial entities, funders, researchers and institutions will need to be baked into the project's fundamental structure.

### Conclusion

This piece is the result of a number of conversations between the authors and others operating in the metrics field. The writing process was punctuated by the [EU Consultation on Metrics](#) and, more recently, the announcement of the [UK's Responsible Metrics Forum](#). These initiatives informed our thinking but, as outlined above, did not fully encompass the scale of action and community involvement we argue is necessary for the paradigmatic change required.

We propose that a coordinated, cross-sector and international effort is required, which operates openly and shares data, resource and expertise across the stakeholders represented. As a next step, we call for the establishment of a group (and adequate funding) to take the first steps by developing the scope and structure of the major project outlined above through community consultations. We hope to see a consensus building round a reputable, transparent entity representing and spear-heading the further development and safeguarding of scientometrics. Powered by the community, this entity would bear the responsibility to take actions that would address the range of concerns and requirements outlined above.

This community entity might take any number of forms. A few examples include:

1. an **independent non-profit membership organisation** (e.g. like ORCID) managed by a cross-sector board and executive.
2. an **independent research metrics foundation** – funded by a consortium of national and independent research funding agencies, whose aim would be to deliver establishment of
3. an **independent, international office of research metrics** - funded by national governments and organisations, whose remit would be to develop standards and deliver research metrics – including to provide ‘a **Frascati Manual**’ of definitions and standards for research/science metrics. This could include an ongoing programme of research (including ability to commission research) to keep pace with developments in science and research practice.
4. an **international, distributed hub of experts** (similar to a learned society) that could, for example, commission and

that can both deliver and advise on scientific indicators and commission work or work with an existing independent funding agency to support a research programme.

More than ever, scholarly research needs effective, trusted research metrics in today’s dynamic communications environment. Each of the actions proposed here is concrete and practical. Yet, they are all united in service of this broad and ambitious goal, so fundamental to the support of the scholarly enterprise at large.

---

#### Author contributions

All the authors contributed equally to this article.

#### Competing interests

No competing interests were disclosed.

#### Grant information

The author(s) declared that no grants were involved in supporting this work.

## References

- Hicks D, Wouters P, Waltman L, *et al.*: **Bibliometrics: The Leiden Manifesto for research metrics**. *Nature*. 2015; **520**(7548): 429–31.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Lane J: **Let’s make science metrics more scientific**. *Nature*. 2010; **464**(7288):

488–489.

[PubMed Abstract](#) | [Publisher Full Text](#)

Wilsdon J, Allen L, Belfiore E, *et al.*: **The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management**. 2015.  
[Publisher Full Text](#)

# Open Peer Review

Current Referee Status: ? ? ?

---

## Version 1

Referee Report 13 February 2017

doi:10.5256/f1000research.11230.r19504

? **Ivan Oransky**<sup>1</sup> , **Alison Abris**<sup>2</sup>

<sup>1</sup> Retraction Watch, The Center for Scientific Integrity, New York, NY, USA

<sup>2</sup> Retraction Watch, The Center For Scientific Integrity, New York, NY, USA

Thank you for the opportunity to review this manuscript. We should note, however, that given the nature of the manuscript, which is really an editorial, we are of the opinion that it probably does not require the kind of rigorous peer review usually applied to original research.

The manuscript presents a good summary of many of the issues facing research into metrics, and offers a plan for addressing them, at a very high level. We would suggest a few additions for improvement:

- A nod to the fact that metrics can always be gamed, and that while making them more scientific could cut down on this risk, it will likely always be possible. It might also bear mentioning that no matter what metrics we end up with, they are no substitute for reading a particular paper. Put another way, metrics may be useful for certain things (eg large-scale productivity) but not others (eg quality).
- A more specific set of next steps. What first steps might funders, scientists, and administrators take? A systematic review and meta-analysis? A gathering to frame the questions and identify funding priorities? Etc.

**Competing Interests:** IO is executive director, and AA is researcher, at The Center For Scientific Integrity.

**We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.**

Referee Report 13 January 2017

doi:10.5256/f1000research.11230.r19300

? **Maria Nedeva**

Manchester Institute of Innovation Research (MIIoR), Alliance Manchester Business School (MBS), The University of Manchester, Manchester, UK

I believe that the title is appropriate and the abstract captures the essence of this opinion piece. I'm emphasising that this is an opinion piece because judgement regarding sources and data would be very different. There is much that I like about this piece and one of the main things is that it puts out there a very important at present discussion: how we use indicators in research evaluation and how we can do this better (or at least in a way that doesn't disadvantage the development of science). The authors are well informed about the state of play and have given serious consideration to what can be done.

I also can see how the very practical proposals in this piece could be implemented and yield some results.

My reservations are about the failure to reach beyond the 'technical' – this is very needed though probably outside of what the authors have set out to achieve here. This is why, I believe that this piece should be published and, possibly, scholars in the UK and beyond encouraged to take part in this kind of discussion.

Hope this helps.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Referee Report 03 January 2017

doi:10.5256/f1000research.11230.r18877



**David J. Currie** 

Department of Biology, University of Ottawa, Ottawa, ON, Canada

In an opinion piece, Lin, Murphy, Taylor and Allen point out that: 1) researchers' contributions to their fields are evaluated for purposes of research grant allocations, career advancement, prizes, etc.; and 2) the quality and validity of metrics that underlie these decisions are not well studied.

Lin and colleagues call for the development of a discipline that will improve the evidence and infrastructure with which science is evaluated. Their point is well-taken. The manuscript was useful (to me) in pointing out some references and links to initiatives that are now underway in this field. However, to tell the reader what needs to be done is much less useful than actually doing something. This manuscript offers some reasonable suggestions about steps that might improve the evaluation of science; the difficulty is that the article does not present any evidence of an advance. More than opinion is necessary to advance the field.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**