



We need a measured approach to metrics

Quantitative indicators of research output can inform decisions but must be supported by robust analysis, argues James Wilsdon.

Metrics evoke a mixed reaction from the research community. A commitment to using data and evidence to inform decisions makes many of us sympathetic to, even enthusiastic about, the prospect of granular, real-time analysis of our own activities. If scientists cannot take full advantage of the possibilities of big data, then who can?

Yet we only have to look at the blunt use of metrics such as journal impact factors, *h*-indices and grant-income targets to be reminded of the pitfalls. Some of the most precious qualities of academic culture resist simple quantification, and individual indicators can struggle to do justice to the richness and plurality of our research. Too often, poorly designed evaluation criteria are distorting behaviour and determining careers. At their worst, metrics can contribute to what Rowan Williams, the former Archbishop of Canterbury, calls a “new barbarity” in our universities. Metrics hold real power: they are constitutive of values, identities and livelihoods.

Since April 2014, I have chaired an independent review of the use of research metrics for the UK government. This week, we publish the results (go.nature.com/smbaix).

They will feed into how British funding bodies will design the next round of research assessment in universities, which is used to allocate around £1.6 billion (US\$2.5 billion) of funding each year. And they will be of interest to any scientist who feels the rising tide of metrics lapping at their ankles. For the research community still has the ability and opportunity — and now a serious body of evidence — to influence how this tide washes through higher education and research.

One certainty is that the lure — and so the fear — of metrics will continue. There are growing pressures to audit and evaluate public spending on higher education and research, and policy-makers want more strategic intelligence on research quality and impact. Institutions need to manage and develop their strategies for research, and at the same time compete for prestige, students, staff and resources. Meanwhile, there is a massive increase in the availability of real-time big data on research uptake, and in the capacity of tools to analyse them.

In a positive sense, wider use of quantitative indicators, and the emergence of alternative metrics for societal impact, could support the transition to a more open, accountable and outward-facing research system. Yet only a minority of the scientists we consulted supported the increased use of metrics. It is clear that across the research community, the description, production and consumption of metrics remains contested and open to misunderstanding.

Our conclusion is that metrics should support, not supplant, expert judgement. Peer review is not perfect, but it is the best form of academic governance we have, and it should

remain the main basis by which to assess research papers, proposals and individuals.

Quantitative indicators can meet their potential only if they are underpinned by an open and interoperable data infrastructure. How underlying data are collected and processed — and the extent to which they remain open to interrogation — is crucial. Without the right identifiers, standards and semantics, we risk developing metrics that are not contextually robust or properly understood.

Universities, funders and publishers need to harmonize their systems of data capture. And they need to make it easier to find and assess fragmented information about research — particularly about funding. If metrics are to be reliable, and not add administrative burden, the priority for the community must be the widespread introduction of unique identifiers, such as ORCID tags, for individuals and research works.

It is tempting to boil down complex judgements to simple scores and numbers, but there is legitimate concern that some quantitative indicators can be gamed, or lead to unintended consequences. Personnel managers and recruitment or promotion panels should be explicit about the criteria they use for decisions about academic appointments and promotions. These criteria should be founded in expert judgement and may reflect both the academic quality of outputs and wider contributions to policy, industry or society.

Such decisions will sometimes be usefully guided by metrics, if the measures are relevant to the criteria in question and are used responsibly. Article-level citation metrics can be useful indicators of academic impact as long as they are interpreted in the light of disciplinary norms and

with due regard to their limitations. Journal-level metrics, such as impact factors, should not be used in this way. To reduce the likelihood of abuse, publishers should stop their unhealthy emphasis on the journal impact factor as a promotional tool.

The research community needs to develop a more sophisticated and nuanced approach to metrics. (Even using the term metrics is a problem, because it implies precision and specificity. ‘Indicators’ is better.) Discussion is crucial, and I invite *Nature*’s readers to share good and bad uses of metrics at our new blog www.ResponsibleMetrics.org. Borrowing from the *Literary Review*’s ‘Bad Sex in Fiction’ award, every year we will award a ‘Bad Metric’ prize to the most egregious example of an inappropriate use of quantitative indicators in research management. Sadly, I imagine there will be plenty to choose from. ■

James Wilsdon is professor of science and democracy at the University of Sussex, UK, and chair of the Independent Review of the Use of Metrics in Research Assessment & Management. e-mail: j.wilsdon@sussex.ac.uk

THERE IS LEGITIMATE
CONCERN
THAT SOME
QUANTITATIVE
INDICATORS
CAN BE
GAMED.

➔ NATURE.COM
Discuss this article
online at:
go.nature.com/ldyree