

Statement on the use of metrics in the assessment of research

May 2020

Introduction

Quantitative metrics¹ can provide valuable data to supplement qualitative measures centred on peer review and expert assessment. They can be used to address qualitative bias tendencies by providing appropriate factual evidence. Decision-making must never be based entirely on quantitative measures, and should always support and not supplant expert peer assessment. This statement lays out the background and context to the use of metrics and provides principles to inform which metrics are appropriate for such use in supporting research assessment.

Background and context

The use of metrics in the assessment of research has an established history². A host of measures have been applied to different aspects of research over time as efforts have been forged to demonstrate research excellence and/or value for money, with varying degrees of success³. Some, like the Journal Impact Factor, which was originally created for the purposes of identifying journals to include in the Science Citation Index rather than with any intention to rank journals, have gained international traction within the research community⁴. The development and use of metrics grew rapidly within the research community from about 2001, when the Web of Science, with its associated citation analysis tools, became widely available online. This use of metrics then surfaced publicly in Gordon Brown's budget speech in 2006 when he declared that the peer review Research Assessment Exercise would be replaced with a system of metrics to assess research quality⁵. Whilst this shift has not taken place in the subsequent development of the Research Excellence Framework, with research assessment recognising the continued primacy of peer review and expert assessment as the most effective measure, interest in how metrics can be used, and used responsibly, has continued to grow.

The University of Hull is no different in wishing to better understand how it can apply metrics appropriately in managing its research activity. There are now many different metrics that can be used (e.g., see those available at The Metrics Toolkit, <http://www.metrics-toolkit.org/>) and the challenge can be as much to identify which are most useful as to applying them. To inform developments in how metrics are applied to research assessment this document lays out a set of principles to guide how the University, and its constituent Faculties and Institutes, can identify and make use of appropriate metrics to address the questions they wish to answer. Some of these will be for local purposes, e.g., promotion and recruitment; some will be to aid benchmarking with other institutions; and others will support national or international reporting. In each case there is a need to look at what is required and then consider which metrics can be used, most often in combination, to address these requirements.

¹ The term 'metrics' is here applied to all quantitative measures for assessing research, whilst recognising that the separate terms 'bibliometrics' and 'altmetrics' are commonly used when referring to specific measures related to research outputs.

² Wilsdon, J., et al. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. DOI: 10.13140/RG.2.1.4929.1363 – see Chapter 2

³ Ibid, Chapter 3

⁴ Garfield, E. *The History and Meaning of the Journal Impact Factor*. JAMA. 2006;295(1):90–93. DOI:10.1001/jama.295.1.90

⁵ Hodges, L. (2006). *The RAE is dead - long live metrics*. The Independent, Thursday 13th April. <https://www.independent.co.uk/news/education/higher/the-rae-is-dead-long-live-metrics-6103828.html>

Our Preferred Methodology for Making Use of Metrics: The Leiden Manifesto

The University's use of metrics will be based upon the Leiden Manifesto, and informed by UK reports, The Metric Tide and UK Progress Towards the Use of Metrics Responsibly. These approaches are outlined briefly below.

In establishing principles for the University there is, as with most metrics, prior thinking within the research community that the University can benefit from. This statement takes as its lead the Leiden Manifesto⁶, a set of principles that were created in 2015 at the University of Leiden in The Netherlands, where there has been ongoing research into the use of metrics for many years. The principles are presented here with a commentary relevant to local usage to set out how each can be applied within the University. The principles themselves do not define the metrics that should be used, but inform the conversations that can lead to these. **To that end, Faculties and Institutes are invited to use this document to inform the development of their own metrics toolset**, informed by local, University, and broader requirements.

As well as using the Leiden Manifesto as its basis, this statement aligns developments in the use of metrics at the University with the recommendations in two reports:

- The Metric Tide, a report published by the Independent Review of the Role of Metrics in Research Assessment and Management in 2015⁷
- UK Progress towards the use of metrics responsibly, a follow-up report in 2018 from a successor body, the UK Forum for Responsible Research Metrics⁸.

A separate statement of principles, DORA, has also attracted interest within some research areas⁹. DORA is a more detailed set of principles targeting specific shifts in metric practice. It will be of more direct relevance to some areas of research within the University than others. Hence, a decision on whether to sign the declaration or not is delegated to those disciplines that can best utilise DORA as part of their metrics toolset.

Research information management and data infrastructure

Alongside recommendations related to metrics, The Metric Tide and its follow-up also make recommendations about the development of the data infrastructure that will aid the effective use of metrics. In particular, the uses of globally used identifiers within research information management systems is highlighted as enabling consistency in the way data is captured, and hence improving the effectiveness in how the data can be used for assessment. These identifiers are:

- ORCID (Open Researcher and Contributor ID) – a standard person identifier for researchers¹⁰
- ISNI (International Standard Name Identifier) – an identifier standard for the names of people and organisations (with the emphasis on the latter in the context of metrics)¹¹
- DOI (Digital Object Identifier) – a standard for the identification of research outputs¹²

⁶ Hicks, D., et al. (2015). *Bibliometrics: The Leiden Manifesto for research metrics*. Nature 520, 429–431 (23 April 2015). DOI: 10.1038/520429a

⁷ The Metric Tide, <https://re.ukri.org/news-events-publications/publications/metric-tide/>

⁸ UK Progress towards the use of metrics responsibly, <https://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Pages/forum-for-responsible-research-metrics.aspx>

⁹ San Francisco Declaration on Research Assessment, <https://sfdora.org/>

¹⁰ Open Researcher and Contributor ID, <https://orcid.org>

¹¹ International Standard Name Identifier, <http://www.isni.org/>

¹² Digital Object Identifier, <https://www.doi.org/>

All informed commentators suggest that the consistent use of identifiers for an academic's research is beneficial to funding bodies, universities and the academics themselves. In applying the principles in this statement, the use of identifiers will be an integral component of any toolset created so that the data used is consistent and can be used with assurance.

Principles in the use of metrics as part of research assessment

The key principles that have been taken into account are as follows:

- 1) Measure performance against the research missions of the institution, group or researcher

The metrics that will be most appropriate in different circumstances should be agreed at the level of research activity at which they will be applied: they should be fit for the purpose for which they will be used¹³. They should be closely aligned with the strategy or objectives being pursued and the associated outcomes from these so the metrics can play a clearly defined role in demonstrating whether those outcomes have been achieved. There should be a clear purpose for each metric being used and a clear link to the area being measured.

- 2) Recognise the systemic effects of assessment and indicators

Multiple metrics should also be used to alleviate any bias that can come about simply from using metrics. Behaviour can change based on what is being measured, hence a collection of metrics will prevent this affecting an assessment and also avoid gaming.

- 3) Keep data collection and analytical processes open, transparent and simple

Data being collected or acquired to provide the basis for metrics should be open and transparent so all those involved can understand what the data is and how and why it will be used. This ensures trust and appreciation of the value of the metric being used. Keeping metrics simple is important as this aids transparency and understanding; however, in doing so the use of simplistic metrics should be avoided, as they lack the substance to provide a valid measure.

- 4) Base assessment of individual researchers on a qualitative judgement of their portfolio

The data available to calculate metrics from is influenced by career stage and gender as well as discipline. For example, a popular metric, the h-index, increases with age regardless of current research activity. Hence, when assessing individuals, metrics can inform and add to an overall picture, but ultimately assessment should be based on a qualitative analysis of their portfolio of work as a whole. Over the years Research Excellence Framework exercises, UKRI funding councils and all of the significant research assessment processes continue to rely on this qualitative analysis.

- 5) Allow those evaluated to verify data and analysis

To further establish trust in the assessment of research, those being assessed, and those carrying it out, should be able to have access to the data to be used so they can verify that this is correct and

¹³ Many metrics related to research outputs are based on data from the predominant corpus of western English language publications. Where research is focused on languages or in cultures that are not widely reported within this corpus there will be a need to recognise this when assessing its performance, identifying relevant metrics that remove bias within the data.

accurate. There should also be open processes in place to allow for challenge where errors or gaps are identified, so that the data quality can be improved.

6) Account for variation by field in publication and citation practices

Identifying appropriate metrics also involves recognising the disciplinary field that is being assessed and its specific metrics. For instance, if assessing a field that produces a particular type of research output, e.g., articles or books, then metrics relevant to that type of output will be most valid.

7) Avoid misplaced concreteness and false precision

Some metrics will suggest they convey a strong message if they score highly. It will be important to ensure that no metric is overly relied upon on the basis of perceived substance, but that multiple metrics are used to validate each other. Metrics that provide excessive detail in their precision should also be avoided so metrics are not used simply for the sake of measuring. Where metrics honestly offer evidence with error margins, these should be transparent.

8) Scrutinise indicators regularly and update them

The factors that will be used to inform which metrics work best described in this document will themselves change over time. Hence, the metrics identified will benefit from regular scrutiny to ensure they continue to be fit for purpose over time, and changed if they are not.
