

Policy for research evaluation using quantitative data

This document outlines a set of principles by which research evaluation and assessment should be conducted at the University of York focusing on the responsible use of quantitative data and indicators.¹

The policy has been informed by (and aligns with) the [Leiden Manifesto](#) and has been developed through consultation with subject matter experts amongst current staff. It is intended to act as an enabling code of good practice and provide clarity for staff on evaluation activities.

Introduction

The University recognises that quantitative indicators on research are now sufficiently well-developed that their usage is becoming more frequent. While such analysis may be established practice in some research disciplines, it is not in others. There is therefore a need for the University to provide clear guidance in this area.

Peer review remains the method of choice for assessment of research quality. By providing guidance on good practice, however, the principles outlined herein support those who wish to use quantitative evaluation measures as a complement.

Context

Bibliometrics is a field of 'research about research' that focuses on scholarly publication and citation data using the latter as a proxy for research quality. Bibliometric data have been used by governments, funding bodies and charities, nationally and internationally, as part of their research assessment processes and are being considered by the UK government as an optional component of the next Research Excellence Framework (REF) exercise where the research discipline is appropriate. It should be noted that bibliometric data are most informative in the sciences and social sciences and less so for arts and humanities disciplines.

More recently the field of 'altmetrics' has emerged in relation to scholarly publications. Altmetrics focus on the online communication and usage of research and can include download data, discussions on research blogs, citations in policy documents and social media mentions.

Other types of quantitative data one might use in research assessment include research grants, research income, industrial partnerships, postgraduate training and commercial activities (e.g. patents, spin-outs, knowledge-transfer partnerships - KTPs).

Application

The policy applies to collective assessment of performance at the level of departments, faculties and the University as a whole. The assessment of individual research performance using solely quantitative indicators is not supported. Such analysis is problematic both in principle and in practice and should be avoided.

¹ The term 'metrics' in relation to quantitative data has gained currency in the UK higher education sector. It is important to note, however, that there are very few true 'metrics' of research performance - they are more accurately 'indicators' e.g. citations are an *indicator* not a *measure* of research esteem. We have therefore used the term 'indicator' throughout.

The principles are not intended to provide recommendations on the application of specific quantitative measures. It is recognised, however, that there is a need for advice in this area and further guidance will be made available in due course. Nor do the principles cover the use of altmetrics or indicators around non-academic impact which are less well-developed, difficult to benchmark and not always applicable to more narrative outcomes such as these.

Policy and principles

Listed below are nine principles for research evaluation and assessment at the University of York.

Principle 1: Quantitative indicators should be used to support not supplant peer-review

The expert judgement and narrative context provided by peer review² is a well-embedded part of the research and publication process. Quantitative indicators, however, can be useful to challenge preconceptions and to inform overall decision-making. As such, the expectation is that both should be used when assessing research quality. It is recognised that the balance between quantitative and qualitative approaches will vary by discipline.

Principle 2: Research evaluation should have clear and strategic objectives

There should always be clearly articulated reasons for the incorporation of quantitative indicators and these should align with relevant departmental, Faculty and University strategies. The expectation is that this alignment be specifically stated in any analysis.

Principle 3: Differences between research disciplines should be accounted for

Contextual information on any disciplinary differences in research indicators should be provided to those undertaking assessment (e.g. average grant sizes, common publication routes, citing conventions) and explicitly acknowledged by them. It should be recognised when it is not appropriate to provide certain types of quantitative data; for example, citation data are not reliable for arts and humanities disciplines. It is recommended that appropriate caveats regarding likely differences between research fields should be acknowledged in any analysis.

Principle 4: Journal-level indicators should not be used exclusively to determine the quality of papers

Journal-level indicators (e.g. JIF) assess journals and should not be used solely to predict the quality of individual papers. High-impact papers can be found in low-impact journals and vice versa. While there is likely to be a broad correlation of journal quality and paper quality it is not necessarily prescriptive. Furthermore, calculation of the Journal Impact Factor does not account for any of the following: publication type (reviews tend to be cited more frequently than articles), research field (e.g. biomedical research is published more frequently and is quicker to accumulate citations than engineering research), journal publication frequency, career stage, skewed underlying data (citation counts do not follow a normal distribution).

It is recommended that paper quality should be assessed using peer review and where appropriate for the discipline, informed by normalised citation impact data.

² The evaluation of academic work by others working in the same field.

Principle 5: A combination of indicators should be used, not a single measure in isolation

It is important that research assessment seeks a variety of perspectives; for this reason we recommend that a suite of quantitative indicators be used rather than a single measure in isolation. The latter is highly unlikely to provide the nuance required for robust evidence-based decision making. The expectation is that multiple indicators are used in any analytical approach.

Principle 6: Data sources should be reliable, robust, accurate and transparent

Source data should be made available where possible. For example, if a department is evaluating its publication portfolio, researchers should be given information on how publications have been sourced (e.g. Scopus, Web of Science) and able to see the publication and citation data included. They should also be given guidance on how to request corrections via these systems. Similarly, researchers should have access to research grants data for the awards with which they are associated and the internal routes for error correction clearly advertised. It is recommended that such information be provided on the appropriate Information Directorate (Library) webpages.

Principle 7: Data analysis processes should be open, transparent and simple and researchers should be given the opportunity to verify their data

Where possible, the criteria of evaluation should be available to researchers and the quantitative indicators used should be easily reproducible. Awareness of potential factors that could bias interpretation of the data should be raised. Existing training for individual researchers and small groups is delivered by the Library Research Support Team within the RETT programme; no training currently exists on strategic usage.

Principle 8: Research indicators and data sources should be regularly reviewed and updated

The systems of evaluation that are used should be sensitive to the evolving needs as an institution, responsive to the changing nature of the research landscape and reflexive. As the institutional understanding of quantitative indicators increases, the institution should also seek to enhance the measures used. The expectation is that a recommended list of indicators be provided to departments and reviewed annually.

Principle 9: There should be a shared understanding of best practice in research evaluation

Institutional webpages should be used to share best practice and the pitfalls of unreliable indicators (e.g. h-index). Those undertaking evaluation using quantitative indicators should have basic statistical training and an understanding of the limitations of the data sources being used. There should be avoidance of false precision; for example, a particular indicator may, in theory, be calculated to three decimal places to avoid ties but, the nature of the underlying data can render discriminating between such values pointless. It is recommended that appropriate information and training be developed in a Faculty-specific context by staff with the necessary expertise. As noted above, the Library's Research Support Team provides one-to-one support for users of bibliometric indicators, and training for small groups, within the RETT programme.

References

1. Hicks, D. *et al.* (2015), The Leiden manifesto for research metrics, *Nature*, 520(7548):429-31. DOI: [10.1038/520429a](https://doi.org/10.1038/520429a)
2. Stern, N (2017) Building on Success and Learning from Experience: An independent review of the Research Excellence Framework ([the Stern Review](#)), *HEFCE*
3. Wilsdon, J., *et al.* (2015), The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management, *HEFCE*, DOI: [10.13140/RG.2.1.4929.1363](https://doi.org/10.13140/RG.2.1.4929.1363)
4. Wilsdon, J., *et al.* (2017) Next generation metrics: responsible metrics and evaluation for open science. Report of the European Commission Expert Group on Altmetrics, *EC Directorate-General for Research and Innovation*. DOI: [10.2777/337729](https://doi.org/10.2777/337729)
5. [San Francisco Declaration on Research Assessment \(DORA\)](#), American Society for Cell Biology (ASCB), 2012