

Imperial College London's response to the call for evidence from the independent review of the role of metrics in research assessment

The main policy objective of research assessment must be that any exercise is able to sufficiently recognise and assess the highest quality research and to fund it accordingly. Equal to this is the desire to reduce the institutional (administrative and academic) and panel burden. We believe that a lighter touch exercise than REF2014 is needed, but it is essential that such an exercise continues to enable the identification of the highest levels of research quality and their differentiation from lesser quality research. Research assessment is used to determine future research funding, and therefore it is important that the methodology used is also able to adequately identify and fund future potential.

Methodology should be decided in consultation with institutions and be introduced only after an appropriate pilot process. Metrics should be normalised for size and subject discipline. In order to deliver the desired reduction in burden for higher education institutions, all metrics, where possible, should be collated from external sources and verified (only) by higher education institutions. A peer review moderated, intelligent use of metrics may be sufficiently robust for the assessment of research output quality. A metrics-based exercise has the additional benefit of assessing the whole research output of an institution or of a subject area within an institution, which also makes it easier to identify "gaming". However, the approach must preserve the trust the sector has in the outcomes of research assessment, and should not incentivise behaviour that is contrary or potentially damaging to normal research practice. In addition, the preparations for research assessment must not hinder the global success of UK research institutions that are competing in an international market against competitors abroad who do not have exercises of such magnitude.

Identifying useful metrics for research assessment

It is essential that any metric used for research assessment should genuinely be able to measure and reflect excellence and quality in its varying forms. A basket of metrics would need to be developed, with limitations understood and contextualised through expert peer review and comparisons made on a subject-specific basis. It will also be important to consider the potential equality implications of using metrics-based systems for research assessment. For example, there would need to be a mechanism for considering specific local circumstances which may have impacted on research outputs and productivity. It should additionally be noted that generally the use of metrics, and particularly citation-based metrics, tends to introduce a bias towards established researchers and towards mature disciplines to the detriment of staff who are establishing their research record and of emerging disciplines. Assuming that the above are addressed, possible metrics include:

Citations: Citations were used as contextual information in the assessment of research outputs in REF2014, but further development is required if citations are to be used as an individual metric in research assessment. There are a number of well-known concerns and issues with using citation data in research assessment. For example, the measurement of citations has limited value for assessing the quality of recently published outputs and the time lags involved vary by discipline, meaning that it is a measure of the impact of research activity that took place some time ago, and is not appropriate to recent activity. Different measures, therefore, might be needed for more recent publications. Other concerns and issues include: different citation patterns in different fields of research; limited applicability to inter-disciplinary research; procedures for considering multi-author outputs; the potential impact on publication practices as the use of such data encourages researchers to reduce their publication volume and only publish work that is likely to be outstanding and highly cited; and the inaccuracies of subject "norms" if applied to smaller submissions or pools of papers. In addition to this, the use of citation metrics in research assessment requires improved data quality in the source systems. It should be noted that in REF2014 large numbers of journals were incompletely indexed in the Scopus database and mismatched outputs were common, meaning that significant work was required on the part of institutions to ensure that the correct citations were listed against submitted outputs. The use of identifiers within data sources (e.g. using author IDs such as ORCID) would help to ensure data quality and enable data to be used to provide useful information. Citations data should not be sourced from just one source, as using a range of sources is the best way to consider the full picture. While different sets of databases could be used for different subject areas, it will be important to ensure consistency between institutions within the same subject areas. Finally, while journal impact factors are useful in identifying influential journals in different research areas, it should not be automatically assumed that if a paper is published in a high impact journal then the quality of the research presented in that paper is high.

Impact: The true impact of research is particularly difficult to measure via metrics, though some forms of impact, e.g. health or economic impacts, are more easily measured than others. The majority of impact “metrics” measure inputs (e.g. patents filed, numbers of attendees/audience members, documented policy, company scale). It is important that, where possible, the actual change or benefit is measured (e.g. patient outcomes, the development of new drugs, products sold, company growth). Exploitation activity must be considered together with evidence of the supporting excellent research and resulting ground-breaking impact. While impact is not easily measured by a metrics-based system, in REF2014 the inclusion of impact in the form of narratives was a particularly significant burden on the academic community. It would be helpful for an analysis to be undertaken of the “impact” part of the REF2014 submission to see to what extent the information could have been captured by metrics.

Measures of esteem and external recognition: These could include external, prestigious fellowships, significant honours and awards, plenary and key note addresses, and memberships of committees. As far as possible these should recognise competitive peer review. Where possible, such measures should not be supplied by institutions but collected externally for institutions to verify.

Research students: As a measure of vitality and sustainability, this could include research student registrations and progression, subsequent destinations of research students after graduation, and sources of funding for PhD students (with contextual information provided).

Key investments in research: This could include investments in research projects, buildings, and sponsored posts, from all funding sources. Such data would need to be normalised by subject area.

Research grants and contracts income: This could include research income from external grants and contracts, particularly from sources which are peer reviewed, as well as research income-in-kind from use of shared research facilities as in REF2014. However, it should be noted that such metrics, used indiscriminately, could potentially penalise blue skies research and other research which traditionally does not attract as much external funding as more applied research. There could also be an incentive for institutions to offer research at a lower full economic cost recovery rate, threatening financial sustainability.

How should metrics be used in research assessment?

A basket of metrics should be used in conjunction with expert peer review in order to reliably identify the highest levels of research quality and their differentiation from lesser quality research. Metrics should be normalised for size and subject discipline, with panels of subject experts having an important role to play in this. In order to deliver the desired reduction in burden for higher education institutions, all metrics, where possible, should be collated from external sources and verified (only) by higher education institutions. It will be important to ensure that any approach preserves the established trust the sector currently has in the outcomes of research assessment, and does not incentivise behaviour that is contrary or potentially damaging to normal research practice. Methodology should be decided in consultation with institutions and should be introduced only after an appropriate pilot process.

“Gaming” and strategic use of metrics

It is inevitable that if metrics are used as the basis for allocating funding, institutions will be incentivised to adapt their practices in an attempt to secure more of that funding. “Gaming”, however, could be minimised by ensuring that:

- Metrics are genuinely able to measure and reflect excellence and quality in its varying forms.
- Metrics do not incentivise behaviour that is contrary or potentially damaging to normal research practice.
- Metrics are used only in conjunction with expert peer review.
- The potential consequences or side effects of incentivising any particular metrics are thoroughly considered.