IOP Institute of Physics

The role of metrics in research assessment

Institute of Physics response to an independent review

A full list of the Institute's submissions to consultations and inquiries can be viewed at www.iop.org

27 June 2014

27 June 2014

IOP Institute of Physics

Dear Sir/Madam,

The role of metrics in research assessment

The Institute of Physics is a leading scientific society. We are a charitable organisation with a worldwide membership of more than 50,000, working together to advance physics education, research and application. We engage with policymakers and the general public to develop awareness and understanding of the value of physics and, through IOP Publishing, we are world leaders in professional scientific communications.

The IOP welcomes the opportunity to respond to the independent review of the role of metrics in research assessment. Our response to the issues listed in the call for input is presented in the attached annex. Please note that our response focuses on the use of bibliometrics in research assessment.

If you need any further information on the points raised, please do not hesitate to contact me.

Yours faithfully,

Tajinder Panesor MInstP
Manager, Science Policy

IOP Institute of Physics

The role of metrics in research assessment

Bibliometrics certainly could have a role to play in research evaluation, even though they are not sufficiently mature to be used formulaically. At a high level of aggregation (e.g. country, institution, field, etc.) bibliometrics correlate well with other indicators of performance (e.g. publications, total research grant funding awarded from all sources), but at the level of an individual, the analysis is not so straightforward. There are also instances where bibliometrics could be used incorrectly – for example, if a Journal Impact Factor is used as an indicator of the quality of an individual paper. Furthermore, citations are largely a measure of the popularity of a paper and do not necessarily correlate with the quality of a piece of work. Moreover, there are a number of factors that affect the citation rate of an article ranging from the subject in which the paper is published to the number of authors on a paper.

When considering individuals, bibliometrics may still offer value but need to be considered alongside a portfolio of other metrics which reflect other aspects of research. Hence, the Institute of Physics is of the view that bibliometrics, alongside an appropriate form of peer review, have the potential to play a supporting role – as a sanity check – in the assessment of research quality.

In recent years, the Institute has commissioned and published two reports that have benchmarked UK physics against international competitors using bibliometric analyses[1]. Using slightly different criteria and methodologies, both evaluated UK physics as a whole and on the basis of individual physics research areas. The feedback from the physics community was of support for the outcomes for the evaluation of physics research as a whole, which showed that the relative *quantity* of UK physics research in terms of world share of publications was decreasing (as was the case for many other leading nations) but in terms of quality indicators was either first or second, and in both cases, above the US.

However, concern was expressed about the analysis at the individual research area level, in particular with how well the journal categories map to specific research areas (important in the context that REF physics sub-panels will have their own breakdowns of physics and there will be little or no scope to expand journal entries to cover a broader range of physics research areas), and that some of the research areas were too large (e.g. large-scale, high-energy physics) leading to substantial citation differences between research areas. There was also the issue of theory vs. experiment and whether citation practices, even within the same journal, might be different for those two areas. Finally, there was the issue around pan-science journals, such as Nature, which are understood to have different citation rates for different branches of science.

[1] Bibliometric evaluation and international benchmarking of the UK's physics research. Summary report prepared for the IOP by Evidence, Thomson Reuters (http://www.iop.org/publications/iop/2012/page_53959.html); and The UK's performance in physics research: National and international perspectives. A report prepared for the IOP, EPSRC and STFC by Science-Metrix (http://www.iop.org/publications/iop/2014/page_63080.html).

One of the reasons we commissioned these studies was a concern that many presentations made by government officials contained bibliometric analyses based on *all* publications from a subject, which showed that in many cases, UK physics did not compare favourably with other subjects. However, we were aware from previous analyses that one can get a different answer by looking at the *total* published output, as opposed to the best, say, 25% from each country, which is the methodology employed for our most recent study by Science-Metrix (i.e. 25% of the most cited publications in physics were used to examine international collaboration among leading organisations in order to identify world leaders not only based on raw output, but also according to high-impact research). This approach is similar to that of the RAE/REF, reviewing only the best papers from each submitted member of staff. As stated previously, in terms of supporting research evaluation, the Institute is of the view that bibliometrics are not sufficiently mature or standardised to be used formulaically, but can be used to support the outcomes of peer review, if bibliometrics are restricted to the four papers submitted and not the global number of a particular research area.

The second of the studies that we commissioned from Science-Metrix used two different methods for the analysis of physics papers; full counting, and fractional counting.

Full counting means that if authors collaborate on a single paper, then when the papers are counted, each author will have one whole paper counted towards his or her country or area. Fractional counting means that if authors collaborate on a single paper, then when the papers are counted they are each awarded their share of the paper towards their country or area. For example if eight authors were to collaborate on a paper then they would each be awarded one eighth of a paper towards their country or area. The key difference between the two techniques is that full counting is influenced by the rate of collaboration. For example, if lots of authors decide to collaborate, then the number of papers produced will be shown to increase. In contrast, fractional counting is likely to show less growth or possibly negative growth in output with more collaboration. For example, the UK's share of the paper output has shown a steeper decline based on fractional counting than based on full counting.

Therefore, in terms of research evaluation, clear parameters need to be established from the outset on what methodologies and techniques are to be employed to complement the peer review process and how variances in outcomes for individual research areas will be addressed. For instance, physics is a field with large variations in authorship, ranging from few authors in theoretical physics to large teams of researchers in "big science" projects using multi-billion-pound equipment. To alleviate the distortions created by a large number of authors in physics, it is preferable to rely on fractional counting for output indicators.

The Institute has listed some issues relating to the possible use of bibliometrics to assess robustly the broad spectrum of physics research alongside peer review in any future REF; it is important to note and consider that the variances in methodologies and techniques can't always be normalised out in a robust way due to the choice of keywords and the inherent biases and uncertainties that the remainder of this response lists. These were developed with input from the Institute's science advisory committee partly in response to a consultation from the higher education funding councils on the development of the REF exercise, which included detailed input on the potential use of bibliometric data that perhaps may still be of relevance to the current review.

Normalisation of citations for different sub-fields

This is perhaps one of the biggest and most significant challenges. Even within sub-fields such as particle physics, papers that develop techniques to be used by large theoretical or experimental groups tend to receive far fewer citations than papers from the groups that adopt these techniques to make interesting measurements or computations. Furthermore, in particle physics theory, where arXiv is the main source of material for all researchers, the impact factor of the journal is less relevant than in biology, for example. This is probably not the case in other areas.

Types of document

There is significant variation in the role of different types of publication in different fields. Care must be taken in choosing whether, and how, to include these in citation analyses.

One example is conference proceedings; in some areas conference presentations are primarily a training activity for PhD students and young researchers, but in others they are a priority means of publication. Conference proceedings journals are not awarded Impact Factors by Thompson Reuters but are still vitally important part of communication in some fields.

Archive-published, non-peer reviewed journals, internet publishing, open access and other alternative publishing methods are already significant in some areas, such as particle physics and astronomy. Blogs and video abstracts are increasingly being used to make research more accessible and to raise the impact and visibility and these are not measured using traditional citations. It will be important to bear in mind the changing publishing landscape which has impacts at the field or sub-field level.

Citation database

The Thomson Web of Science (WoS) or Scopus will not be appropriate for all areas of physics. For particle physics, the primary citation database is SPIRES (the Stanford Public Information REtrieval System), and for astronomy and astrophysics, the primary citation database is the Astrophysical Data System (ADS). It is understood that while most of the natural sciences have 80-95% coverage in the WoS, there have been cases in the physics community where it covered only around 40% of a department's citations compared with those covered by SPIRES, or where 25% of an individual researcher's citations have simply been missed. There is also a concern that the WoS field categorisation is not sufficiently fine grained. This will mean that publications in some journals emerge with a low number of recorded citations compared with the average normalised for that field.

Additionally, many are of the view that Google Scholar is a more reliable database than either WoS or Scopus in some areas. Nonetheless, surveys of citations can reveal enormous variability for a particular paper or author among all three of these databases (and the more specialist ones such as SPIRES). In most areas of physics one could find a large spread of results, where one database will give a paper 0 citations whereas others might give 20; of further concern is that sometimes one database could give 5 and others 45. There is a significant difference between 5 and 45 in perception, and one might think it is a database failure if a paper has 0 citations, but if it erroneously gives 5 or a similar small number, the result will be misleading.

Self-citation

While there are obvious and strong reasons for excluding self-citations, this could disadvantage both small and emerging disciplines, as well as research groups leading larger sub-fields. Excluding self-citation is also problematic when papers

have a large number of authors, for example, in experimental particle physics where there may be several hundred authors for one paper. Additionally, there is also the issue of, for example, crystal growers, where you could be the only person that can make something, so your name will be on most of the papers, giving the illusion of self-citation.

Author numbers

As already stated, most physics papers have multiple authors with substantial variation in the average number. Multiple authorship causes various problems for a bibliometrics approach. In assigning a proportional rating for joint authorship, there will be a question over major or minor authors: not all authors listed will have contributed the same amount of time or expertise to the paper. Additionally, there will be difficulties in the treatment of student authors under the supervision of an academic. The extreme cases will be the LHC experiments with more than 2000 authors.

Seminal papers and other anomalies

Papers with important results that are useful to a large number of researchers have citation rates that drop rapidly as the work simply becomes an accepted part of standard knowledge in that area. Other potential anomalies in citation rates could come from researchers responsible for maintaining a piece of equipment at a major facility/site, who will be highly cited regardless of research excellence.

Period of review

It is well known that different subjects show markedly different behaviour with regard to the timescales for citations. Overall, across all subjects, citations peak in the 3rd and 4th years after publication, and therefore a five-year window for the citation period should be the minimum acceptable. However, given the different behaviour it is clear that a single period is not appropriate and there seems no good reason to insist on uniformity. In passing we note that there appears to be no field-by field analysis of these average times.

Publication type

It is crucial that publications such as review papers, news and views articles can be normalised effectively against publications of the same type. These articles are important in communicating science and raising the profile of UK research, and researchers should not be discouraged from publishing them.

Open access publication in arXiv

For the core subject areas covered by arXiv, many researchers post different versions of an article at different stages of the peer review process, with final publication in a journal being one of many steps. Therefore, in some areas such as high energy physics, the impact factor of the final journal may be less relevant but, more importantly, not all of the relevant citations are captured by analyses as they stand. This is because they do not typically count citations of the arXiv preprint. Yet these can be the most significant ones especially if the timing of publication in a journal is outside of the window of the REF but the preprint is within the window.

As new forms of publication evolve, normalised metrics will need to be developed to ensure that meaningful comparisons can be made year on year. The measurement of the reach, impact and quality of research output is becoming wider than just publication in a journal and this could be where alternative metrics could provide a wider, richer picture than using a single metric of citation in isolation.

The Institute of Physics is a leading scientific society. We are a charitable organisation with a worldwide membership of more than 50,000, working together to advance physics education, research and application. We engage with policymakers and the general public to develop awareness and understanding of the value of physics and, through IOP Publishing, we are world leaders in professional scientific communications.

IOP Institute of Physics

76 Portland Place
London W1B 1NT

Tel: +44 (0) 20 7470 4800
Fax: +44 (0) 20 7470 4848
Email: physics@iop.org
Website: www.iop.org
Registered Charity No. 293851