Bibliometrics and research impact

A quick guide for academic staff
1. What is bibliometrics?

Bibliometrics is the measuring influence or impact in the journal literature. It is the **quantitative** analysis of research literature, based upon citations, and can be used to evaluate the impact on the academic community of a research paper, an individual researcher, a research group or institution, or a journal. Although bibliometrics cannot be used to measure the quality of research, it can be useful for understanding the significance or impact of research.

2. Why use bibliometrics?

Bibliometrics can be used to demonstrate the impact of your own research or that of your research group or institution. Bibliometrics can also be used to identify top performing journals in a subject area, in order to:

- decide where to publish
- learn more about a subject area
- identify emerging areas of research

3. Limitations of bibliometrics

a. Quality

A large number of citations does not automatically mean that a work is of high quality. A work may be heavily cited because other authors are refuting its research.

Bibliometrics does not measure quality. It is important to put the data in context using a combination of metrics and other qualitative information where appropriate, such as funding received, number of patents, awards granted and qualitative measures such as peer review when evaluating quality of work.
b. Discipline Variation

Citations patterns differ greatly between disciplines so direct comparisons cannot be made. Bibliometrics predominantly focuses on journal article citations, but some disciplines such as the arts, humanities and social sciences publish research in different types of publication.

Different fields of research publish at different rates. For example, in biomedicine, there is generally a much stronger culture of publishing in journals and citing the work of peers than in engineering which makes more use of conference papers.

c. Database Variation

The bibliometric databases do not cover all research areas and do not index all publications e.g. conference proceedings or reports are often poorly covered.

Results will vary depending on the database you use, so do not rely on just one.

d. Bias and Discrepancies

Citation bias. People may inappropriately cite their own work, their colleagues, or work from the journals in which they publish. A number of bibliometric tools allow you to exclude self-citations.

Experienced researchers have an advantage over early career researchers as they will have produced more outputs over a period of time and so will have more citations.

There is a bias towards English language material.

Time is needed before a meaningful citation analysis can be made, so new journals tend to fare badly.

Bibliographic tools cannot always reliably differentiate between researchers who share the same surname and initials, meaning that citation counts may be inflated. Researchers can use unique researcher IDs, such as ORCID, to reduce the risk of this.
4. Bibliometric tools

At the University of Cumbria the main citation tool is the Web of Science. Web of Science can provide article level citation data. The Journal Citation Reports (JCR), contained in Web of Science will provide you journal level metrics.

Google Scholar also offers some citation searching, click on the cited by link under search results. Google Scholar also provides some journal level metrics (discussed later).

5. The main metrics

a. Publication Counts - This is simply the number of publications produced by an individual, a research group or an institution. This is the most basic metric used to measure productivity.

b. Citation Counts - Citations are used to measure the impact or influence of a paper or group of papers. It is important to understand that the ‘count’ will vary depending on the database used.

c. Journal Impact Factors - The Journal Impact Factor (JIF) is a metric applied to individual journals. JIF represents the average citation counts of papers published in a journal over a two year period. JIF is produced using data from the Web of Science. The JIF is updated every year and published in the Journal Citation Reports (JCR). A high impact factor is reliant on an article being cited many times in the journals indexed in JCR.

JIFs are created by an algorithm that produces a score based on number of citations of published items in a particular journal divided by the number of items published in that journal over a preceding two year period e.g. this is how the impact factor score is calculated:

Cites in 2014 to papers published in 2012 and 2013 = 165
Number of papers published in 2012 and 2013 = 67
Calculation: 165 divided by 67 = 2.463
Only factoring the last two years enables the factor to be more relevant as it disregards articles which are commonly cited e.g. as background or classic texts.

**An example to show why JIFs should be used with caution:**

Here are JIF scores for two journals:

- *Journal of Criminal Justice*: 3.154
- *New England Journal of Medicine*: 55.873

You will notice that one journal would appear to have a much higher JIF score. However, this does not mean that one is a better journal than the other, in fact both journals represent the title with the highest JIF in their subject category. Therefore it is essential that journals are compared like for like within the context of the title’s JIF data.

d. **Eigenfactor Score** - Journal Citation Reports also provides an Eigenfactor score for journals.

Like the Impact Factor, the Eigenfactor Score is essentially a ratio of number of citations to total number of articles. However, unlike the Impact Factor, the Eigenfactor Score:

- Counts citations to journals in both the sciences and social sciences.
- Eliminates self-citations. Every reference from one article in a journal to another article from the same journal is discounted.
- Weights each reference according to a stochastic measure of the amount of time researchers spend reading the journal.

Eigenfactor scores are scaled so that the Eigenfactor scores of all journals listed in JCR sum to 100. Thus if a journal has an Eigenfactor score of 1.0, it has 1% of the total influence of all indexed publications. In 2013, the journal Nature has the highest Eigenfactor score, with a value of 1.603.
d. H-index

The H-index (a single number) is used to assess an individual, a research group or institution by taking into account productivity (paper counts) AND impact (citations).

If an individual has a h-index of 7, this means that 7 of their papers have been cited at least 7 times each.

As with other metrics, the h-index can only be used effectively for comparing like with like, for example: individuals in a similar discipline and at a similar stage in their career.

e. Google Scholar h5 index

Google scholar provides a journal metric as an alternative to the impact factor called the h5 index.

It is the equivalent to the Hi-index above, but calculated for a journal rather than an author, over a 5 year period. So, an h5 of 10 means that during the past five years a journal has published 10 articles which were each cited at least ten times (and many more articles which were cited fewer than 10 times).

f. Altmetrics

Altmetrics are based on the number of times an article is shared, downloaded or mentioned on social media, blogs or newspapers. In today's digital world, researchers share materials online, and the wider public can engage with research outputs via a variety of different media such as Twitter, blogs, news reports etc.

Often this means impact can be seen more quickly, with research having a broader reach.

Altmetrics aims to capture this alternative impact from digital communications. It is an emerging method of analysing the impact of an article in social media.
6. The benefits of a unique researcher ID

A permanent unique researcher ID will ensure that publications are matched correctly to their respective authors.

- Web of Science provide a service called Researcher ID
- The University of Cumbria encourages researchers to create an ORCID ID. Open Researcher and Contributor ID (ORCID) provides a persistent digital identifier that distinguishes you from every other researcher

7. Summary

The main points to consider when conducting bibliometric analyses are:

a. Always compare ‘like with like’, for example:
   - Groups or individuals in the same or very similar disciplines
   - Groups or individuals at the same stage in their academic careers
   - Journals in the same discipline or category
   - Similar size institutions

b. Don’t rely on one single tool as results can vary depending on the tool used

c. Be aware that some disciplines rely less on publishing in journals than others and therefore will fare badly when conducting bibliometric analyses using some of the main tools available

d. Use a combination of metrics and other qualitative information where appropriate

This guide is a derivative work based on an Open Education Resource by Skills@Library, University of Leeds:
https://library.leeds.ac.uk/tutorials/bibliometrics/introduction/
Further information and contacts

**Web of Science**
http://www.cumbria.ac.uk/StudentLife/Learning/Resources/Eresources/WebOfScience.aspx

**Google Scholar Metrics**
https://scholar.google.co.uk/citations?view_op=top_venues&hl=en

**Web of Science Researcher ID**
http://wokinfo.com/researcherid/

**ORCID**
http://orcid.org/

**Altmetric**
https://www.altmetric.com/audience/researchers/

**Research Office:** research.office@cumbria.ac.uk

**Library and Academic Advisers:** Staffnet/Services/LiSS/Library and Academic Advisers

LIBRARY AND STUDENT SERVICES