

Guidance on comparing UCL to other institutions

August 2020

bibliometrics@ucl.ac.uk

Overview

The [UCL bibliometrics policy](#) sets out some principles for the use of citation metrics in research assessment at UCL. As part of [the overall guidance](#), this paper sets out some general advice on appropriate ways to use metrics on a larger scale, comparing UCL to other institutions.

One of the most common uses for metrics is for overall benchmarking purposes – how is UCL doing in comparison to other institutions? What other institutions are broadly comparable to us in terms of research output? However, there are some potential pitfalls in comparing institutions directly, and it is important to try and correct for size and for disciplinary mixture. Comparing UCL to Imperial, for example – where both are large universities working across a wide range of fields – would be easier than comparing UCL to LSE, which is substantially smaller and has a very different mix of research activity.

Where can we find appropriate metrics?

We recommend you use the InCites tool. This aggregates data from Web of Science, and is available to all UCL users. InCites can be accessed through a link in the top bar of Web of Science, or else by going to <https://incites.clarivate.com>. It requires registration when you first use it – we recommend you use your UCL address here. You will need to be on the UCL network when you first register, so be sure to use a remote desktop or VPN if you are off-site.

To get papers for a single institution, select “organisations” and then, under “dataset” in the top right, ensure it is set to the default “InCites Dataset” and tick “with ESCI” (this ensures all possible papers are added; the ESCI is the optional “emerging sources” dataset of newer journals). If appropriate, set your preferred timeframe – InCites goes back to 1980, and up to the current year – and then scroll down to the “By Attributes” filter, and enter the institution under “Organisational Name” - you can enter several at once here if desired, and you will get a list for each one. Institutions are listed hierarchically, so that eg “University of London” will include all papers from all of the University of London institutions. For historical reasons, InCites lists the UCL Medical School as well as UCL; however, all medical school papers are included in the main result for UCL, so this can be ignored.

If you would prefer to see a broader list – eg “all British universities” – then rather than entering them one by one, you can set the organisation-type and location filters (“Academic” and “United Kingdom”). There is also an “Association” filter, which can be used to return only institutions which are a member of a particular grouping. This is more heavily used for American institutions; the only association relevant for the UK is the Russell Group.

Once you have found the relevant institutions, you need to select appropriate metrics. By default, the table of results displays the total documents and total times cited, which are not desperately helpful. Click the cogwheel icon on the top left of the table of results, and select the desired metrics from the list. A detailed list is given below, but we recommend the **Category Normalized Citation Impact** (under “impact”) and **% Documents in Top 1%** (under “productivity”)

The resulting table now contains the metrics for all institutions in the sample, calculated for each institution separately, and optionally includes benchmark data. These can be selected from the “benchmarks” option on the top bar, and two are available as standard – the “global baseline” and “baseline for all items”. The **global baseline** is calculated for all papers in InCites as a whole, and gives the world average for all papers. The **baseline for all items** shows the average for the documents which appear somewhere in our analysis – so if we have filtered it to just show results for UK universities, it will contain the average metrics for all papers associated with any UK university.

Organization Type		Search 154 results...				Benchmarks	Download
Location		Name	Rank	Web of Science Documents	Category Normalized Citation Impact	% Documents in Top 1%	
X Academic		Global Baseline	n/a	63,922,669	0.98	1.04%	X
X UNITED KINGDOM		Baseline for All Items	n/a	4,076,247	1.36	1.71%	X
▶ Author Position (2008-2020)		<input type="checkbox"/> ▶ University College London	1	350,880	1.64	2.48%	
▶ Rank		<input type="checkbox"/> ▶ University of Oxford	2	322,119	1.81	2.8%	
▶ Association		<input type="checkbox"/> ▶ University of Cambridge	3	289,398	1.77	2.74%	
▶ By Research Network		<input type="checkbox"/> ▶ Imperial College London	4	275,681	1.63	2.44%	
▶ Collaborations with People							
▶ Collaborations with Organizations							
▶ Collaborations with Locations							

In this example, we can see that there are around 64 million documents in the InCites dataset, and four million documents associated with a UK university, of which around 350,000 are associated with UCL – the largest single institution. We can also see some citation metrics; for example, the UCL papers have a lower CNCI than Oxford, but higher than the baseline for the UK as a whole and substantially higher than the baseline for the world at large.¹

When dealing with aggregated groups of papers, it can be important to decide whether to *exclude* any records. By default, InCites provides all publications, but often it is more useful to concentrate on key item types such as articles and reviews. In certain fields, conference proceedings may be more appropriate than articles/reviews; in others, it may be important to include books and book chapters, though InCites citation data on these is more limited.

For simplicity and to ensure consistency, many institution-scale analyses only consider articles and reviews, although this can penalise institutions that are particularly strong in, for example, engineering or the social sciences.

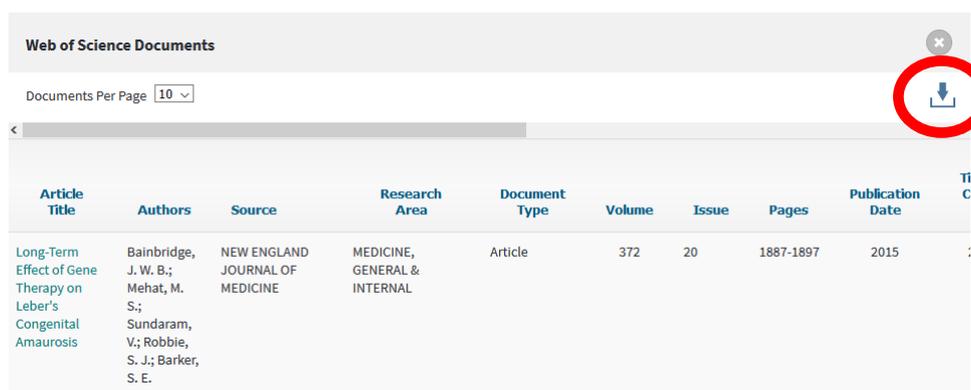
If you wish to filter by item types, you can do this using the left-hand sidebar. It is probably better to select desired item types rather than filter out undesired ones, as InCites indexes a very large number of types of record. In most cases we would recommend articles, reviews, and possibly conference proceedings or book chapters.

¹ It may seem counterintuitive that the global averages for normalised metrics are not *exactly* 1; this is an artefact of the way they are calculated.

Downloading lists for analysis

You can download the aggregate data, as displayed on the table, by using the download button next to “benchmarks” in the top right of the list. Graphs can also be downloaded using the corresponding button.

If you would like to access a list of all papers belonging to an institution, you can see it by clicking the number under "Web of Science Documents" in the data results table. Download it with the arrow icon on the top right of this list. This list is limited to 50,000 records; if you have more than this for an institution, consider filtering down by year and downloading it in smaller chunks.



The screenshot shows a table titled "Web of Science Documents" with a "Documents Per Page" dropdown set to 10. A red circle highlights a download icon (a downward arrow) in the top right corner. The table has the following columns: Article Title, Authors, Source, Research Area, Document Type, Volume, Issue, Pages, Publication Date, and Ti C. The first row of data is:

Article Title	Authors	Source	Research Area	Document Type	Volume	Issue	Pages	Publication Date	Ti C
Long-Term Effect of Gene Therapy on Leber's Congenital Amaurosis	Bainbridge, J. W. B.; Mehat, M. S.; Sundaram, V.; Robbie, S. J.; Barker, S. E.	NEW ENGLAND JOURNAL OF MEDICINE	MEDICINE, GENERAL & INTERNAL	Article	372	20	1887-1897	2015	:

Once you have this list downloaded, you can manipulate it in Excel to find details for individual papers, or to conduct further analysis. Bear in mind that not all the individual metrics quoted in the download are useful.

Interpreting metrics

The key metrics we recommend you use for groups of papers are the Category Normalised Citation Impact, and the share of documents in the top 1% or 10%.

- **Category Normalised Citation Impact (CNCI).** This is a relative citation ratio, comparing the number of cites this paper has received to the mean of all papers in the same field, same document type, and same publication year. A CNCI of 1 is equal to the world average; the overall CNCI for most UCL publications is around 2, varying by subject area.
- **% of documents in top 1%/10%.** This gives the relative share of papers that exist in the top 1% or top 10% of comparable papers by citation count. The list of comparable papers is defined as those with the same field, same document type, and same publication year.

The CNCI and “top share”, taken together, can give a good sense of where a paper fits in comparison to the rest of the field. We recommend looking at both to help avoid anomalies – for example, a group of papers with a very high CNCI but a more normal-looking share of papers in the top 10% probably indicates that there is one or two very highly cited papers, and the rest are more in line with the field average.

As noted above, InCites also offers a range of other metrics, normalised to some degree or another, which may be worth assessing. The **percentage of international collaborations** can be useful, but the expected level of collaboration varies dramatically by field, and so care should be taken using it at an institutional level. It is most meaningful when used to show developments over time, or when activity in one particular field is being considered. InCites offers a **percentage of industry**

collaborations, but due to the way they index authors, this only identifies major industrial or commercial collaborators and misses smaller organisations. It is thus unreliable as an overall figure and should be avoided.

The simple **number of publications per year** can be useful in some cases, but it is not normalised by field, and is directly linked to the size of the institution so may not be meaningful for direct comparisons. The **h-index of an institution** should always be avoided, as it is not possible to scale this by the field of research or the underlying size of the institution.

The **mean citation impact** (and **impact relative to world**) should be treated with caution as it is not normalised by field or year. It may be appropriate if the output for a particular discipline is being considered but should not be used institution-wide or over a multi-year period. The **percentage of documents cited** is likewise field-dependent and year-dependent.

For assessment of recent papers, it is important to bear in mind that the small number of citations involved can make statistics over a short period (the past 1-2 years) unreliable. In particular, for papers less than a year old, relative citation numbers can be very strongly affected by unrelated factors like the exact month of publication.

If you have not filtered by document type, be alert for individual anomalies. Including all records can produce problems for normalised metrics by introducing dramatic outliers – for example, documents classed as “editorial material” are very rarely cited. The vast majority will have zero citations, and the rare items that are cited will have improbably high normalised values – if the average is only one editorial in twenty being cited, then an editorial with just one citation can shoot up to a CNCI of 20, vastly more than most research articles will receive.

FAQS:

1. *I'm interested in looking at a UCL department, not UCL as a whole?*

InCites has a separate tool – InCites MyOrganisation – designed to give us analysis information on individual UCL departments. We have [a separate guide](#) on using it for departmental data.

However, it is not possible to get department-level data on other institutions – while this allows us to find data for the UCL geography department, for example, we would not be able to identify papers associated with the corresponding department at Imperial or Edinburgh. Depending on the subject, you could try filtering by subject area and comparing institutions based only on their papers in specific subjects. This may not work in all cases.

2. *This download includes the Journal Impact Factor and something called the "Journal Normalized Citation Impact" - what about those metrics?*

We recommend you **do not** use these. The Journal Normalised Citation Impact shows how a paper compares to others in the same journal, not the field as a whole; this is not very informative. The Journal Impact Factor is an averaged measure of citations for all papers in the journal, and does not necessarily indicate anything about that one paper. Research England has strongly discouraged use of the impact factor for REF selection, and UCL has committed to not using it. See the [UCL Metrics Policy](#) for more information.