Using SciVal Responsibly

A Guide to Interpretation and Good Practice

Authors: Ian Rowlands with input from the LIS-Bibliometrics Committee

Version 3.1, Date: January 2020

About this Guide

This guide is designed to help those who use SciVal to source and apply bibliometrics in academic institutions. It was originally devised in February 2018 by Dr Ian Rowlands of King's College London as a guide for his university, which makes SciVal widely available to its staff. King's does this because it believes that bibliometric data are *best used in context* by specialists in the field. A small group of LIS-Bibliometrics committee members reviewed and revised the King's guide to make it more applicable to a wider audience. SciVal is a continually updated source and so feedback is always welcome at LIS-Bibliometrics@jiscmail.ac.uk.

LIS-Bibliometrics is keen that bibliometric data should be used carefully and responsibly and this requires an understanding of the strengths and limitations of the indicators that SciVal publishes. The purpose of this Guide is to help researchers and professional services staff to make the most meaningful use of SciVal. It includes some important `inside track' insights and practical tips that may not be found elsewhere.

The scope and coverage limitations of SciVal are fairly widely understood and serve as a reminder that these metrics are not appropriate in fields where scholarly communication takes place mainly outside of the journals and conference literature. This is one of the many judgment calls that need to be made when putting bibliometric data into their proper context.

One of the most useful features of SciVal is the ability to drill down in detail using various filters. This allows a user to define a set of publications accurately, but that may mean generating top level measures that are based on small samples with considerable variance. Bibliometrics distributions are often highly skewed, where even apparently simple concepts like the `average' can be problematic. So one objective of this Guide is to set out some advice on sample sizes and broad confidence intervals, to avoid over-interpreting the headline data.

Bibliometric indicators should always be used in combination, not in isolation, because each can only offer partial insights. They should also be used in a `variable geometry' along with other quantitative and qualitative indicators, including expert judgments and non-publication metrics, such as grants or awards, to flesh out the picture.

Contents

QUANTITATIVE METRICS GOOD PRACTICE

The below advice should be considered before using SciVal or any bibliometric indicators.[1].

The principles of the Leiden Manifesto[2] and DORA[3] should be referred to before undertaking any quantitative bibliometric analysis. Users in an institution with a responsible metrics policy or that is a signatory of DORA will be expected to uphold these standards.

We would recommend that you always try to provide multiple indicators, rather than standardising on just one. Giving multiple indicators helps reduce the effect of outlier papers affecting the calculations, while also allowing different aspects of research quality to be assessed.

Essential advice for selecting and interpreting indicators

- Use at least two indicators in any metric-based assessment.
- Avoid using journal-based indicators (such as impact factor or journal rank quartiles) to assess the quality of individual papers.
- Citation-based indicators are time-dependent and thus must only be used to compare articles of similar ages. Likewise, citation-based indicators are discipline specific and should not be used to compare articles of different subject disciplines.
- Using discipline filters (SciVal's Subject Area, Topic or Topic Cluster) might not give a full picture of multidisciplinary research or work in newly-emerging disciplines.
- Any indicator calculated by a simple average will be susceptible to being skewed by outliers. Check for outliers in your dataset and consider whether they might skew the indicators you wish to use.
- Accept that the pre-set SciVal metrics might not, in every case, match your analysis need. If there are not suitable metrics, users could consider exporting data for analysis in a different programme. Similarly, avoid selecting metrics on the basis that they may fit the conclusions you are expecting to draw.
- Avoid false precision – while SciVal quotes many metrics to two decimal places, this can give a false sense of accuracy, and very small differences between groups or over time are unlikely to be meaningful. Consider rounding values when interpreting them.
- When presenting the data, include the date of either the most recent SciVal update and/or the date of which you collected the data, alongside any filters applied to the source data.

---

[1] Adapted from Gray, A.; Price, R. Using InCites Responsibly. 2020 3 https://doi.org/10.25561/75946
[2] Hicks, D.; Wouters, P.; Waltman, L.; de Rijcke, S.; Rafols, I. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* (2015) 429-431 doi:10.1038/520429a
[3] https://sfdora.org/read/

AWARDED GRANTS

Definition

Annual data for selected major funders, showing the number and value of awards made to an institution.

Key strengths	Key weaknesses
Useful for high level benchmarking of major funding support between similar institutions.	SciVal's coverage of funders is limited and excludes some major sources , including National Institute for Health Research (NIHR) and many charitable sources.
Useful for tracking institutional trends in terms of grant capture over time.	Levels of external funding vary greatly across disciplines.

Interpretation

Awards Volume does not take account of price inflation and is recorded in US dollars. Both volume and number of awards are linked to a single institution only - the one where the Principal Investigator was located at the time of the award. No account is taken of other beneficiaries (co-investigators at other institutions), which of course cuts both ways. Awards Volume considers aggregated values of awards over the award lifetime. In other words, it considers the total value awarded at the time of award and not the value (to be) spent in any particular time period.

For these reasons, the figures from SciVal will differ, perhaps considerably, from those reported internally within the institution.

Practical insights

As of January 2020, SciVal provides data for the following funders (from 2009 onwards):

UK
- Wellcome Trust (WT)

- Arts and Humanities Research Council (AHRC)
- Biotechnology and Biological Sciences Research Council (BBSCR)
- Economic and Social Research Council (ESRC)
- Engineering and Physical Sciences Research Council (EPSRC)
- Medical Research Council (MRC)
- Natural Environment Research Council (NERC)
- Science and Technology Facilities Council (STFC)

US
- National Institutes of Health (NIH)
- National Science Foundation (NSF)

Australia
- Australian Research Council (ARC)
- National Health and Medical Research Council (NHMRC)

SciVal permits analysis of award data using subject fingerprints derived from the abstract of the successful application.

Good practice tips

• It is essential to be clear about the limitations outlined above, and to record these.

• SciVal awards data cannot be linked to outputs or organisational structures such as faculties.

SCHOLARLY OUTPUT

Definition

Number of outputs during a specific publication time window.

Key strengths	Key weaknesses
An objective proxy for research production.	Despite large differences in numbers of publications between disciplines, Scholarly Output is not field- normalised.
A useful tool to identify entities that are insufficiently research-active in comparison with peers in the same field.	Collaboration and multi-authorship make it difficult to assess the contribution of an institution, group or individual to a team effort.

Interpretation

SciVal uses whole publication counts, so an output jointly-authored by researchers from two different institutions is credited to both institutions with a value of 1 each. This means that many institutional comparisons are beset by issues of multiple counting. SciVal does not currently support more sophisticated indexes that take multi- authorship into account using fractional methods[4]. Counting conventions matter because they frame how we should interpret the numbers. Whole publication counts are best interpreted as a measure of *participation* in science; fractional counts as a measure of *contribution*.

Practical insights

It is possible to identify `kilo' papers with more than 10, 50 or 100 authors from SciVal publication lists and to limit further analyses by excluding these, if this is appropriate.

---

[4] See Roberto Todeschini and Alberto Baccini, *Handbook of Bibliometrics Indicators.* Weinheim: Wiley Verlag, 2017: 54-74 for a detailed discussion of alternatives to whole publication counts.

Scopus and SciVal index some, but by no means all, journals cover-to-cover, thereby including non- substantive content such as letters to the editor, errata, obituaries, etc., on a non-representative basis.

Good practice tips

- We recommend that editorial material should be excluded from any analysis.
  - Consideration should be given to which publication types should be included (e.g. conference papers), and this limitation must be stated.
- It *may* be sensible to use SciVal's new features to screen out 'kilo' papers with extensive author lists, but this approach needs to be justified (and stated).

CITATION COUNT

Definition

Number of citations received during a specific citation time window.

Key strengths	Key weaknesses
A reasonably objective proxy for research impact, in the limited sense of its utility for actively publishing researchers.	Strong bias towards journals and conferences as a source of citations, so an inappropriate measure in fields where scholars mainly communicate through other channels.
The diffusion of citations across time, fields and geographies can offer unique insights into structural processes within science.	Citation rates vary in different areas and this indicator is not field-normalised. Outputs also take much longer to reach citation maturity in some fields.

Interpretation

Strictly speaking, we should use the term `reference counts' not citation counts, because neither Scopus nor SciVal record the number of times a reference is actually invoked within the citing work. The justification for using citations as a proxy for `quality', `impact' or `excellence' is somewhat questionable in terms of construct validity. There is however considerable evidence that citation counts correlate with peer review judgments sufficiently in some fields to justify their use in research evaluation[5]. The extent to which citation counts are a credible measure of impact is best judged by experts who have an understanding of citation norms and behaviours within their field or specialty.

---

[5] For example, for REF2020 Research England sanctioned the limited use of citation data for the following Sub-Panels:

Sub-panel 1: Clinical Medicine
Sub-panel 2: Public Health, Health Services and Primary Care
Sub-panel 3: Allied Health Professions, Dentistry, Nursing and Pharmacy
Sub-panel 4: Psychology, Psychiatry and Neuroscience
Sub-panel 5: Biological Sciences
Sub-panel 6: Agriculture, Food and Veterinary Sciences
Sub panel 7: Earth Systems and Environmental Sciences
Sub-panel 8: Chemistry
Sub-panel 9: Physics
Sub-panel 11: Computer Science and Informatics
Sub-panel 16: Economics and Econometrics

Practical insight

Scopus is updated with publication and citation data more regularly than SciVal. SciVal is typically always 1-2 weeks behind Scopus data and as a result, Scopus may record more citations at a given point in time.

Good practice tips

• When reporting Citation Count, the source database and the date on which the counts were made (Scopus)or the date of the last update (SciVal) should always be included.
• We generally recommend that self-citations are be included in all analyses. There are very legitimate reasons to cite one's own work, and we are not clear precisely how the exclusion filter actually works.
• Only compare Citation Count within the same publication year and within the same field.

.

CITATIONS PER PUBLICATION

Definition

The number of citations received to specific time window by a set of outputs divided by the number of those outputs.

Key strengths	Key weaknesses
A size-independent indicator that reveals citation impact relative to the number of outputs.	Caution needs to exercise to compare like-with-like on the basis of outputs with similar age and subject characteristics.
Facilitates benchmarking of broadly similar entities.	The mean of a skewed distribution can be misleading due to statistical outliers.

Interpretation

Citations to outputs are not normally distributed, like human height, with most observations clustering around the mean. In contrast, most outputs tend to attract few citations while a small minority receive very many. You should therefore expect to find statistical outliers in your sample and these will `artificially' inflate the mean. SciVal does not provide confidence intervals for this indicator but, as a rule of thumb, it is probably sensible to avoid drawing firm conclusions on samples of less than 500 outputs.

Practical insight

SciVal provides the number of times an output has been cited *from its publication to the present.* Older publications tend to perform much better on this indicator for obvious reasons.

Good practice tips

- Only compare Citations per Publication *within* the same publication year and *within* the same field.

- Be careful when making comparisons based on samples of fewer than 500 outputs, without first screening for extreme statistical outliers.

- As with Citation Count, consider whether it is justifiable to exclude self-citations

FIELD-WEIGHTED CITATION IMPACT

Definition

Field-Weighted Citation Impact (FWCI) is the ratio of the actual number of citations received by an output to date and the `expected' number for an output with similar characteristics. `Expected' refers to average citations (over the previous three years) for all outputs of the same age, document type and field. Where a journal appears in more than one field category, each field contributes equally to the calculation.

Key strengths	Key weaknesses
FWCI corrects for differences in citation rates attributable to age, document type and subject differences in citing behaviour.	Field delimitation is based on a journal-level subject classification and may not accurately reflect the actual content of the output.
FWCI considers a great deal of information and effectively benchmarks an output against its immediate peers in the world literature.	FWCI is poorly designed because it uses the mean to characterise heavily skewed citation distributions.

Interpretation

FWCI is a useful indicator for benchmarking purposes because it anchors an output's performance relative to its peers in the scientific literature. An output with an FWCI value of 1 can be interpreted as having `world average impact'. Like many other citation indicators, the distribution of FWCI is highly skewed and this property makes it a useful tool for spotting extreme outliers: papers that have perhaps been cited 1,000 times more than expected.

On the other hand, if you want to use mean FWCI to compare two sets of publications, this indicator (a ratio of a ratio based on inappropriate use of the mean) suffers from poor design. FWCI is best used to benchmark large publication sets, such as whole institutions, where the effects of statistical outliers are dampened. The table below offers some indicative guidance on the appropriate interpretation of mean FWCI at different scales.

Stability intervals and interpretation

20,000	Norway	+/- 3.3%	2.20 [CI95 2.13-2.27]
15,000	Greece	+/- 3.8%	2.20 [CI95 2.12-2.28]
10,000	University of Cambridge	+/- 4.2%	2.20 [CI95 2.11-2.29]
5,000	King's College London	+/- 7.4%	2.20 [CI95 2.04-2.36]
2,500	University of Leicester	+/- 8.8%	2.20 [CI95 2.01-2.39]
1,000	Open University	+/- 11.9%	2.20 [CI95 1.94-2.46]
500	De Montfort University	+/- 14.6%	2.20 [CI95 1.88-2.52]
100	Large research-active department	+/- 24.8%	2.20 [CI95 1.65-2.75]
50	A small research-active department	+/- 65.1%	2.20 [CI95 0.77-3.63]

For more information on how these estimates were derived, please see:

Rowlands, Ian. (2017) SciVal's Field weighted citation impact: Sample size matters! *The Bibliomagican [Blog]*. 11 May 2017. https://thebibliomagician.wordpress.com/2017/05/11/scivals-field-weighted-citation-impact-sample-size-matters-2/

Unlike the Citation Count for an output, which can only remain the same or increase, the value of FWCI may fluctuate up or down over time as the citation rate for a given output speeds up or slows down relative to the literature[6].

Practical insights

SciVal recommends that FWCI values should not be taken too seriously for recent papers, because citations take time to settle down. As a rule of thumb, FWCI values should be not be used for the two publication years preceding the April of the following calendar year. So, as an example, only from April 2018 one should consider outputs published during 2016 or earlier.

After four years, SciVal stops recalculating FWCI and the other metrics at each batch update and the values are then permanently locked down. This limits the value of the indicator for older outputs, whose citation impact may be sustained over many years.

Good practice tips

• Despite the fact that FWCI normalises for field differences in citation rates, this is an approximate art, and it makes little sense to compare mixed and diverse publication sets across very different

---

[6] For further reading on the volatility of the FWCI, see Rowlands, I. *Six weeks is a long time in bibliometrics@ Stability and Field-Weighted Citation Percentile. The Bibliomagician [Blog].* 21 November 2019. https://thebibliomagician.wordpress.com/2019/11/21/six-weeks-is-a-long-time-in-bibliometrics-stability-and-field-weighted-citation-percentile/

fields (comparing, say, King's and LSE at an institutional level). Comparisons should be made on the basis of similar research mission and scale (comparing, say, King's and Imperial outputs in Biochemistry, Genetics & Molecular Biology).

- Average FWCIs should be interpreted within the confidence intervals indicated set out in the table above.

- FWCI has limited value for outputs more than four years old and should not be used for very recent outputs.

OUTPUTS IN TOP CITATION PERCENTILES

Definition

Counts the number (or percentage) of outputs that fall into the top 1%, 5%, 10% or 25% of the world's most highly cited outputs. Field-weighting may be applied optionally.

Key strengths	Key weaknesses
A highly intuitive indicator that takes account of a lot of contextual information (age, document type and, if desired, field).	SciVal rarely uses the full scale (1-100) and percentile values for very recent papers may be seriously misleading.
Focuses attention at the top end of the impact distribution ('highly cited papers')	May lose a sense of absolute differences in citation impact and therefore undervalue extraordinary outputs (FWCI is more useful here).

Interpretation

SciVal's Outputs in Top Percentiles (OTP) is a highly intuitive indicator and one that lends itself well to comparative analysis because, at least in theory, it uses a common scale (1-100) unlike FWCI whose values can range over many orders of magnitude.

The same strictures that were voiced for FWCI concerning recent outputs apply to OTP. If anything, the health warning needs to be amplified. At the time of writing, an *uncited* review paper from the current year in emergency medicine actually appears at the third percentile. This is because the OTP depends on being able to divide the publications into 100 percentiles, and this level of division is not always possible: especially for items just published. The issue is exacerbated in this particular case because there happen to be relatively few review papers in this specialty. While the metric is `correct' it may prove to be hopelessly optimistic in the longer run.

Details of the methodology used to calculate Outputs in Top Percentiles are not in the public domain, but our observation is that SciVal rarely uses the full 100-point scale. As a result, it is good practice to include Citation Count alongside this indicator.

As is the case for FWCI, the value of OTP may rise or fall over time. This of course means that a marginal output may fall either side of a hard threshold, such as the top 10%, as its OTP value changes over time.[7].

Practical insights

Outputs in Top Percentiles uses the convention that 1 is the highest impact percentile. In contrast, and a source of endless confusion, Scopus provides a `reverse percentile' where 99 represents the highest impact. The two indicators are created using different citation time windows (12 months for SciVal, 18 for Scopus). It is therefore very important to be clear in any communication which indicator is being used.

SciVal reports OTP as an integer value. For outputs in the current year, SciVal records a null value until the first data snapshot on or after 1 July.

Good practice tips

• Always state whether this metric was collected using SciVal or Scopus and give the dates.

• Always provide Citation Count alongside Outputs in Top Percentiles.

• We do not recommend that this indicator should ever be averaged: far better to compare entities on the basis of the number and proportion of outputs at the standard percentile thresholds: 1%, 5%, 10% and 25%.

• We generally recommend that the option to apply field-weighting should be applied, because fields with high citation rates will otherwise have a significant advantage. The document type(s) chosen will also affect the results.

• The same concerns around self-citation, very recent and much older papers apply here as for FWCI.

---

[7] Rowlands, I. (2019) Six weeks is a long time in bibliometrics: Stability and Field-Weighted Citation Percentile *The Bibliomagician. [Blog].* 21 November 2019. https://thebibliomagician.wordpress.com/2019/11/21/six-weeks-is-a-long-time-in-bibliometrics-stability-and-field-weighted-citation-percentile/

H-INDEX

Definition

The value of the *h*-index refers to the number of outputs, *h,* in a bibliography that have each been cited by at least *h* other outputs. For example, an author claiming an *h*-index of 20 must be able to demonstrate 20 papers that have been cited 20 or more times each.

Key strengths	Key weaknesses
Attempts to demonstrate both quality and quantity in a single convenient measure.	Attempts to demonstrate both quality and quantity, but lacks construct validity with regard to either.
Is insensitive to both very highly cited outliers and poorly cited papers.	Accurate determination of *h* requires a complete bibliography and its value is significantly affected by length of career and field, which are respectively difficult and near impossible to normalise for.

Interpretation

SciVal generates its *h*-index over a researcher's whole published record and so it should not be used to infer differences in performance between people at different stages in their career. It has also introduced an *h5* index that only considers outputs published over the past five years.

Field-normalisation is not available for any of these indicators and this seriously limits their value: this indicator must *never* be used for comparisons between different specialties. It can be shown, for example, that if you take account of differing citation rates between fields, then a mathematician with an *h*-index of 20 is of equivalent standing to a microbiologist with an index of 58[8]. As a result, the *h* family of indicators must be used, if at all, with a great deal of caution.

One undesirable property of the *h*-index, if understood as a performance metric, is that it is monotonic: its value can rise or remain the same, but it can never decrease. For some authors, *h* will continue to grow for many years after they have ceased to be research-active as their older papers

---

[8] Juan E Iglesias and Carlos Pecharroman (2007). Scaling the *h*-index for different scientific ISI fields, *Scientometrics* 73(3): 303-320.

gradually accrue citations and pass the threshold. Like many citation indicators, this can significantly discriminate against women[9].

Once popular, the *h*-index can no longer be regarded as a useful indicator: it purports to do the near impossible: reducing quality and quantity to a single number. Its method is dubious[10], not least because *h* discards a great deal of relevant information: ignoring both very highly and poorly cited papers, for example.

The validity of the *h*-index is increasingly being challenged, on both empirical and theoretical grounds[6]. The clearest evidence of why *h* does not measure what it claims to can obtained easily using a pocket calculator. Look up the total number of citations to date for an author and take half the square root: this will almost certainly give you their *h*-index or at least a very close approximation. So *h* is best regarded as an estimate of total impact, not as a nuanced performance measure that trades off quantity and quality of outputs. In which case, why not just use Citation Count as that measure, not a convoluted alternative[11]?

Practical insight

SciVal calculates its *h*-index from 1996 onwards, so more senior researchers may find their *h*-index does not take account of earlier papers. Missing even a small number of highly cited papers will have a direct effect on the *h*-index, and in this respect it really makes for hard work.

Good practice tips

• Urgently consider alternatives: the h-index is difficult to work with, it does not measure what it claims to, and its credibility is increasingly under attack in the literature.

• If you must use the h-index, you need to be absolutely sure that your list of publications is complete, and then only compare researchers working in very similar areas and at a similar point in their careers. You should also consider any career breaks that may have affected a researcher's publication trajectory.

• Look at the researcher's outputs by publication year to spot any possible career breaks.

• Never, ever, calculate or present average h values.

---

[9] Geraci, L., Balsis, S. & Busch, A.J.B. Scientometrics (2015) 105: 2023. https://doi.org/10.1007/s11192-015-1757-5
[10] See Cameron Barnes (2016) The construct validity of the *h*-index. *Journal of Documentation* 72:878-895 for a devastating critique.
[11] Rowlands, I. Is it time to bury the *h*-index? *The Bibliomagician. [Blog].* 23 March 2018. https://thebibliomagician.wordpress.com/2018/03/23/is-it-time-to-bury-the-h-index/

PUBLICATIONS IN TOP JOURNAL PERCENTILES

Definition

Publications in Top Journal Percentiles (PTJP) is a journal-level indicator, unlike Outputs in Top Percentiles which, as we saw earlier, measures the citation impact of individual outputs. PTJP is expressed as the number (or proportion) of outputs which are published in the most heavily cited journals within a field - irrespective of their individual citation impact. The thresholds for journal impact are the same as for article impact: top 1%, 5%, 10% and 25%.

Key strengths	Key weaknesses
A useful indicator for understanding past publishing behaviour and for developing and monitoring future publication strategy.	The citation impact of a journal must never be used as a proxy for the value of a given paper in research evaluation: but the two are often conflated.
Provides a very different perspective than article-level measures and extends the bibliometrics vocabulary.	The relationship between a journal's citation impact and the rigour of its peer review process is unclear.

Interpretation

An output in a top journal may remain uncited for many years, so beware the ecological fallacy that publication in a high impact journal necessarily means that an output must be highly cited as well.

Publications in Top Journal Percentiles is a confusing indicator because SciVal offers a choice from *three very different approaches* for determining which are the top journals in a field: CiteScore, Scimago Journal Rank (SJR) and Source Normalised Impact per Paper (SNIP). These should not be confused with the classic journal impact factor which is proprietary to another organisation: Clarivate Analytics.

Practical insight

CiteScore, SJR and SNIP correlate rather strongly with each other (rank-ordered coefficients in the range 0.74 to 0.90) but the fact remains that each method yields different lists of high impact journals.

We cannot recommend one approach over another at this stage, because we do not and cannot know which is `the best'. Our advice is to try out all three and then form a judgment based on the range of outcomes and look at the journals that are represented in the top categories.

Good practice tips

• Publications in Top Journal Percentiles should only ever be used in the context of thinking about publication strategy, never as a proxy for the impact of individual outputs.

• Try out all three methods and look at the range of outcomes.

• We recommend using the percentage option for comparative purposes

COLLABORATION

Definition

SciVal defines four mutually exclusive levels of research collaboration based on co-authorship patterns:

- **International** (co-authors from more than one country)
- **National** (co-authors from more than one institution, in the same country)
- **Institutional** (more than one author, in the same institution)
- **Single author** (no collaboration)

A fifth category, which cuts across these, is Academic-Corporate Collaboration: any output with co-authors from an academic institution and a non-academic body such as government, a charity or the private sector.

Key strengths	Key weaknesses
Focuses attention on research as a collective activity involving external partners.	Informal modes of collaboration remain invisible.
Indicates the degree to which a set of authors are integrated into national or international scientific networks.	These measures do not tell us anything about specific roles or contributions.

Interpretation

Collaboration is a very important aspect of modern research and it provides essential context in research evaluation. It is not an end in itself, however, hence our description of Collaboration as a process measure, not a measure of research impact. This categorisation of outputs by collaboration level can be used on their own as numbers or preferably percentages to get a better understanding of the size and shape of collaborative activity. The levels may be combined with a measure of research impact (e.g. Citations per Publication) in SciVal's Benchmarking module.

Practical insights

Depending on the context of the research question, it may or may not be appropriate to make use of the functionality in SciVal facility to filter out `kilo' papers (see p. 5).

An issue that is particular important to understanding Academic-Corporate Collaboration is that SciVal only indexes author's host institutions once they reach a critical mass of outputs. Many non-academic institutions publish few papers, so this indicator may well understate the extent of an author or institution's networks with industry, government and the professions.

Good practice tip

• Apply the field-weighted option if your set of outputs is diverse (for example, across the whole of your organisation).

• It is generally more instructive to use percentages rather than numbers of outputs.

ALL SCIENCE JOURNAL CLASSIFICATION

SciVal uses a journal level subject classification as the basis for its field normalisation. The default is the All Science Journal Classification (ASJC) whichcomprises five broad clusters, 27 major fields, and 334 minor fields. Multiple classifications per journal are possible and in these cases, the field normalisation applies equal weighting to each field.

ASJC categories do not map at all well onto departmental or faculty structures because academics do not limit their publishing behaviour in this way, so the obvious temptation to use these categories for cross- institutional benchmarking should be tempered.

The field names and codes are provided by Elsevier here:
https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/

COMPARING METRICS IN SCIVAL AND INCITES

There are multiple providers of research data and analytics. A rival product to SciVal is Clarivate Analytic's InCites.[12] The purpose of this section is not to compare the two as overall products but to provide guidance at the individual metric level[13].

When thinking of where equivalents exist or don't exist between two providers of metrics, the Snowball Metrics initiative[14] can be useful. Where either InCites or SciVal declare they use a Snowball Metrics-compliant methodology for an indicator it will be indicated with an (*) below.

InCites indicator	SciVal equivalent	Comments
# Web of Science Documents *	Scholarly Output*	Both will include journal publications (including various article types) as well as books and book chapters. As of June 2019 InCites update, InCites indexes close to 60.5 million records and SciVal 74 million in the same month. Both data sources come from each's respective primary databases, Web of Science Core Collection and Scopus.[15]
Times Cited*	Citation Count*	
Citation Impact*	Citations per Publication*	See above.
Category Normalised Citation Impact (CNCI)*	Field Weighted Citation Impact (FWCI)*	SciVal state that an output's FWCI is 'locked' after four years. Follow the advice in the CNCI indicator section of this document for details on how to access 'locked' values in InCites
Journal Normalised Citation Impact	n/a	
% of Documents in Top 1% and % of Documents in Top 10%*	Outputs in Top Citation Percentiles*	
Percentile in Subject Area	n/a	Can be considered as near-similar to Outputs in Top Citation Percentiles with 'Show as field-weighted' included as described above.
Impact Relative to World	n/a	
Highly Cited Papers	n/a	
Hot Papers	n/a	

% Percentage of Documents Cited	n/a	
International Collaborations	Collaboration Impact*	SciVal presents Collaboration Impact in a % and broken down to international, national and single authorship. Then you must select each of these 'levels' of collaboration in Scholarly Output to view the publications by individual country.
% of Industry Collaborations	Academic-Corporate Collaboration*	
h-index	h-index*	

Disclaimer

SciVal® is a registered Trademark of Elsevier Properties SA. This guide is not endorsed by Elsevier, and the compilers do not warrant the accuracy or completeness of the information provided. Our intention is to offer guidance on the responsible use of metrics contained in SciVal, their meaning, application and limitations.

About the Authors

Dr Ian Rowlands was formerly the Research Information & Intelligence Specialist at King's College London.

About LIS-Bibliometrics

LIS-Bibliometrics was formed in 2010 to support Librarians and other practitioners in the effective use of bibliometrics for research evaluation. Join the discussion list at https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=LIS-BIBLIOMETRICSand follow The Bibliomagician blog at https://thebibliomagician.wordpress.com/.