Bibliometrics Breakdown: An Insider's Insight

Do you know your h-index from your m-quotient? Quantitative analysis of the scientific literature has changed dramatically over the last decade and bibliometrics has become an important aspect of the assessment of your scientific productivity. Despite there being a proliferation of bibliometric parameters, their actual meaning and how they might be used is not generally well understood.

Familiarity with bibliometric parameters can aid you in the effective quantitative assessment of publications generated by individual researchers and research teams as well as journals. We offer some insights from the Niche team on how you may interpret these parameters and employ them effectively.

Before you start

Existing and emerging bibliometric tools are frequently used to provide an assessment of the value of scientific research. These give you a quantitative score for author's, research groups and publication vehicles.

Journal ranking is performed to provide a quantitative measure of a journal's performance in its field. Of the various bibliometric parameters measured, perhaps the most frequently cited is the Journal Impact Factor.

The impact factor takes into account all indexed citations received by a given journal divided by the number of 'citable' articles published by a journal in the previous 2 years.

The h-index is the most frequently used bibliometric parameter to assess an author's scientific contribution. The h-index is considered to be a more robust assessment of scientific contribution than metrics based on simple publication counts.

Prepare to succeed

As different fields can have vastly different publication and citation rates, it is not appropriate to use the impact factor to make a quantitative comparison of journals from different fields.

The Impact Factor is imprecise and subject to bias and manipulation. A journal's score is an average, it says little about the quality of any single piece of research. It continues to be the most widely used metric.

One widely cited article appearing in a journal can artificially inflate that journal's score, even if it is a controversial article where it may then be cited in a host of criticism.

The h-index can be influenced by self-citation, which accounts for a significant portion of all citations. These may occur as a consequence of cumulative individual research, a need for personal gratification or an attempt to increase the author's scientific visibility.

Key Insights

Over the last decade a new field of study has exploded onto the scientific landscape – that of bibliometrics. Qualitative analysis of the scientific literature is changing rapidly with the creation of new evaluation tools, parameters and normative data. These parameters can be categorised into author or journal focused metrics. We recently highlighted the potential use of these data to grade research in our Insider's Insight into the identification and profiling of key opinion leaders [1].

It is now possible to generate a simple 'number' that can be used to give a 'value' for the contribution that individual researchers have made to the scientific literature itself and, by association, their scientific standing. In addition to the more traditional publication count and number of citations, we now have the h-index, m-quotient, h-index, e-index, g-index and i-10 index to assess authors. These values were contrived in an attempt to overcome the well-recognised imperfection of the traditional publication count and number of citations.

There are also several quantitative parameters that can help measure the academic strength of scientific journals [Table 1]. The Journal Impact Factor (JIF; 2-year) is the most widely known of these parameters. Others include the Eigenfactor, article influence score (AIS), SCImago journal rank (SJR) and source-normalised impact per paper (SNIP). Each of these parameters has strengths and weaknesses from both theoretical and practical standpoints.

Abbreviations

AIS – Article influence score
PoP – Publish or perish
SJR – SCImago journal rank
SNIP – Source normalised impact per paper
WOS – Web of science
RG – ResearchGate
JIF – Journal Impact Factor

In a recent Insider's Insight into targeted journal selection, we provided some guidance on the utility that these parameters can have for tasks such as targeting the right journal for your publications [2].

It is immediately obvious that publication metrics fail to give a true assessment of research 'quality', whatever that may be; it is not yet possible to derive a simple indication of the scientific contribution that a piece of work may make. You still have to determine that for yourself.

Bibliometric Tools

The term bibliometrics was coined by Alan Pritchard in a paper from 1969 [3]. He defined the term as "the application of mathematics and statistical methods to books and other media of communication". Although not new, the need for manual calculation in the past made citation analysis time-consuming requiring a significant commitment. For better or worse, automated algorithms have made bibliometric parameters more readily available. The first computer algorithm for automated citation extraction and indexing was created by Cite Seer. Various options are now available via on-line tools that can be used to generate bibliometric data.
A few are described here:

PubMed

Created by the National Library of Medicine and launched in 1997 as a freely available interface to the MEDLINE database, PubMed has become one of the most popular and widely used search engines for the medical literature. Articles are confined to biomedical and life science journals and there is very little actual analysis of the citations.

Launched in 2004 by Elsevier, Scopus [4] is the largest online bibliometric database and includes information on journal articles from all major disciplines published from 1966 onward, these include articles from the social and physical sciences that do not appear on PubMed.

Scopus

Citation analysis is more robust with Scopus than with PubMed and is available for articles published after 1996; there are plans to extend archiving back to 1970. A unique advantage of the Scopus database is individual author identification, whereby articles by an author are grouped on the basis of affiliation and co-authors. Unlike PubMed, Scopus is not free to users.

WEB OF SCIENCE™

Although the Thomson Reuters WOS database contains fewer articles than Scopus, it includes articles dating back to 1900 [5]. Slightly fewer disciplines are covered by WOS than Scopus. It includes a robust citation analysis, although a recent study identified a fifth more articles following a citation analysis comparing Scopus and WOS; WOS comes into its own when you want to evaluate research reported before 1996. Like Scopus, access to WOS is not free to access.

Google Scholar is a free database that offers an impressive search capacity. It is most likely the best way to access obscure information, such as articles published in journals that have yet to be indexed in other databases. Google Scholar includes citations from books, online sources and conference proceedings.

Google scholar

Although there is very little in the way of citation analysis or author identification it will create a profile for an author and provide a list of their articles. Google Scholar can calculate the h-index and i-10 index for that author as well as summarise citations over the last 5 years.

ResearchGate

ResearchGate might be described as a social networking site for scientists and researchers. Members have a user profile and can upload research output including papers, data, book chapters, negative results, patents, research proposals, methods, presentations and software source code. ResearchGate publishes values for h-index (and a h-index excluding self-citations) as well as providing a citation impact measurement in the form of an 'RG Score'. RG Scores correlate with existing citation impact measures but have been criticized as the way it is calculated is unqualified.

Author Evaluation

Counting the number of publications provides the simplest bibliometric parameter. Usually this only involves counting peer-reviewed articles from journals listed in at least one of the journal search databases – it may include case reports but not book chapters, editorials or opinion pieces. Calculation doesn't take authorship position into account or quality of the journal. No distinction is given to ground-breaking work over less impactful articles. Articles appearing in 'poor' journals are counted equally as those in 'top' journals.

Citation counting often weighs articles based on their influence on subsequent publications. This approach has the limitation that, for example, widely read educational or otherwise informative articles may not receive citation in the literature that might reflect its value. The same can be said for case reports that provide important guidance for rare conditions only encountered infrequently by physicians and most likely only receive a few citations. Neither does it differentiate between positive or negative citations. For example, an article that is often cited in a critical manner, as a result of some inherent error or deficiency, would receive an excellent citation score despite its low value. Self-citation by authors is also counted. Although this practice may be appropriate when an author builds on their previous work in a scientific field of study, the system is vulnerable to abuse – whether accidental or deliberate.

h-index

Proposed to counter the recognised limitations with publication and citation counts, the h-index attempts to encompass in one value the quality and quantity/impact of a given author's scientific contribution based on their most frequently cited articles. The h-index equates to the point where the number of citations of an author's work equates to the number of publications (see Figure 1).

It tends to 'normalise' data in that articles with particularly high or low citation counts are given less weight. In addition, incremental increases in the h-index become progressively more difficult to achieve. Thus, it is more difficult for an author with a h-index of 10 (10 articles with 10 citations) to progress to a h-index of 11 (11 articles with 11 citations) than it is for an author with a h-index of 2 to progress to an h-index of 3. The h-index is a more robust assessment of scientific contribution than metrics based on simple publication counts and has become somewhat utilitarian. However, it assumes a level playing field in that the author on any specific manuscript will not influence its citation irrespective of its academic contribution. The h-index can also be influenced by self-citation, although it is deemed to be more difficult for authors to impact significantly on their own h-index score.

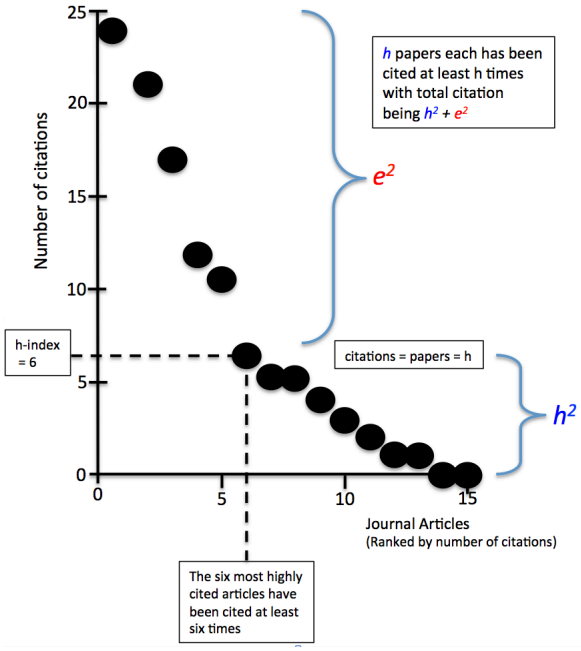

Figure 1: h-index calculation

The contemporary *h*-index (h_c-index)

Proposed as an alternative to the h-index, it incorporates a time-weighted aspect where older published articles are given less weight [6]. The h_c-index is calculated by multiplying the citation for an article by four and then dividing the number by the number of years since the article was published – these data are used in the ranking. Thus, the citation count for articles published in the current year would be multiplied by four, whereas that of an article published 6 years ago would be multiplied by four and divided by six.

Author Evaluation (cont.)

m-quotient

This is a variant of the h-index. It is calculated by dividing an author's h-index by the number of years since their first publication [7]. It reflects how the average amount of an author's h-index has increased over their publishing career and can, for example, be used to differentiate between two authors with similar h-indexes but who may have different career lengths.

e-index

Since incremental increases in the h-index becomes progressively more difficult, additional citations of articles that constitute the h-index do not get recognised. The e-index is an adjunct parameter that summarises the excess citations on top of those not counted in the h-index. Thus, once the articles (h) that make up the h-index are identified and h^2 used as the calculating parameter, the remaining number that are considered excess citations are used to calculate the e-index where: $e = (\text{total citations} - h^2)^{0.5}$.

g-index

The g-index is defined as a number such that the top 'g' articles are cited an average of 'g' times (or are cited g^2 or more times). Whereas, the e-index attempts to complement the h-index by addressing excess citation beyond 'h' (citations ignored by the h-index) the g-index includes all citations for the top 'g' articles [9].

i-10 (i-n) index

The number of publications that have been cited 10 or more times which thus ignores less 'important' work. This approach can be modified to an 'i-n' index where n can be any value – such as i-5 or i-100 which could be used to differentiate between junior or senior authors or departments, respectively.

Self citation

Authors citing their own work accounts for a significant portion of all citations (approximately 1 in 15 citations of articles in high-profile general medicine journals [8]).

These self-references may result from the cumulative nature of individual research, the need for personal gratification or the use of self-citation as a rhetorical and tactical tool in the struggle for displaying or establishing scientific authority.

There is no actual penalty for regular self-citers – the effect of self-citation remains positive even when an author is recognised as having an unusually high rate of self-citation. Studies most vulnerable to self-citation tend to be those with more authors, small sample sizes and conducted in the fields of cardiovascular medicine or infectious disease.

the missing metric

The current system that we use to maintain the process of scientific publications relies on a service that recesses little recognition. Those who review manuscripts and provide guidance on how to best to prepare manuscripts for publication get no compensation or citeable academic recognition for the application of their time or expertise. The R-index has been proposed as a simple way to quantify a scientists efforts as a reviewer [10].

Publish or Perish

Publish or Perish (PoP) is a software program that can be used to retrieve and analyse academic citations [11]. It can calculate an author's h-index, g-index, and e-index, as well as many other bibliometric parameters.

Researchers can use PoP to determine the impact of their research. It primarily uses the Google Scholar database and is free of charge for personal non-profit use. Authors can search for their articles by author name, just as they would with Google Scholar. Identifying articles to be used in a PoP calculation can be time consuming for authors with the same name as another author and/or one with a large number of published articles.

Journal Ranking and Evaluation

Journal ranking provides a quantitative measure of a journals performance in its field. There are many different factors that influence a journal's ranking and a summary of the various bibliometric parameters is provided in Table 1. It is important to remember when conducting a measure of impact that it is not an exact science. As such, bibliometric values should be viewed with a 'pinch of salt' and used with an understanding of their limitations. For example you should consider that:

- Metrics measuring journal impact cannot necessarily be compared across different subject areas as they have different citation rates and behaviour
- Review articles often attract more citations irrespective of their quality
- Good quality articles and journals often go uncited
- Colleagues may cite friends to increase their visibility
- The quality of a particular article cannot necessarily be judged by the journal it is published in

Table 1: Summary of journal bibliometric parameters

Journal Impact Factor (2 yr or 5 yr)	Although the most widely known and used metric not everyone is fully aware of it's limitations [12].
Eigenfactor	Gives weight to citations in more widely read journals but differences in the ranking between high and low profile specialty fields may challenge the robustness of any comparison of journals across a range of disciplines [13].
Article Influence Score	Provides a percentage of all scientific articles published by a specific journal, however, it has limitations that are similar to those of the Eigenfactor [14].
SCImago Journal Rank	Gives greater weight to journals such as those sharing a similar theme or specialty that frequently cite each other. Difficulty in comparing numbers across disciplines and values can be skewed in articles cite other articles published by less widely read journals [4, 5, 15].
Source normalised impact per paper	Greater weight is given to citations from the same field [16].

Impact Factor

Developed by Eugene Garfield and the Institute for Scientific Information (acquired by Thomson Scientific and Healthcare in 1992), JIF takes into account all indexed citations received by a given journal (target window) divided by the number of 'citable' articles published by a journal over a set time (census period). By convention, the Impact Factor usually refers to data from the previous 2 years of publication, although a 5-year impact factor is sometimes cited [12].

> *"Not everything that can be counted counts, and not everything that counts can be counted."*
>
> Albert Einstein

Because different sciences can have vastly different publication and citation rates, it is not appropriate to use JIF to make a quantitative comparison of journals from different fields.

Eigenfactor

Developed by researchers at the universities of Washington and California, the Eigenfactor differs from the impact factor in two important ways [13]. First, citations from more widely read journals, as determined by the citing journal's Eigenfactor score, are given greater weight. This limits the emphasis placed on articles in low-impact journals. Second, although there is a 1-year census period (as with the Impact Factor), the target window is 5 years and consequently the Eigenfactor appears less susceptible to rapid fluctuations or manipulation.

The weight it gives to citations from more widely read journals is determined by eigenvector centrality, which is a measure of a websites traffic and therefore its perceived importance. Similarly, citations from more active journals are given more weight. This should give the Eigenfactor a little more credibility than JIF. It is difficult for one journal to artificially increase the number of citations it receives from a more popular and widely read journal.

Biasing the Journal Impact Factor

Journal impact factors have been published yearly since 1972 for journals indexed in the Journal Citation Reports (an annual publication) and are widely promoted as describing the 'importance' of a journal. The distorting effects of this narrow and ill-informed view of research have been discussed many times [17. 18]. For example, the definition of what constitutes a citable article in a journal can be manipulated to decrease the denominator and increase the impact factor.

One widely cited article in a journal can artificially inflate a journal's impact factor, even if it is a controversial article that is subsequently cited in criticism. Case reports are often cited only infrequently, which has resulted in many journals opting to discontinue publishing this important type of article for concern that they may have a negative impact on their Impact Factor.

SCImago Journal Rank

Like the Eigenfactor, the SJR uses a page-rank algorithm to determine which citations are from more widely read journals, with these citations being given more weight. The main difference between the Eigenfactor and the SJR is that the former relies on the Institute for Scientific Information WOS database, whereas the SJR relies on the Scopus database [15].

An updated version of the SJR, known as the SJR2, was introduced in 2012, where SJR2 measures the cosine of the citing and cited journals to determine the thematic relationship of the journals. Journals that often cite each other are considered to be thematically close and are given greater weight. In addition, unlike any other bibliometric parameters, the SJR2 divides the prestige gained by a journal by the number of citable documents. The more often that related journals cite a specific journal, the more prominence that journal is given in its respective discipline. This approach was adopted to address a fundamental concern that affects many other bibliometric measures – as more journals and articles are added to research databases, bibliometric parameters are 'diluted', and comparison of numbers across time becomes limited.

CiteScore

Late in 2016, the publishers Elsevier introduced CiteScore, its own system of ranking journals. It is similar to the Impact Factor but covers twice as many journals and is based on the Scopus database. Unlike JIF, CiteScore incorporates editorials, letters and news items in its calculations - articles that are seldom cited. Thus, for example, The Lancet ranks fourth in the world under the JIF system but ranks below 200th in the CiteScore.

Elsevier claims that as new titles can receive CiteScore metrics the year after they are first indexed by Scopus the metric is more up to date. There is also no complex application process or unclear journal omissions. Should CiteScore ever reach the level of prominence that JIF currently holds, journals will face strong incentives to reduce or eliminate news articles that appeal to many readers.

Article Influence Score

The AIS is derived from the Eigenfactor [14]. The number of articles published by a journal over a 5-year period is divided by the total number of articles published by all journals during the same period. This gives an idea of what percentage of the total number of scientific articles were published in a given journal. The Eigenfactor score is then divided by this percentage, and the number is normalised to 1. An AIS greater than 1 means that each article in that journal has above-average influence, a score below 1 indicates below average influence.

Source-normalized Impact per Paper

Source-normalised Impact per Paper (SNIP) gives weight to citations that are made from the same scientific discipline [16]. Citations in fields that have fewer overall citations are given more weight. In essence, SNIP divides a journal's citation count per paper by the 'citation potential' in a given discipline. Thus, a citation from an article with 100 references will contribute less than a citation from an article with only 10 references. As SNIP takes this citation potential into account, it serves as a more robust comparator of journals from different disciplines. As with SJR, SNIP makes use of the Scopus database.

An interview with one of our Managing Director

Q Where do you stand on bibliometrics?

A I stand with Eugene Garfield, who started the 'metrics mania' with the proposal of the journal impact factor (JIF). He clearly stated that it was not suitable as a measure of the worth of individuals. He has largely been ignored and the JIF has come to dominate the lives of many researchers, despite decades of evidence of the harm it can do [17, 19]. Perhaps Aaron Levenstein put it most eruditely when he said "Statistics are like bikinis. What they reveal is suggestive, but what they conceal is vital."

Q Is there any practical value in the field of bibliometrics?

A You shouldn't mistake criticism for lack of appreciation. At Niche we are often asked to make qualitative comparisons and these quantitative parameters, that can be generated quickly and simply by computers, help inform our decisions (as long as we do this with the understanding of their limitations). I like Sir Arthur Conan Doyle's quote "I never guess. It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

Q Can you see things changing?

A There is a new breed of open-access journals that are free to access and have low running costs and overheads. Born on the web, they can accept all papers that meet appropriate quality standards, with no artificial caps. Many are edited by working scientists, who assess the worth of papers without regard for citations. There are new ways of continually evaluating the contribution of individual pieces of research. These represent an opportunity to re-democratise the scientific literature

Q Do you have a pet hate?

A The influence of the 'big journals' and the perverse distortion and bias they bring to the scientific literature. These journals aggressively curate their brands, in ways more conducive to selling subscriptions than to stimulating the most important research. It should be relatively easy to reduce the influence that these purely commercial players have on the publishing environment, except for the fact that many institutions emphasise the importance publication in these journals have on career development.

Altmetrics

Altmetrics is the latest buzzword in the vocabulary of bibliometricians. These new techniques attempt to describe the 'impact' of a piece of research by counting the number of times that it is mentioned in tweets, Facebook pages, blogs, YouTube and news media. To a scientist it is likely that this sounds like a rather unnatural source of data for scientific evaluation. Some studies have shown correlations between altmetric scores and journal impact factors [2, 20]. However, it may be easy to manipulate scores simply through the application of resource, suggesting these values are not unbiased.

Twitter may be an excellent tool for journalism, it's good for debunking bad science and for spreading links. But, as pointed out on scientific discussion boards, these communication channels may be too brief for serious discussions and rarely useful for appropriate scientific discussions [21].

"Altmetrics are numbers generated by people who don't understand research, for people who don't understand research. People who read papers and understand research just don't need them and should shun them [21]."

And finally...

Although there are limitations with bibliometric analyses we shouldn't throw the baby out with the bath water. Bibliometric analyses adds a quantitative aspect to an otherwise somewhat qualitative process. Moving beyond simple tallies of publication totals and impact factors, modern analytic tools have emerged to improve on prior, manual methods of assessment. There is no single ideal tool, however an accurate understanding of bibliometric parameters can aid in effectively evaluating individual authors, departments and institutions, as well as individual articles and journals. Under the right circumstances the ability to utilise this information can provide you with a powerful insight into specific aspects of the publishing landscape.

Nevertheless, it is critical to remember that academic productivity, whether of an individual, department, or journal, cannot accurately be reduced to a single number on a linear scale. Any imperfect metric is subject to manipulation and an understanding of the strengths and weaknesses of different bibliometric parameters is required to detect attempts at manipulation. It should be remembered that all these parameters are simple surrogate markers and you should never let one single metric distract from developing your own informed opinion on the value of a piece of research.

Post publication peer review is now happening, in comments on published papers and through sites like PubPeer, where it is already clear that anonymous peer review can work really well. New journals like eLife have open comments after each paper, though authors do not seem to have yet adopted the habit of using them constructively. They will!

Although the h-index seems to be a rather crude tool for measuring academic performance, it has stood the test of time and correlates well with the more elaborately mathematical parameters. The same cannot be said of the various bibliometric parameters used for assessing journals. The Journal Impact Factor has been widely criticised as being imprecise and subject to manipulation, yet it continues to be the most widely used metric. Other metrics, such as the Eigenfactor, AIS and SJR, offer more robust analysis and are starting to gain traction within the scientific community.

References

1. Insider's Insight into the identification and profiling of Key Opinion Leaders, http://www.niche.org.uk/asset/insider-insight/Insider-KOLIP-P.pdf [accessed 10 June 2016].
2. Insider's Insight into Targeted Journal Selection, http://www.niche.org.uk/asset/insider-insight/Insider-Targeted-Journal-Selection.pdf [accessed 10 June 2016].
3. Pritchard, Alan, 'Statistical Bibliography or Bibliometrics?', Journal of Documentation, 1969; 25:348-349.
4. Scopus. Elsevier Web site. http://www.scopus.com. [accessed 20 November 2013].
5. ISI Web of Knowledge. Thomson Reuters Web site. http: //www.isiwebofknowledge.com. Published 2013 [accessed 1 March 2013].
6. Sidiropoulos A, Katsaros D, Manolopoulos Y. Generalized Hirsch h-index for disclosing latent facts in citation networks. Scientometrics 2007;72:253–280.
7. Hirsch JE. Does the h index have predictive power? Proc Natl Acad Sci USA 2007;104:19193–19198.
8. Kulkarni AV et al. Author self-citation in the general medicine literature. PLoS ONE 2011; 6: e20885. doi:10.1371/journal.pone.0020885
9. Egghe L. Theory and practice of the g-index. Scientometrics 2006;69:131–152.
10. Cantor M, Gero S. The missing metric: quantifying contributions of reviewers. R Soc open sci 2015; 2:140540.
11. Harzing AW. Publish or Perish. http://www.harzing.com/pop.htm. Published 2007 [accessed June 10, 2014].
12. Garfield E. Citation analysis as a tool in journal evaluation. Science 1972;178(4060):471–479.
13. Bergstrom CT, West JD, Wiseman MA. The Eigenfactor metrics. J Neurosci 2008;28:11433–11434.
14. West JD, Bergstrom TC, Bergstrom CT. The Eigenfactor metrics: a network approach to assessing scholarly journals. Coll Res Libr 2010;71:236–244.
15. González-Pereira B, Guerrero-Bote VP, Moya-Anegón F. A new approach to the metric of journals' scientific prestige: the SJR indicator. J Informetrics 2010;4:379–391.
16. Moed HF. The source normalized impact per paper is a valid and sophisticated indicator of journal citation impact. J Am Soc Inf Sci Technol 2011;62:211–213.
17. Vale RD. Evaluating how we evaluate. Mol Biol Cell 2012; 23:3285–3289.
18. Vosshall LB. The glacial pace of scientific publishing: why it hurts everyone and what we can do to fix it. FASEB J 2012; 26:3589–3593.
19. Colquhoun, D. Challenging the tyranny of impact factors Nature 2003 423 (6939), 479.
20. Scotti V, De Silvestri A, Scudeller L, et al. Novel bibliometric scores for evaluating research quality and output: a correlation study with established indexes. Int J Biol Markers 2016 Jun 8:0. doi: 10.5301/jbm.5000217.
21. http://www.dcscience.net/2014/01/16/why-you-should-ignore-altmetrics-and-other-bibliometric-nightmares/ [accessed 10 June 2016].

Next Steps

We created this Insider's Insight into bibliometrics to share some helpful pointers. We hope you found it useful. If you would like advice on practical ways you can use this data please contact me at the email address below.

Dr Justin Cook
Head of Medical Writing
justin.cook@niche.org.uk

Get in touch

niche science&technology

+44 (0)20 8332 2588
www.niche.org.uk