Online Information Review

An indicator of the impact of journals based on the percentage of their highly cited publications

Sara M. González-Betancor, Pablo Dorta-González,

Article information:

To cite this document:
Sara M. González-Betancor, Pablo Dorta-González, (2017) "An indicator of the impact of journals based on the percentage of their highly cited publications", Online Information Review, Vol. 41 Issue: 3, pp.398-411, https://doi.org/10.1108/OIR-01-2016-0008
Permanent link to this document:
https://doi.org/10.1108/OIR-01-2016-0008

Received 11 January 2016
Revised 19 December 2016
Accepted 31 March 2017

An indicator of the impact of journals based on the percentage of their highly cited publications

Sara M. González-Betancor and Pablo Dorta-González
Metodos Cuantitativos en Economia y Gestion,
University of Las Palmas de Gran Canaria, Las Palmas, Spain

Abstract

Purpose – The two most used citation impact indicators in the assessment of scientific journals are, nowadays, the impact factor and the *h*-index. However, both indicators are not field normalized (vary heavily depending on the scientific category). Furthermore, the impact factor is not robust to the presence of articles with a large number of citations, while the *h*-index depends on the journal size. These limitations are very important when comparing journals of different sizes and categories. The purpose of this paper is to propose an alternative citation impact indicator, based on the percentage of highly cited articles in the journal.

Design/methodology/approach – This alternative indicator is empirically compared with the impact factor and the *h*-index, considering different time windows and citation percentiles (levels of citation for considering an article as highly cited compared to others in the same year and category). The authors use four journal categories (Clarivate Analytics Web of Science) which are quite different according to the publication profiles and citation levels (*Information Science & Library Science*, *Operations Research & Management Science*, *Ophthalmology*, and *Physics Condensed Matter*).

Findings – After analyzing 20 different indicators, depending on the citation percentile and the time window in which citations are counted, the indicator that seems to best homogenize the categories is the one that considers a time window of two years and a citation level of 10 percent.

Originality/value – The percentage of highly cited articles in a journal is field normalized (comparable between scientific categories), independent of the journal size and also robust to the presence of articles with a high number of citations.

Keywords Citation analysis, *H*-index, Impact factor, Citation impact indicator, Journal assessment, Percentage of highly cited articles

Paper type Research paper

Introduction

Scientific communication is currently developed by means of the internet. There are a growing number of scientific journals that either publish scientific articles in open access or at least allow access to the article through the subscription mode. Thus, researchers and people, in general, are in need of tools that help them discriminate the relevance and quality of content in scientific journals. In this regard, the present study proposes a suitable tool to achieve this goal, which can be added to other well-known indexes.

To evaluate the quality of a research – in comparison to other research – is one of the purposes of research assessment. As quality is subjective and difficult to measure, citations are used as a proxy. Citations play an important role in scholarly communication and are a significant component in research evaluation. The assumption is that highly cited work has influenced the work of many other researchers and hence it is more valuable.

At present two families of citation impact indicators for scientific journals are often used. The first being the journal impact indices, which consider an average of citations per publication for a given census and citation time window (Garfield, 1972). This family includes, among others, the journal impact factors (for two and five years) of database Journal Citation Reports (JCR) owned by Clarivate Analytics (Bensman, 2007; Bergstrom, 2007; Moed *et al.*, 2012), the maximum impact factor (Dorta-González and Dorta-González, 2013c), and the SJR impact indexes (González-Pereira *et al.*, 2009) of the Scopus database owned by Elsevier.

Online Information Review
Vol. 41 No. 3, 2017
pp. 398-411
© Emerald Publishing Limited
1468-4527
DOI 10.1108/OIR-01-2016-0008

For decades, the Journal Impact Factor (JIF) has been an accepted indicator in ranking journals. However, there are increasing arguments against the fairness of using the JIF as the sole ranking criteria (Waltman and Van Eck, 2013). The JIF is defined as the average number of citations to each journal in a current year with respect to "citable ítems" published in that journal during the two preceding years (Garfield, 1972). Nevertheless, it has been criticized due to arbitrary decisions in its construction. The definition of "citable items" including letters together with the peer reviewed papers (research articles, proceedings papers, and reviews), the focus on the two preceding years, the incomparability between fields, etc., have been discussed in the literature (Bensman, 2007; Moed *et al.*, 2012) and have given many possible modifications and improvements (Althouse *et al.*, 2009; Bornmann and Daniel, 2008). In response, Clarivate Analytics has incorporated the five-year impact factor, the eigenfactor score, and the article influence score (Bergstrom, 2007) to the JCR journals. All these indicators consider a five-year citation window and are useful for comparing journals in the same subject category. However, although in many cases, the five-year impact factor is greater than the two-year impact factor, both indicators lead statistically to the same ranking (Leydesdorff, 2009; Rousseau, 2009).

The second of these families of citation impact indicators for scientific journals are the *h* indices (Hirsch, 2005), which consider that *h* maximum integer value for which we can say that there are *h* publications with *h* or more citations, all within a given time window. This indicator estimates the number of important works published by a journal, increasing the requirement while increasing its value. This is a robust indicator that considers both quantitative and qualitative aspects. However, although this indicator has proven useful for detecting the most prestigious journals in an area, there is empirical evidence which does not discriminate between those at intermediate levels, and penalize selective journals in relation to the major producers (Costas and Bordons, 2007a, b; Dorta-González and Dorta-González, 2011; Egghe, 2013).

Both families of citation impact indicators are useful for comparing journals within the same field. However, they are not appropriate when comparing different scientific fields. Different scientific fields have different citation practices and citation-based bibliometric indicators need to be normalized for such differences in order to allow for journal comparisons. This problem of field-specific differences in citation impact indicators comes from institutional research evaluation (Leydesdorff and Bornmann, 2011; Van Raan *et al.*, 2010). For example, research institutes often have among their missions the objective of integrating interdisciplinary bodies of knowledge which are generally populated by scholars with different disciplinary backgrounds (Leydesdorff and Rafols, 2011; Wagner *et al.*, 2011).

There are statistical patterns which are field specific and allow for the normalization of the JIF. Garfield (1979) proposes the term "citation potential" for systematic differences among fields of science, based on the average number of references. For example, in the biomedical fields long reference lists with more than 50 items are common, but in mathematics short lists with less than 20 references are the standard (Dorta-González and Dorta-González, 2013a). These differences are a consequence of the citation cultures and can produce significant differences in the JIF, since the probability of being cited is affected. In this sense, the average number of references is the variable that has most frequently been used in the literature to justify the differences between fields of science, as well as the most employed in source normalization (Leydesdorff and Bornmann, 2011; Moed, 2010; Zitt and Small, 2008). However, the variables that to a greater degree explain the variance in the impact factor do not include the average number of references (Dorta-González and Dorta-González, 2013a) and, therefore, it is necessary to consider other sources of variance in the normalization process, such as the ratio of references to journals included in the JCR, the field growth, the ratio of JCR references to the target window, and the proportion of cited to citing items. Given these large differences in citation practices, the development of bibliometric indicators that allow for between-field comparisons is clearly a critical issue (Waltman and Van Eck, 2013).

Traditionally, normalization has also been based on a classification system of journals. This is the case, for example, with the categories in the Web of Science database (Egghe and Rousseau, 2002), the relative position with respect to these categories (Bordons and Barrigón, 1992) and the quartile where each journal belongs when ranked in decreasing order according to their impact factor. However, the delimitation of the scientific fields and disciplines is a problem not adequately solved in bibliometrics, as these boundaries are diffuse and dynamically developed over time (Leydesdorff, 2012, p. 359). Alternatively, the idea of source normalization has been proposed. In this approach, the normalization is performed according to the citing journals (Dorta-González *et al.*, 2014; Leydesdorff and Bornmann, 2011).

The two families of citation impact indicators most commonly used in the evaluation of journals (impact factors and *h* indices) depend strongly on the scientific field they belong to, which makes them non-comparable across disciplines. In addition, the *h*-index also depends on the size of the journal, while the impact factor is not robust with the presence of a small number of highly cited articles. Due to these limitations when comparing journals of different sizes and fields, it is necessary to consider other indicators of impact for journals, to enable comparisons between fields, which do not depend on the size of the journal and be at the same time robust as previously mentioned (Waltman and Van Eck, 2013).

An alternative to this issue is the percentage of highly cited articles in the journal, considering the term article in its strictest sense. Being a percentage, it is a relative value, so this indicator does not depend on the size of the journal. High citation is determined by comparing with the other items in the same field and year at international level, so this indicator does not depend on the field. In addition, it is robust because the inclusion of a new widely cited article does not significantly affect the value of the indicator.

This paper empirically compares this index with the impact factor and the *h*-index, considering different time windows and citation percentiles, i.e. levels of citation for considering an article as highly cited compared to others in the same year and field.

Percentage of highly cited articles in a journal

Highly cited articles are those which have received a number of citations that equals or exceeds the citations of the paper that occupies the q percentile position for their category and year of publication. This q-value may vary depending on the purpose intended. In this work, we have set the following citation percentiles 10, 20, 25, 30, and 40 percent as benchmarks. Two of them – 10 and 20 percent – coincide with Clarivate Analytics Essential Science Indicators baselines for percentiles. The other three could be reasonable threshold for considering an article as highly cited, instead of jumping to the median, which would be the next percentile baseline considered by Clarivate Analytics.

Having the minimum number of citations necessary to belong to the group of highly cited articles in a category and year of publication, then how many articles meet this requirement in each journal of the category can be determined. Putting this information in relation to the total number of articles published that year by the journal (supplementary material 1), an impact indicator of the journal's scientific production is obtained.

Since the total number of citations of an article is a value that grows over time, it is necessary to set an observation time window. This paper looks at four possible time windows, covering two to five years.

Let $(pArt\_q\_t)_t^j$ be the impact indicator for the j-journal and the y-year, which measures the percentage of articles in the q percent of the most cited, considering a time window of t years. That is, for a given year and journal, $pArt\_q\_t$ compares the citations of the articles published in the journal in the period [$y-t+1$, y] with those of other articles in the same category and year of the same database, determining what percentage of them go within q percent of the most cited.

For example, $(pArt\_q\_t)^{OIR}_{2013}$ identifies the percentage of articles within the first quartile of the most cited, considering a time window of four years, for the *Online Information Review* (*OIR*) in 2013. This indicator compares the citations of the articles published by *OIR* in the period 2010-2013 with the citations of the rest of articles of the category *Information Science & Library Science (IS & LS)* within the Web of Science database during the same period, determining what percentage of them are within 25 percent of the most cited.

As it is a relative value, the indicator does not depend on the size of the journal. Nor does it depend on the category, as high citation is determined by comparison with other articles within the same category and year. In addition, it is robust because the inclusion of a new widely cited article does not significantly affect the indicator value. This is because all those citations above the ones needed for being highly cited are not considered for the calculation.

This paper empirically examines this indicator and compares it with the two-year and five-year impact factor, and the three-year and five-year *h*-index. Different time windows and benchmarks for a document to be considered highly cited are compared using a total amount of 278 journals and four different categories. The percentage of highly cited articles is determined for each journal in terms of five percentage benchmarks (10, 20, 25, 30, and 40 percent) and four time windows (2012-2013, 2011-2013, 2010-2013, and 2009-2013). Thus, 20 indicators (5 benchmarks × 4 windows) are compared with the two-year and five-year impact factor of 2013, and the three-year and five-year *h*-index of 2013. The aim of the study is to identify which of these indicators are field normalized and valid for comparing journals of different categories.

Materials and methods

The Clarivate Analytics Web of Science classifies journals (approximately 12,000) into 251 scientific categories, which are grouped into 151 research areas and 22 scientific fields. For this research it was decided to work at the level of scientific category and select four categories which are quite different according to the publication profiles and citation levels. Specifically, the categories selected were *IS & LS*, *Operations Research & Management Science (OR & MS)*, *Ophthalmology*, and *Physics Condensed Matter (Physics CM)*. Furthermore, it was decided to work with the total population of journals of these categories, instead of a sample.

The information collected from the Clarivate Analytics databases for each journal was the following:

- total number of research articles published between 2009 and 2013;
- five-year *h*-index (considering the data from 2009 to 2013) and three-year *h*-index (considering the data from 2011 to 2013);
- two-year and five-year impact factor for 2013; and
- number of articles in the journal with enough citations to belong to the group of the 10, 20, 25, 30, or 40 percent most cited in its category and year of publication, for each of the five years (2009-2013).

With this information we created the database shown in the supplementary material, which is the basis for our research.

It was decided not to include the year 2014, because at the time in which the database searches were made, the impact factor of 2014 had not yet been published. All searches were made between February and March 2015.

First, we evidence graphically the differences in the average number of citations between scientific categories that hinder direct comparisons among them. Then, we test if the proposed index – percentage of highly cited articles – is field normalized, allowing to compare between journals of different categories, and if it is also valid to measure the impact of the journal. To check these two characteristics, we perform every analysis comparing the results with those obtained using the most common indicators (JIF, *h*-index).

To check field normalization, we start analyzing the descriptive statistics of every index in each of the four considered categories, and continue performing an analysis of variance (ANOVA) for every index to test if the means are equal among categories – enabling comparisons among them – or not.

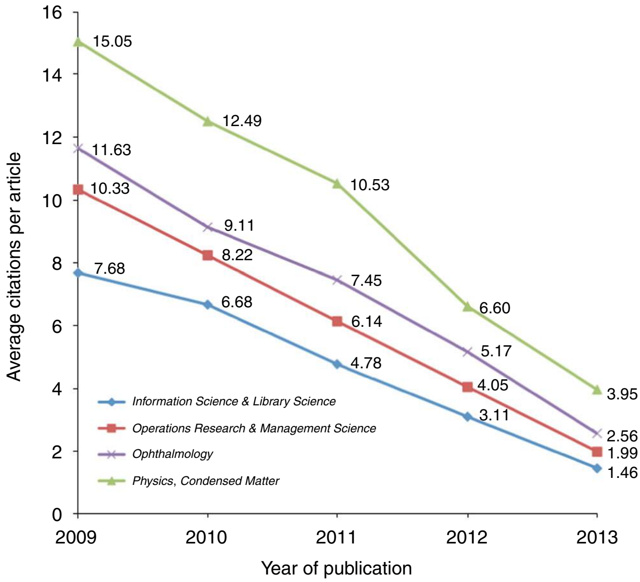
Finally, to check if the proposed indexes actually measure the impact of the journals, we compare the ranking of journals generated within each category by these indexes with those obtained by the JIF or the *h*-index, calculating Spearman's rank correlation among them.

Results

Differences between scientific categories

Figure 1 shows, based on the information of the Web of Science database, the average number of citations per research article, in the period 2009-2013 for each of the four scientific categories analyzed. This average citation rate is the ratio between the number of citations since publication until 2013 and the number of articles published in each particular year. For example, the value 10.53 in *Physics CM* in 2011 indicates that, on average, the research articles of 2011 in this category have received a total of 10.53 citations between 2011 and 2013.

As shown in Figure 1, the differences in the average number of citations between categories are very important. Thus, the average citations received in *Physics CM* are about twice those of *IS & LS*. Ordering the journal categories from the highest to the lowest average citations, the ranking obtained is the following: *Physics CM*, *Ophthalmology*, *OR & MS*, and *IS & LS*.



Note: Observe the important differences among journal categories
Source: Clarivate analytics web of science

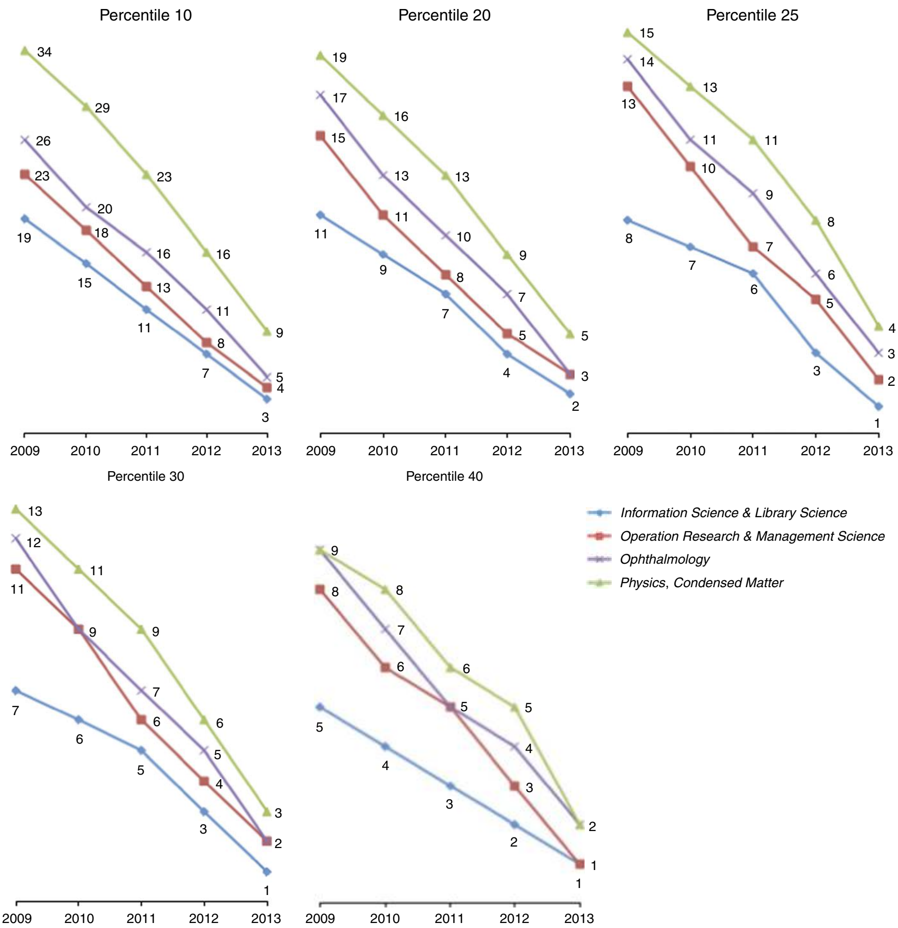
Figure 1.
Average citations per article by year of publication

The citation percentiles of Figure 2 show the minimum number of citations required to reach a certain percentage benchmark in each journal category and year. For example, in the case of *Physics CM*, a value of 23 in the tenth percentile for 2011 indicates that 10 percent of the most cited articles in this category have received at least 23 citations during the period 2011-2013.

As can be seen, there are again major differences between categories. Ordering the scientific categories from the highest to the lowest, according to the required citations needed to reach a specific citation percentile, the same sequence as in the above figure is again obtained: *Physics CM*, *Ophthalmology*, *OR & MS*, and *IS & LS*.

Field normalization (homogeneity)

A valid indicator should have the following properties related to homogeneity: it should be homogeneous between scientific categories (between-groups) but heterogeneous within the same scientific category (within-groups). The between-group homogeneity ensures



Note: Observe again the important differences between journal categories
Source: Clarivate analytics web of science

Figure 2. Minimum number of citations necessary to ensure that an article is within 10, 20, 25, 30, and 40 percent of the most cited in its journal category and year of publication

comparability of the indicator between journals of different categories, while the intra-group heterogeneity guarantees the discrimination capacity of the indicator.

Table I shows the descriptive statistics of the analyzed variables (JIF, *h,* and *pArt_q_t*) by scientific category. Regarding *pArt*, and within each of the categories, the dispersion of all the proposed indicators is quite large compared to its average (Pearsons' coefficient of variation varies between 0.538 – for *pArt_40_2* in *OR & MS* – and 1.978 – to *pArt_10_2* in *Physics CM*). This is motivated by the important differences between scientific journals in relation to the number of citations received by each of them, and highlights the discrimination capacity of the indicator (intra-group heterogeneity). This relative variability decreases in all scientific categories, as the citation percentile considered grows. Naturally, the same applies when considering all the journals together without differentiating between scientific categories (see Table II). That is, as the citation percentile decreases, the indicator variability increases, showing that the index is more discriminative.

The average, minimum, and maximum values increase as the citation percentile increases. But just the opposite occurs with the skewness and the kurtosis, that is, these descriptive statistics decrease as the citation percentile increases. This phenomenon is directly related to the definition of the indicator. Note that increasing the citation percentile, the number of articles that meet this condition in every journal increases progressively, reaching the total amount of published articles in the 100-citation-percentile. When this occurs (100 percentile), all indicator values match 1, so that their mean is 1 and their standard deviation is 0. Therefore, as the citation percentile increases, the indicators' distribution tends to become more similar to a standard normal distribution, and hence what we observe about the skewness and kurtosis values.

The minimum values evidence in all categories that, when the citation percentile considered is up to 25, there are journals with no articles among the most cited in its scientific category. In the case of *IS & LS* that is extensible to at least the 40th percentile.

Moreover, the maximum values evidence the concentration of highly cited articles in some journals. This phenomenon is more common in the *Ophthalmology* category, where some journals have 70 percent of their articles among the 10 percent of the most cited in its category, compared to the 49 percent maximum in the *OR & MS* category.

Table II shows the descriptive statistics of the analyzed variables for the aggregated data. The dispersion of each of the indicators analyzed, considering all the journals together, is made up of the differences between categories (between-groups) as well as the differences between journals within a category (within-groups). Considering both parts independently, it is observed that the percentage of the variance explained by the differences between groups (categories) is practically zero in all indicators, so that almost 100 percent is explained by the differences among journals in the same category. This shows that all *pArt* indicators analyzed are homogeneous between groups and heterogeneous within them, so that direct comparison of the indicator value itself between categories is ensured.

Table II also shows an ANOVA for each of the 20 indicators of the *pArt* family, in order to test the null hypothesis that the mean of all categories are equal, vs the alternative that some of them are different (*t*-test for independent samples). In all cases, except of *pArt_25_2*, there is no statistical evidence to reject the null hypothesis at any level of significance, so that we can speak of homogeneity among groups of indicators. In the case of indicator *pArt_25_2* the category with statistically significant mean difference, compared to other categories, is *IS & LS*. Since the tests are carried out for the 278 journals together with no missing values in any group, the degrees of freedom are kept constant, so – on the basis that all indicators (except *pArt_25_2*) are equally homogeneous – the best indicator is the one with the highest probability associated with the *F* statistic, or in other words, the one with the smallest value of *F*. In this sense, the best indicator is the *pArt_10_2*. It really is the most

Variable	Obs	Mean	SD	CV	Min.	Max.	Asymmetry	Kurtosis
Information Science & Library Science								
JIF5	73	1.397	1.442	1.032	0.019	8.157	1.988	8.368
JIF2	82	1.030	0.988	0.959	0.034	5.405	1.913	7.518
h5	82	10.634	8.097	0.761	1	39	1.486	5.256
h3	82	6.695	4.941	0.738	1	27	1.596	6.135
pArt_10_5	82	0.072	0.104	1.460	0	0.556	2.201	8.700
pArt_10_4	82	0.071	0.103	1.452	0	0.539	2.175	8.471
pArt_10_3	82	0.071	0.104	1.460	0	0.497	2.045	7.257
pArt_10_2	82	0.072	0.101	1.413	0	0.470	2.008	7.202
pArt_20_5	82	0.147	0.169	1.144	0	0.710	1.222	3.803
pArt_20_4	82	0.147	0.168	1.144	0	0.703	1.203	3.749
pArt_20_3	82	0.150	0.169	1.124	0	0.675	1.146	3.476
pArt_20_2	82	0.156	0.170	1.088	0	0.679	1.109	3.472
pArt_25_5	82	0.221	0.206	0.933	0	0.814	0.885	2.762
pArt_25_4	82	0.225	0.207	0.920	0	0.819	0.855	2.716
pArt_25_3	82	0.237	0.210	0.884	0	0.801	0.769	2.517
pArt_25_2	82	0.275	0.221	0.806	0	0.813	0.559	2.147
pArt_30_5	82	0.241	0.217	0.898	0	0.824	0.768	2.476
pArt_30_4	82	0.244	0.217	0.891	0	0.832	0.759	2.463
pArt_30_3	82	0.249	0.215	0.864	0	0.812	0.712	2.400
pArt_30_2	82	0.275	0.221	0.804	0	0.813	0.560	2.150
pArt_40_5	82	0.317	0.250	0.789	0	0.882	0.487	1.971
pArt_40_4	82	0.320	0.250	0.781	0	0.888	0.454	1.919
pArt_40_3	82	0.321	0.248	0.772	0	0.874	0.443	1.931
pArt_40_2	82	0.324	0.246	0.759	0	0.873	0.408	1.887
Operations Research & Management Science								
JIF5	73	1.590	1.119	0.704	0.337	7.718	2.645	13.879
JIF2	78	1.215	0.796	0.655	0.220	4.478	1.630	6.584
h5	78	15.500	9.511	0.614	4	52	1.578	5.793
h3	78	9.808	5.961	0.608	2	35	1.513	5.819
pArt_10_5	78	0.060	0.080	1.324	0	0.495	3.031	15.041
pArt_10_4	78	0.060	0.079	1.307	0	0.469	2.957	14.133
pArt_10_3	78	0.061	0.079	1.286	0	0.420	2.647	11.741
pArt_10_2	78	0.066	0.085	1.285	0	0.457	2.417	9.965
pArt_20_5	78	0.135	0.126	0.936	0	0.702	1.815	7.693
pArt_20_4	78	0.136	0.125	0.917	0	0.663	1.736	7.125
pArt_20_3	78	0.135	0.121	0.894	0	0.596	1.585	6.230
pArt_20_2	78	0.134	0.118	0.878	0	0.585	1.451	5.532
pArt_25_5	78	0.174	0.145	0.831	0	0.726	1.323	5.091
pArt_25_4	78	0.177	0.144	0.817	0	0.691	1.265	4.784
pArt_25_3	78	0.179	0.143	0.798	0	0.670	1.167	4.367
pArt_25_2	78	0.183	0.141	0.770	0	0.676	1.027	3.974
pArt_30_5	78	0.210	0.160	0.763	0.012	0.764	1.056	3.963
pArt_30_4	78	0.211	0.159	0.750	0	0.737	1.032	3.858
pArt_30_3	78	0.216	0.156	0.725	0	0.701	0.895	3.401
pArt_30_2	78	0.217	0.153	0.703	0	0.702	0.814	3.252
pArt_40_5	78	0.318	0.190	0.597	0.024	0.851	0.515	2.569
pArt_40_4	78	0.323	0.189	0.587	0.013	0.834	0.472	2.538
pArt_40_3	78	0.328	0.186	0.568	0.019	0.787	0.375	2.421
pArt_40_2	78	0.350	0.189	0.538	0	0.793	0.188	2.363
Ophthalmology								
JIF5	55	2.154	1.705	0.792	0.192	11.207	3.038	15.957
JIF2	57	2.067	1.550	0.750	0.163	9.897	2.661	13.171

(*continued*)

Table I. Descriptive statistics of variables (*JIF*5, *JIF*2, *h5*, *h3*, *pArt_q_t*, $q = 10, 20, 25, 30, 40$; $t = 5, 4, 3, 2$) in journal categories

Variable	Obs	Mean	SD	CV	Min.	Max.	Asymmetry	Kurtosis
h5	58	21.310	12.895	0.605	4	61	1.062	3.606
h3	58	14.121	8.323	0.589	3	45	1.555	6.031
pArt_10_5	58	0.075	0.102	1.358	0	0.701	4.170	25.564
pArt_10_4	58	0.075	0.102	1.360	0	0.701	4.211	25.941
pArt_10_3	58	0.076	0.104	1.366	0	0.713	4.197	25.875
pArt_10_2	58	0.081	0.103	1.279	0	0.707	3.979	24.185
pArt_20_5	58	0.154	0.138	0.894	0	0.841	2.316	11.789
pArt_20_4	58	0.156	0.138	0.886	0	0.843	2.295	11.707
pArt_20_3	58	0.160	0.139	0.871	0	0.851	2.259	11.585
pArt_20_2	58	0.165	0.139	0.846	0	0.853	2.238	11.446
pArt_25_5	58	0.180	0.149	0.826	0	0.879	1.943	9.613
pArt_25_4	58	0.179	0.148	0.825	0	0.881	2.002	9.979
pArt_25_3	58	0.179	0.147	0.818	0	0.881	2.021	10.171
pArt_25_2	58	0.181	0.144	0.796	0	0.867	1.995	9.985
pArt_30_5	58	0.242	0.162	0.668	0.008	0.904	1.265	6.281
pArt_30_4	58	0.246	0.163	0.662	0.010	0.910	1.237	6.186
pArt_30_3	58	0.250	0.162	0.647	0.012	0.911	1.215	6.126
pArt_30_2	58	0.257	0.160	0.623	0.012	0.907	1.188	5.989
pArt_40_5	58	0.299	0.176	0.587	0.008	0.924	0.767	4.381
pArt_40_4	58	0.300	0.174	0.582	0.010	0.925	0.783	4.467
pArt_40_3	58	0.300	0.172	0.573	0.012	0.921	0.800	4.487
pArt_40_2	58	0.288	0.165	0.573	0.012	0.907	0.896	4.828
Physics, Condensed Matter								
JIF5	59	3.296	6.216	1.886	0.098	41.775	4.492	26.468
JIF2	60	3.115	5.485	1.761	0.109	36.425	4.302	24.478
h5	60	34.667	37.272	1.075	4	175	2.425	8.411
h3	60	23.117	24.587	1.064	2	112	2.313	7.696
pArt_10_5	60	0.084	0.162	1.943	0	0.635	2.359	7.162
pArt_10_4	60	0.082	0.160	1.957	0	0.635	2.427	7.551
pArt_10_3	60	0.080	0.158	1.965	0	0.635	2.461	7.711
pArt_10_2	60	0.078	0.155	1.978	0	0.624	2.513	7.981
pArt_20_5	60	0.160	0.217	1.355	0	0.821	1.997	5.847
pArt_20_4	60	0.159	0.214	1.345	0	0.821	2.004	5.943
pArt_20_3	60	0.159	0.212	1.334	0	0.821	2.042	6.132
pArt_20_2	60	0.160	0.209	1.309	0	0.825	2.047	6.199
pArt_25_5	60	0.195	0.229	1.174	0	0.866	1.799	5.270
pArt_25_4	60	0.193	0.227	1.176	0	0.866	1.798	5.315
pArt_25_3	60	0.192	0.224	1.169	0	0.866	1.830	5.481
pArt_25_2	60	0.195	0.222	1.140	0	0.863	1.816	5.467
pArt_30_5	60	0.247	0.241	0.976	0.001	0.913	1.508	4.486
pArt_30_4	60	0.251	0.242	0.964	0.002	0.913	1.454	4.353
pArt_30_3	60	0.255	0.240	0.940	0.002	0.913	1.432	4.345
pArt_30_2	60	0.268	0.239	0.891	0	0.912	1.353	4.168
pArt_40_5	60	0.337	0.251	0.743	0.007	0.948	0.939	3.327
pArt_40_4	60	0.341	0.251	0.736	0.007	0.948	0.887	3.219
pArt_40_3	60	0.349	0.248	0.711	0.010	0.948	0.834	3.188
pArt_40_2	60	0.355	0.245	0.690	0.014	0.943	0.840	3.219

Table I.

Source: Supplementary material

restrictive of all proposed indicators (percentage of articles within the tenth percentile of citation in a citation window of two years).

Furthermore, performing the same hypothesis testing for JIF and *h*, it follows that they are not field normalized (Table II). The tests for *JIF5* and *JIF2* indicate that there are

Table II. Descriptive statistics of variables for aggregated data, and analysis of variance (ANOVA)

Variable	Obs	Mean	SD	CV	% of variance explained by groups	*F*	Prob.	Min.	Max.	Asymmetry	Kurtosis
JIF5	260	2.042	3.274	1.603	0.051	4.49	0.004	0.019	41.775	7.902	87.549
JIF2	277	1.747	2.836	1.623	0.094	8.14	0.000	0.034	36.425	7.889	86.477
h5	278	19.414	21.277	1.096	0.208	19.09	0.000	1	175	4.272	26.861
h3	278	12.662	14.035	1.108	0.226	21.09	0.000	1	112	4.224	25.606
pArt_10_5	278	0.072	0.113	1.577	0.000	0.51	0.678	0	0.701	3.014	13.319
pArt_10_4	278	0.071	0.112	1.570	0.000	0.45	0.719	0	0.701	3.047	13.646
pArt_10_3	278	0.071	0.112	1.564	0.000	0.36	0.779	0	0.713	2.987	13.308
pArt_10_2	278	0.073	0.111	1.512	0.000	0.24	0.869	0	0.707	2.881	12.717
pArt_20_5	278	0.148	0.163	1.105	0.000	0.30	0.827	0	0.841	1.943	7.265
pArt_20_4	278	0.148	0.162	1.094	0.000	0.29	0.834	0	0.843	1.920	7.228
pArt_20_3	278	0.150	0.161	1.077	0.000	0.35	0.788	0	0.851	1.890	7.130
pArt_20_2	278	0.153	0.161	1.052	0.000	0.51	0.678	0	0.853	1.856	7.036
pArt_25_5	278	0.194	0.185	0.957	0.000	0.99	0.399	0	0.879	1.536	5.293
pArt_25_4	278	0.195	0.185	0.949	0.002	1.13	0.337	0	0.881	1.514	5.222
pArt_25_3	278	0.199	0.185	0.930	0.010	1.71	0.165	0	0.881	1.468	5.040
pArt_25_2	278	0.212	0.190	0.898	0.047	4.37	0.005	0	0.867	1.345	4.485
pArt_30_5	278	0.234	0.197	0.843	0.000	0.54	0.654	0	0.913	1.235	4.396
pArt_30_4	278	0.237	0.197	0.834	0.000	0.61	0.611	0	0.913	1.209	4.328
pArt_30_3	278	0.241	0.196	0.811	0.000	0.63	0.595	0	0.913	1.162	4.247
pArt_30_2	278	0.253	0.197	0.777	0.005	1.33	0.265	0	0.912	1.083	4.007
pArt_40_5	278	0.318	0.220	0.691	0.000	0.30	0.829	0	0.948	0.712	3.002
pArt_40_4	278	0.321	0.219	0.683	0.000	0.35	0.790	0	0.948	0.679	2.942
pArt_40_3	278	0.325	0.217	0.668	0.000	0.49	0.687	0	0.948	0.643	2.923
pArt_40_2	278	0.330	0.216	0.653	0.003	1.24	0.295	0	0.943	0.595	2.864

Source: Supplementary material

differences between *Physics CM* and *IS & LS*, as well as between *Physics CM* and *OR & MS*. The tests for *h5* and *h3* indicate that there are differences between *Physics CM* and the other three categories, and between *Ophthalmology* and *IS & LS*.

Indicator validity

In the previous section, we have deduced that the *pArt* family of indicators is homogeneous among scientific categories, i.e. it has no differences due to the scientific category of the journal. However, in addition to homogeneity, an indicator must be valid, that is, having the ability to actually measure the impact of the journal. To approach the validity of the *pArt* indicator we have compared the ranking of journals generated within each category using the aforementioned, with those obtained using the most common indicators (*JIF5*, *JIF2*, *h5*, and *h3*).

The indicator with the highest level of homogeneity also proved to be the most restrictive – *pArt_10_2*. In order to assess the validity we decided to analyze the ranking generated through this indicator, and that generated through the least restrictive indicator – *pArt_40_5* – as it also has the same requirements: homogeneity between-groups (with an associated probability among the highest) and intra-group heterogeneity (supplementary material 2).

The indicator *pArt_10_2* has null values for 54 of the 278 journals. That is, the number of journals with no articles within the tenth citation percentile, taking a time window of two years, is 54 (26 in *IS & LS*, 14 in *OR & MS*, four in *Ophthalmology*, and ten in *Physics CM*). Therefore, sorting the journals within each category according to the *pArt_10_2* value, the indicator does not discriminate between journals once the relative positions 57, 65, 55, and 51 are achieved. However, using a higher percentile and time window, such as *pArt_40_5*, the number of journals with nulls is reduced to only two cases, specifically in the *IS & LS* category. This means that reducing the percentile citation and the time window makes the indicator less discriminative and it cannot differentiate between lower-impact journals.

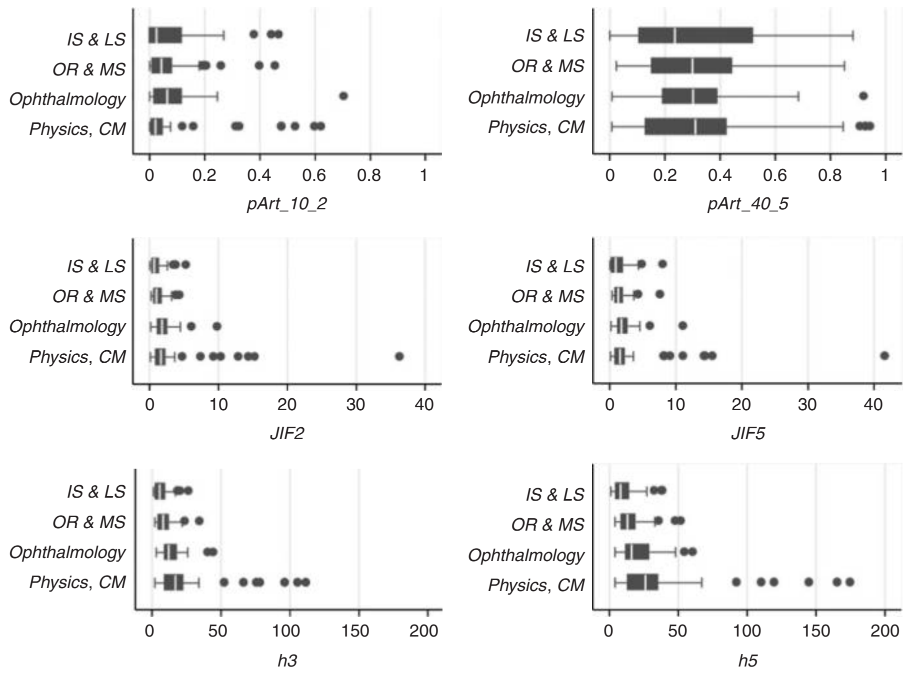
Figure 3 provides a comparison between the two proposed indicators and the four most common indicators. The percentage of highly cited articles (*pArt*) show closer, and therefore more comparable, distributions among scientific categories.

Table III shows the Spearman's rank correlation between all analyzed indicators, all of which are significant at 99 percent. It should be noted that the Spearman's rank correlation is less sensitive than Pearson's to extreme small or large values in the tails of the distribution, as it limits the analysis of such data to their position in the ranking.

All correlations are quite high and without significant differences between categories. The *pArt_q_t* indicator (percentage of highly cited articles) is robust to parameter changes in q and t. Indeed, correlations between *pArt_10_2* and *pArt_40_5* are above 0.80 in all categories. Moreover, in general, the correlation between these indicators and *JIF* or h are also high, exceeding 0.79, except for *OR & MS*. It seems, therefore, that the *pArt* indicator provides a similar dimension to that shown by h and *JIF*, and also that the *pArt* indicator is robust with respect to any of its parameters.

Conclusions

The *JIF* and h-index allow comparisons between journals of the same scientific category, but are not valid indicators to compare journals of different categories. The proportion of highly



Note: Observe that *pArt_40_5* produces distributions closer between categories and therefore is more comparable

Figure 3. Comparative analysis of indicator distributions among journal categories

	JIF2	JIF5	h3	h5	pArt_10_2	pArt_40_5
Information Science & Library Science						
JIF2	1					
JIF5	0.96	1				
h3	0.85	0.85	1			
h5	0.88	0.91	0.96	1		
pArt_10_2	0.84	0.83	0.88	0.85	1	
pArt_40_5	0.88	0.91	0.90	0.92	0.90	1
Operations Research & Management Science						
JIF2	1					
JIF5	0.94	1				
h3	0.86	0.85	1			
h5	0.83	0.88	0.95	1		
pArt_10_2	0.68	0.69	0.64	0.65	1	
pArt_40_5	0.87	0.94	0.82	0.85	0.80	1
Ophthalmology						
JIF2	1					
JIF5	0.95	1				
h3	0.84	0.81	1			
h5	0.83	0.82	0.97	1		
pArt_10_2	0.82	0.79	0.82	0.80	1	
pArt_40_5	0.85	0.85	0.86	0.86	0.84	1
Physics, Condensed Matter						
JIF2	1					
JIF5	0.97	1				
h3	0.83	0.83	1			
h5	0.83	0.85	0.98	1		
pArt_10_2	0.93	0.91	0.82	0.82	1	
pArt_40_5	0.97	0.97	0.81	0.81	0.92	1

Notes: All are significant at 99 percent. Very high correlation between all indicators and therefore very similar journal rankings within each category

Table III. Spearman's rank correlation between variables in journal categories

cited articles in a journal (*pArt*) appears to be an alternative citation impact indicator to achieve this end, since it is a relative measure that lacks the known limitations of other indicators in the literature, such as the journal size or the sensitivity to the high citation of a small number of articles. In fact, it makes it possible to identify what proportion of highly cited articles every journal publishes every year. So that we find journals with 0 percent of highly cited articles vs journals with more than 90 percent of highly cited articles, which can be a clear indicator of the impact of a journal.

The dilemma focuses on the limits set to consider an article as part of the group of the most cited, and the time window to be considered in the citation counts. After analyzing the behavior of 20 different indicators, depending on the citation percentile considered (10, 20, 25, 30, and 40 percent) and the time window in which the number of citations is counted (two to five years), the indicator that seems to best homogenize the categories is the one that considers a time window of two years and a citation level of 10 percent (*pArt_10_2*). This indicator is limited by its inability to discriminate within the set of the least cited journals in which none of the articles is within the most cited, due to its restrictions. However, as evidenced in the empirical application, the parameters chosen are not very important, as the results – in terms of homogeneity between categories – have been quite similar in the 20 scenarios studied.

The main advantage of the proposed index is that, in any case, each of the 20 indicators of the *pArt* family is a very simple but powerful indicator, that can be intuitively understood, and that enables comparisons between journals of different categories – as well as between journals of the same category – because of its proofed field normalization and validity.

Citation percentiles could also be used in impact factor and *h*-index, avoiding by this way their limitations related to comparisons among categories. Related to the impact factor, quartiles are the most commonly used baselines. They enable a comparison of quality in journals among fields (those inside the same quartile should be considered as having similar quality), but they do not allow to discriminate between journals of the same quartile, hindering the possibility of making a ranking of journals. Related to the *h*-index there is currently no normalization based on percentiles, and could be an interesting future work.

References

Althouse, B.M., West, J.D., Bergstrom, C.T. and Bergstrom, T. (2009), "Differences in impact factor across categories and over time", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 1, pp. 27-34.

Bensman, S.J. (2007), "Garfield and the impact factor", *Annual Review of Information Science and Technology*, Vol. 41 No. 1, pp. 93-155.

Bergstrom, C. (2007), "Eigenfactor: measuring the value and prestige of scholarly journals", *College and Research Libraries News*, Vol. 68 No. 5, p. 314.

Bordons, M. and Barrigón, S. (1992), "Bibliometric analysis of publications of Spanish pharmacologists in the SCI (1984-89). Part II. Contribution to subcategories other than «Pharmacology & Pharmacy» (ISI)", *Scientometrics*, Vol. 25 No. 3, pp. 425-446.

Bornmann, L. and Daniel, H.D. (2008), "What do citation counts measure? A review of studies on citing behavior", *Journal of Documentation*, Vol. 64 No. 1, pp. 45-80.

Costas, R. and Bordons, M. (2007a), "Una visión crítica del índice h: algunas consideraciones derivadas de su aplicación práctica", *El profesional de la información*, Vol. 16 No. 5, pp. 427-432.

Costas, R. and Bordons, M. (2007b), "The h-index: advantages, limitations and its relation with other bibliometric indicators at the micro-level", *Journal of Informetrics*, Vol. 1 No. 3, pp. 193-203.

Dorta-González, P. and Dorta-González, M.I. (2011), "Central indexes to the citation distribution: a complement to the h-index", *Scientometrics*, Vol. 88 No. 3, pp. 729-745.

Dorta-González, P. and Dorta-González, M.I. (2013a), "Comparing journals from different categories of science and social science through a JCR subject categories normalized impact factor", *Scientometrics*, Vol. 95 No. 2, pp. 645-672.

Dorta-González, P. and Dorta-González, M.I. (2013c), "Impact maturity times and citation time windows: the 2-year maximum journal impact factor", *Journal of Informetrics*, Vol. 7 No. 3, pp. 593-602.

Dorta-González, P., Dorta-González, M.I., Santos-Peñate, D.R. and Suárez-Vega, R. (2014), "Journal topic citation potential and between-category comparisons: the topic normalized impact factor", *Journal of Informetrics*, Vol. 8 No. 2, pp. 406-418.

Egghe, L. (2013), "Theoretical justification of the central area indices and the central interval indices", *Scientometrics*, Vol. 95 No. 1, pp. 25-34.

Egghe, L. and Rousseau, R. (2002), "A general framework for relative impact indicators", *Canadian Journal of Information and Library Science*, Vol. 27 No. 1, pp. 29-48.

Garfield, E. (1972), "Citation analysis as a tool in journal evaluation", *Science*, Vol. 178 No. 4060, pp. 471-479.

Garfield, E. (1979), "Is citation analysis a legitimate evaluation tool?", *Scientometrics*, Vol. 1 No. 4, pp. 359-375.

González-Pereira, B., Guerrero-Bote, V.P. and Moya-Anegón, F. (2009), "The SJR indicator: a new indicator of journals' scientific prestige", *Journal of Informetrics*, Vol. 4 No. 3, pp. 379-391.

Hirsch, J.E. (2005), "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences*, Vol. 102 No. 46, pp. 16569-16572.

Leydesdorff, L. (2009), "How are new citation-based journal indicators adding to the bibliometric toolbox?", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 7, pp. 1327-1336.

Leydesdorff, L. (2012), "Alternatives to the journal impact factor: I3 and the top-10% (or top-25%?) of the most-highly cited papers", *Scientometrics*, Vol. 92 No. 2, pp. 355-365.

Leydesdorff, L. and Bornmann, L. (2011), "How fractional counting of citations affects the impact factor: normalization in terms of differences in citation potentials among categories of science", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 2, pp. 217-229.

Leydesdorff, L. and Rafols, I. (2011), "Indicators of the interdisciplinarity of journals: diversity, centrality, and citations", *Journal of Informetrics*, Vol. 5 No. 1, pp. 87-100.

Moed, H.F. (2010), "Measuring contextual citation impact of scientific journals", *Journal of Informetrics*, Vol. 4 No. 3, pp. 265-277.

Moed, H.F., Colledge, L., Reedijk, J., Moya-Anegon, F., Guerrero-Bote, V., Plume, A. and Amin, M. (2012), "Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in an informed way", *Scientometrics*, Vol. 92 No. 2, pp. 367-376.

Rousseau, R. (2009), "What does the web of science five-year synchronous impact factor have to offer?", *Chinese Journal of Library and Information Science*, Vol. 2 No. 3, pp. 1-7.

Van Raan, A.F.J., Van Leeuwen, T.N., Visser, M.S., Van Eck, N.J. and Waltman, L. (2010), "Rivals for the crown: reply to Opthof and Leydesdorff", *Journal of Informetrics*, Vol. 4 No. 3, pp. 431-435.

Wagner, C., Roessner, J.D., Bobb, K., Klein, J., Boyack, K., Keyton, J., Rafols, I. and Börner, K. (2011), "Approaches to understanding and measuring interdisciplinary scientific research (IDR): a review of the literature", *Journal of Informetrics*, Vol. 5 No. 1, pp. 14-26.

Waltman, L. and Van Eck, N.J. (2013), "Source normalized indicators of citation impact: an overview of different approaches and an empirical comparison", *Scientometrics*, Vol. 96 No. 3, pp. 699-716.

Zitt, M. and Small, H. (2008), "Modifying the journal impact factor by fractional citation weighting: the audience factor", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 11, pp. 1856-1860.

Further reading

Dorta-González, P. and Dorta-González, M.I. (2013b), "Hábitos de publicación y citación según campos científicos: principales diferencias a partir de las revistas JCR", *Revista Española de Documentación Científica*, Vol. 36 No. 4, Article ID en012, available at: http://dx.doi.org/10.3989/redc.2013.4.1003

Corresponding author
Sara M. González-Betancor can be contacted at: sara.gonzalez@ulpgc.es

This article has been cited by:

1. Ricardo Brito, Alonso Rodríguez-Navarro. 2018. Research assessment by percentile-based double rank analysis. *Journal of Informetrics* **12**:1, 315-329. [Crossref]
2. Pablo Dorta-González, Yolanda Santana-Jiménez. 2018. Prevalence and citation advantage of gold open access in the subject areas of the Scopus database. *Research Evaluation* **27**:1, 1-15. [Crossref]