



Does the citation period have any effect on the informative value of selected citation indicators in research evaluations?

Marcel Clermont¹ · Johanna Krolak² · Dirk Tunger^{3,4}

Received: 21 February 2020 / Published online: 17 December 2020
© The Author(s) 2020

Abstract

To date, there has not been any definitive statement for bibliometric analyses from the scientific community as to how long to make the period for analysis in order to guarantee the maximum possible validity of citation analyses. Up to now, it has been particularly unclear as to whether different disciplines and citation indicators have any effect on what period for analysis to select. This is why this study aims to provide answers to these issues. Therefore, publications in the fields of business studies, biology, and medicine from 2007 to 2009 are surveyed. The citations from ten separate consecutive citation periods from 2009 up to 2018 are generated for these publications from Web of Science and selected citation indicators are calculated for them. The results show that the fundamental validity of citation indicators increases with time. However, if the goal of the analyses is to establish a ranking order, then even short citation periods are informative. Our results are stable in terms of the disciplines, although differences do occur in the citation indicators under consideration.

Keywords Bibliometric indicators · Citation metrics · Citation period · German universities · Research performance evaluation · Research ranking · Web of science · Citation time window

Introduction

There are frequent discussions in scientometric literature, as well as in higher education evaluations, about how long the period that a citation analysis is based on should be (e.g. Research Evaluation and Policy Project 2005, p. 20f.). According to Adams (2005), a short-term citation period of 1 or 2 years would make it possible to monitor

✉ Marcel Clermont
marcel.clermont@dhge.de

¹ Duale Hochschule Gera-Eisenach, Am Wartenberg 2, 99817 Eisenach, Germany

² Forschungszentrum Jülich GmbH, Zentralbibliothek, 52425 Jülich, Germany

³ Faculty of Information Science and Communication Studies, Institute of Information Management, TH Köln, Claudiusstraße 1, 50678 Cologne, Germany

⁴ Project Management Jülich, Center of Excellence “Analyses, Studies, Strategy”, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

and evaluate the citation rate at an early stage, while looking at short citation time periods can cause the data to be distorted. Publications have varying levels of ageing on the one hand (Glänzel and Schoepflin 1995) and durability on the other (Costas et al. 2010). With regard to investigating the research impact of German-language business schools, researchers such as Dyckhoff and Schmitz (2007) analyse comparatively long periods of 10 and 14 years, while Dilger (2010) made the citation period relatively short at two and a half years. Long citation periods bring with them the benefit of compensating for short, random fluctuations to a certain degree by lengthening the citation period. The drawback is the resulting greater effort required to survey the citation indicators.

With regard to determining suitable citation periods for generating informative and reliable results, it must be recognized that there is only a modest number of studies and that no uniform consensus regarding a suitable citation period for certain analyses can be derived from these studies. Furthermore, various mathematical methods are applied, such as half-lives, linear regression models, and correlation analyses (see, e.g., Abramo et al. 2019; Rovira-Esteva et al. 2019; Wang 2013). Only simple citation indicators are used in these studies. Beginning with the first citation analysis by Gross and Gross (1927), a wide array of citation indicators have been developed over time. In order to determine the effect of research services, indicators such as the h-index, developed by Hirsch (2005), are used. Because departments have differing habits relating to publication and citation (Glänzel et al. 2008), comparisons can only be made between disciplines if standardized indicators are used. As already noted, the variety of different citation indicator values in bibliometric studies caused by different citation periods has so far been ignored. This raises the following questions:

What citation period for citations of scientific publications will maximize the informative value of research performance analyses? Can differences be seen between different citation indicators?

Furthermore, studies are based either on whole data sets or on individual disciplines. The results of the latter studies vary depending on the discipline that is being analysed. So far, however, there is no comparative analysis of different citation periods for disciplines, which raises the following question:

Does the length of the citation period that can be shown to have the maximum informative value vary between disciplines?

Our paper's contribution is three folded and can be sum up as follows:

- Comprehensive inclusion of different citation periods to analyse probably changing processes during the timeframes
- Analysis of different relevant citation indicators, especially first time inclusion of a journal normalized indicator
- Consequent results consideration from the viewpoint of universities research evaluation

Our paper is structured as follows: The next section gives a brief account of the current state of research with regard to the stated research questions. Additionally, we concretize these research questions. Next, we show the design of our research, before presenting the results and analysing them concerning our research questions. The article concludes with the study's limitations and an outlook for further research.

Literature overview and detailed research questions

As already mentioned, the issue of the effect of the respective citation period has so far been the subject of some empirical research. In the following, we give a brief overview of the existing studies and their conclusions:

By using half-lives for 70,000 publications and a time frame of 55 years, Rovira-Esteva et al. (2019) determined that 50% of citations occurred within 5 years. From the 10th year, an increase in the citations of only a small percentage can be registered. The authors draw the conclusion that a period of between 5 and 10 years is sufficient to reach valid statements regarding the long-term perception of a publication. An approach for differentiated analysis of various fields can also be found within this study.

Abramo et al. (2019) base their analysis on 123,128 publications from Web of Science (WoS). They use a linear regression model and combine the journal impact factor and the number of citations. The authors conclude that a citation period of 3 years is sufficient to determine the long-term effects of scientific publications.

Wang et al. (2019) investigated the correlation between previous citations and new citations for 36 journals over a period of 10 years from 2008 until 2017. They use several citation periods to analyse the period, where the previous citations have the greatest influence on future citation of the regarded publications. It was found that the correlation between previous citations and new citations decrease significantly with increasing time windows. Thus, citations that are less recent in time show a greater influence on future citations of the publications than ‘older’ citations.

Liu et al. (2015) performed a comparison of journal impact factors with different citation periods and “peer-reviewed results” for ophthalmological journals. The highest correlations between peer-reviewed results and journal impact factor were found for a citation period of 3 and 4 years, which leads to the conclusion that a 3-year citation period is sufficient to reflect the actual impact of journals. Leydesdorff et al. (2013) came to the same conclusions, especially considering that a 2-year citation period does not cover disciplines with a slow citation build-up. Dorta-Gonzalez and Dorta-Gonzalez (2013), on the other hand, proposed to use the citation period with the highest average number of citations for the calculation of the journal impact factor.

Wang (2013) studies the correlations between all citations over 31 years and the cumulative citations in all possible periods for 746,460 publications from 1980 available on WoS. The results of the correlation analyses, i.e. the Spearman rank correlation coefficient between the total number of citations and the cumulative citations, of 0.754 for 3 years, 0.871 for 5 years, and 0.948 for 10 years, imply that even a citation period of 3 years can show long-term effects and establish a fundamental trend that will not change.

The delayed impact of publications, the so-called “sleeping beauties”, is used in the scientific community as a counter-argument for the consideration of short or medium-term citation periods for citation indicators. Glänzel et al. (2003) analyse publications from 1980 over a period of 21 years to gain insights into the significance of the share of publications with delayed impact. They find out that 76% of all publications were cited in the first 3 years after the publication year. Publications that are not cited in the period of 3 or 5 years were associated with a lower expected citation impact, so delayed uptake does not lead to a postponement of the citation process by several years. Publications that attract a lot of attention in later periods represent extreme cases (see Glänzel et al. 2003; van Raan 2004a), e.g., concerning Glänzel only 0.00014% of almost 450,000 publications were sleeping beauties. In 2008, Glänzel stated “that the particular choice of a standard citation

window cannot be made responsible for possibly negative results of an otherwise correct bibliometric evaluation study”.

Gonzalez and Gonzalez (2016) regard the long-term impact of publications. They postulate that the length of a citation period with the highest significance varies over time and by discipline. In the literature, citation periods with a fixed start and end are predominantly used. This does not reflect different levels of maturity in specific disciplines, and advantage or disadvantage some publications. From this, the authors conclude that a variable citation period is more suitable, which takes into account the differences between the disciplines to be analysed.

Based on the existing studies, it can be said that no consensus has been registered to date as to what effect the length of the citation period since the publication’s release has on the validity of the used citation indicators. The recommendations vary between 3 and 5 years. Since we wish to pursue these issues in this article, we concretise our research questions by formulating five more detailed ones:

Ketzler and Zimmermann (2013) find that the age of an article has a significant positive effect on the number of citations. It is therefore to be expected that the results of the citation indicators will differ for citation periods of different lengths, since analysis of longer citation periods involves an increased volume of citations. Consequently, if the citation period is extended, the number of uncited publications should decrease, causing the variance in the citations obtained for each publication to increase. Nicolaisen and Frandsen (2019) discovered that the ratio of non-citation frequency of publications reacts sensitively to the length of the citation period. Hu and Wu (2014) found that for a citation period of 12 years or longer, the percentage of papers never cited is stable and the probability of being cited in the future is very low for these publications. Using duration analyses, van Dalen and Henkens (2005) showed, contrary to the prevailing myth, that the chance of being cited decreases with age, non-citability does not play a role in the temporal distribution of citations. Based on these findings, our first research question is as follows:

(R1) Does the informative content of citation indicators increase for longer citation periods?

A publication ages based on the principle of phases of maturing and decline of its citation volume (Glänzel and Schoepflin 1995). We can conclude from this that once the publication transitions from the maturing phase to the decline phase, analysis of a longer citation period provides no additional informative value with regard to determining the impact of a publication. At this point, saturation is reached. This raises the question of whether the trend towards saturation can be detected at a certain point during the maturing process. Due to the heterogeneous citation behaviour of departments or research fields, the maturation of citations up to their distribution peak takes place according to different subject-specific time patterns (see, e.g., Abramo et al. 2011; Lehmann et al. 2006; Glänzel and Moed 2002). Since nobody knows the existence and the time of occurrence of high citation numbers of sleeping beauties, it is difficult to define the length of time until the citation matures (see, e.g., El Aichouchi and Gorry 2018; Song et al. 2018; van Raan and Winnink 2018; Teixeira et al. 2017). Based on this we formulate our second detailed research question:

(R2) Is there a citation period for which the long-term impact of citations can be determined and after which a longer period of analysis of the citation rate does not provide any additional informative value?

Within the framework of the above-mentioned maturing process, it is questionable, how the distribution of citations will change over time. Since Lotka’s law (Lotka 1926) we

know that citation distributions are skewed. There is a large number of studies, which confirm this law. These studies conclude that citations are distributed heterogeneously among publications. This means that a small percentage of publications generate the majority of citations and a large percentage of publications receive zero or only a few number of citations (see, e.g., Ruiz-Castillo and Costas 2014, 2018 Radicchi and Castellano 2012; Albarran et al. 2011; Albarran and Ruiz-Castillo 2011; Radicchi et al. 2008). According to Ruiz-Castillo and Costas (2018), the skewness of the individual citation distributions varies greatly within individual disciplines, but the average skewness in all disciplines is of a similar magnitude. With regard to the observation of the citation distribution over time, Li et al. (2013) found that the forms of citation distributions within each year under consideration are similar. We would like to combine the analysis of Ruiz-Castillo and Costas (2018) and Li et al. (2013) and raise the question:

(R3) To what extent does the uneven distribution of citation numbers differ for the three disciplines under consideration and does it change over time?

The previous mentioned studies are predominantly based on the number of citations. As already established, the research deficit relates especially to the issue of how far the citation period can be limited for other citation indicators. In this article, we will therefore address the h-index as a simple indicator and the *J*-factor as a standardized indicator.

The h-index has the advantage of not increasing uniformly with each additional citation, but instead only when the citations are changed in the h-core, which means that it is more robust in the face of fluctuations and citation peaks (Hirsch 2005). The h-index has already been subjected to a number of empirical studies with regard to its validity and compensating potential distortions by the bibliometrics community (see, e.g., Hirsch 2007; Alonso et al. 2009; Jensen et al. 2009; Honekopp and Khan 2012; Sharma et al. 2013). In order to eliminate the disadvantages (supposedly) inherent in the h-index, a variety of further indicators have been developed. On the basis of several studies into the relationships between the h-index and its variants, it has been established that the h-index is highly correlated with its variants (see, e.g., Bar-Ilan 2008; Schreiber 2008; Bornmann and Daniel 2009; Bornmann et al. 2011). As a consequence, considering another related and/or derived indicator in addition to the original h-index provides no added information value and we therefore do not consider any of these variants. H-index variants that take the currentness of the publications into consideration can also be disregarded, since the publications from the disciplines being studied come from a narrow range of publication ages. Variants that include the duration of the author's publication career or the number of co-authors are obsolete, as this study is not conducted at the level of individual scientists.

Schreiber (2015) analysed the development of the timed h-index as a function of the length of citation periods. He found that in most groups the median of all received citations is reached for a citation period of less than 5 years. With short citation periods of two (Fiala 2014) or 3 years (Van Raan 2006), publications that are still at the beginning of the process of citation accumulation are used to calculate the h-index. Consequently, there are still small differences in the citation volume of the individual publications, which reduces the selectivity of the h-index values. However, Pan and Fortunato (2014) conclude that the values of the h-index for a 5-year citation period are already based on more differentiated and higher citation numbers, which provides a higher selectivity of the h-Index values and fluctuations have less influence on the h-index values.

Based on previous studies, it remains unclear, how the value of the h-index will be affected when extending the citation period. As per the design of the h-index, its value does not necessarily and uniformly increase with each additional citation, since only core

publications have an effect on this index. This means that the increasing volume of citations that comes with looking at growing citation periods does not cause the h-index to increase proportionally. As a consequence, looking at longer citation periods should imply only limited added informative value, which is why we consider the following:

(R4) Does the added informative value of the h-index remains limited in comparison with the absolute number of citations and the citations per publication when the citation period is extended?

To permit interdisciplinary comparisons to evaluate research performance, differences specific to the disciplines are compared while using methods of standardization. A distinction is drawn between field-normalized and journal-normalized methods. In this regard, Glänzel et al. (2008) compare the characteristics of the field-normalized indicator Normalized Mean Citation Rate (NMCR) for 676 European institutions and two citation periods. Such field-normalized indicators compare the citation rate of a publication with all publications from the same discipline. In the above-mentioned study, there are high correlations between the indicators for all institutions for the 3-year and 5-year citation periods. As a consequence, it was concluded that a 3-year citation period is sufficient to reach valid statements on the effects of citations with regard to the NMCR indicator.

A major weakness of field normalization is the need for the a priori classification of disciplines and sub-disciplines. Any delineation between disciplines will always prompt debates despite the presence of different classification systems. Journal-normalized methods, by contrast, are based on the idea of taking mean citation rates for the available publications from the same year of publication and of the same document type of the unit being studied, and correlating these with the mean citation rate of all publications from the respective journal. This process is applied to each journal in question. As a consequence, normalization is independent of the choice of field classification.

Because of this advantage and the fact that a study of journal-normalized citation indicators has not been carried out to date, we use the *J*-factor developed by Ball et al. (2009) in this publication. The *J*-factor analyses the ratio between the citation rates of the publications of a unit *I* being studied (e.g. a department or university) in a journal *z* and the citation rates of the publications from a reference group *St* in the journal *z*. This means that for each unit being studied, the number of citations per publication CPP^I in a particular journal *z* is correlated with the mean number of citations per publication CPP^{St} for all publications in the journal *z* with the same year of publication and document type. This ratio is weighted by the proportion of publications $P^I(z)$ of a unit being studied in this journal in relation to all publications $P^I(\text{all})$ of the unit being studied in the evaluation period. The total of the results for each journal generates the *J*-factor for a unit:

$$J(I, St) = \sum_z \frac{CPP^I(z)}{CPP^{St}(z)} * \frac{P^I(z)}{P^I(\text{all})}$$

The *J*-factor implies a relative evaluation of the specific citation performance in comparison to the citation performance of the reference group. A *J*-factor of 1 means a citation performance equal to the reference group; a *J*-factor greater than 1 indicates an above-average citation performance in relation to the reference group, while a *J*-factor less than 1 signals a below-average citation performance.

One benefit of the *J*-factor is that the analysis can be performed for a period of unlimited length. The disadvantage is that the unit being studied can also publish in journals unrelated to the discipline, which then show different citation behaviours and can cause

distortions of the data as a result. In contrast to simple citation indicators, an increase in the J -factor cannot be achieved simply by increasing the citation numbers, since the higher number of citations also appears in the reference group. This means that the behaviour of the J -factor depends on the citation development of the respective reference group even when the number of citations in a unit being studied increases. Accordingly, the development of the J -factor as the number of citations increases is barely calculable. However, we assume that the J -factor is more informative, the more citations are incorporated into the factor's calculation. The number of citations from both the unit being studied and the reference group will continue to increase as the citation period grows longer. Accordingly, with regard to field-normalized citation indicators, we ask ourselves if looking at longer citation periods leads to added informative value for the J -factor:

(R5) Does the informative value of the J -factor increase as the citation period grows longer?

As already noted, disciplines differ in their citation behaviour, regarding both the citation volume and the timespan in which they display their citation potential. Because of this, short, or even too short, citation periods for some disciplines can lead to an inaccurate evaluation (Glänzel et al. 2008). Accordingly, the length of the citation period, adapted to the citation behaviour of the respective discipline, should be selected (see, e.g., Wang 2013; Waltman et al. 2011). Based on the preceding findings that disciplines differ in their citation behaviour and require different timespans to mature their citation potential as a consequence, we finally question the following:

(R6) Do the results concerning the length of the citation periods with regard to the effect of the research differ depending on the discipline being studied?

Study design

Data collection

The starting point for our data collection process are publications from the so-called 'CHE ranking', carried out by the Center for Higher Education Development (CHE). The CHE provides a list of all chairs and institutes from different disciplines of universities in Germany. For these chairs and institutes, the CHE creates a discipline-specific ranking in a 3-year-cycle. One indicator used is the absolute number of publications during the regarded 3 years (Berghoff et al. 2009). The bibliometrics team at Forschungszentrum Jülich annually prepares the publication data for these chairs and institutes. Therefore, all publications are determined top-down, starting at the university-level and going down to a special institute or chair, according to the work-done-at-method in WoS. If several institutions co-authored a publication, each participating institution is credited once for the publication. A fractional counting is not applied. In this way, publications are assigned to fields without using a classification. For our study, we took the CHE publication list from the chairs and institutions of business administration, biology, and medicine, for the period 2007 to 2009. Only the document types "Article" and "Review" are taken into account.

Based on these publication data, we determined the citation data for the three stated disciplines in the WoS local installation of the Competence Centre for Bibliometrics.¹ We will look at ten separate consecutive citation periods, starting in 2009 up to 2018. The first citation period is for 3 years and runs from 2007 up to and including 2009, the second is from 2007 up to 2010, the third from 2007 up to 2011 and so on. Based on this absolute publication and citation data we calculate for each chair or institution and for each citation period the following indicators: the citations per publication, the *h*-index and the *J*-factor. The process of data collection and the represented level of the study are not on a field level, also it looks like, but instead of this, they are on an institutional level. Furthermore, there is no scientific reason in the survey methodology why the indicators should not work on a field level. As long as the examined subset is part of the overall benchmark, *h*-Index and *J*-factor works also on a field level (see van Raan 2004b; Malesios and Psarakis 2014).

Analysis instruments

In order to determine informative citation periods, various mathematical and/or statistical methods were used in the bibliometric literature, such as cited half-lives, linear regressions, and correlation coefficients. A cited half-life means the timespan after which half of all the citations received by a publication occur (De Bellis 2009). Linear regressions are used to make forecasts and describe relationships between two variables. However, neither method is suitable for testing our hypotheses.

By contrast, correlation coefficients describe the relationship between two or more variables and therefore are generally suitable for addressing our research questions. However, to do this, we need to define what we mean when we talk about a citation indicator's added information value below. This means that increasing an indicator, e.g. increasing the absolute number of citations, will always provide added information value, since this higher figure reflects the cited paper's perception in the community. In this respect, it naturally makes a difference whether the paper has been cited 5 or 500 times.

However, performance indicators are frequently used in higher education evaluations to generate rankings of individual scientists, departments, or universities. For this purpose, as already mentioned, CHE has developed a research ranking, which has undergone critical analysis by Clermont and Dirksen (2016). In rankings, the precise characteristic of the indicator is irrelevant to the viewer. What is important is its rank position. Therefore, there is added information value in this kind of analysis when the rankings differ strongly from each other as the citation period varies. Ranking correlation coefficients are suitable for such a study. There are two known methods for calculating this: Spearman's ρ (Spearman 1904) and Kendall's τ (Kendall 1938).

According to Spearman's method, pairs of ranks are formed and compared with each other (Xu et al. 2013). However, problems occur when there are ties, i.e. when two values have the same ranking (Schendera 2004). As a solution to this, an arithmetic mean is formed, which has negative effects on the results' informative value. Furthermore, Spearman assumes that the differences present in the rankings are equivalent, i.e. that the difference between first and second place is equivalent to the difference between last and

¹ In order to allow the scientific community to perform comprehensive and reliable bibliometric analyses, the German Federal Ministry of Education and Research (BMBF) launched the German Competence Centre for Bibliometrics (reference no. 01PQ17001) in 2008. Further details are available at <http://www.bibliometrie.info>.

second-to-last. We were not able to establish a relationship of this kind with our collected data.

In contrast to Spearman, Kendall does not assume identical differences in the rankings. Kendall's ranking correlation coefficient assumes a given ranking based on an indicator used (e.g. a ranking based on the number of citations between 2007 and 2009). It is measured then, how often a ranking based on another indicator (e.g. a ranking based on the number of citations between 2007 and 2018) "breaks" the initial ranking. The resulting number is divided by the number of all possible rankings, so that the coefficient is between -1 and $+1$. A value of $+1$ means, that all ranking positions are identical; a value of -1 assumes that the ranking positions are reversed.

Let us assume a ranking of all universities under consideration n based on the initial indicator x_1 . Starting from the original ranking, a pair comparison with all subsequent rankings is made, starting with the first ranking position of the initial ranking. Thus $0.5n(n-1)$ pair comparisons are carried out. Concerning all pair comparisons it is analysed, how many concordant and discordant pairs exist. If the ranking positions based on x_2 are greater than the initial regarded ranking position based on x_1 , this pair is concordant. Accordingly, a pair is discordant, if the following ranking position is smaller than the regarded one. If we denote the number of concordant pairs as C and the number of discordant pairs as D , Kendall's τ is defined as follows:

$$\tau = \frac{C - D}{0.5n(n - 1)}$$

If there are bindings, i.e. identical rank positions, the above formula have to be extended, since bindings are neither concordant nor discordant. We refrain from an explicit presentation here and refer to the relevant statistical literature.

An advantage of Kendall's ranking correlation coefficient is that all rankings can be compared with each other, and not just individual ranking pairs. For our study, this would mean that all ten measured values for the ten citation periods can be compared with each other for individual indicators. According to a study by Xu et al. (2013), Kendall's method is also preferable with regard to a smaller number of data sets and outliers in the data. This corresponds to our data situation with a skewed distribution (see also the next section), potential outliers, and a low volume of data from 86 universities for business studies, 68 for biology, and 44 for medicine. On the basis of the considered properties of the two methods and the stated characteristics of our data situation, we use the ranking correlation coefficient by Kendall.

Results

Descriptive analysis

The descriptive analyses of the three considered disciplines with regard to the ten generated citation periods for the individual indicators are shown in Tables 1, 2 and 3. Here, the position of the data is given in more detail with the arithmetic mean, median, minimum, and maximum values for each indicator and citation period. The distributional width of the data is shown based on the standard deviation and variation coefficient.

A consistent pattern emerges for the arithmetic mean of the citation indicators of all disciplines and citation periods. The indicators "number of citations" and "citations

Table 1 Descriptive analysis: business administration

		Arithmetic mean	Median	Minimum	Maximum	Standard deviation	Variation coefficient
Citations 2007-	2009	63	13	0	332	52.166	0.831
	2010	131	25	0	717	109.438	0.834
	2011	216	41	0	1201	180.593	0.834
	2012	311	60	0	1737	261.639	0.842
	2013	414	74	0	2337	348.391	0.841
	2014	521	97	0	2987	443.511	0.852
	2015	640	125	0	3702	550.202	0.860
	2016	765	142	0	4524	666.719	0.871
	2017	882	160	0	5295	776.339	0.881
	2018	913	167	0	5523	807.527	0.885
Citations per Paper 2007-	2009	1.202	0.875	0	11.000	1.395	1.161
	2010	2.446	1.917	0	16.000	2.350	0.961
	2011	4.003	3.167	0	21.000	3.479	0.869
	2012	5.794	4.500	0	30.000	5.308	0.916
	2013	7.751	5.958	0	39.500	7.243	0.935
	2014	9.798	7.313	0	57.200	9.677	0.988
	2015	12.027	8.723	0	78.000	12.236	1.017
	2016	14.416	10.440	0	99.200	15.091	1.047
	2017	16.708	11.429	0	126.000	18.292	1.095
	2018	17.374	12.000	0	136.700	19.502	1.123
h-Index 2007-	2009	2	2	0	8	1.580	0.646
	2010	4	3	1	13	2.461	0.686
	2011	5	4	1	17	3.368	0.738
	2012	5	5	1	22	4.065	0.753
	2013	6	5	1	25	4.590	0.766
	2014	7	6	1	29	5.272	0.806
	2015	7	6	1	31	5.684	0.813
	2016	7	6	1	34	6.188	0.835
	2017	8	7	1	37	6.702	0.862
	2018	8	7	1	38	6.821	0.869
J-Factor 2007-	2009	0.927	0.855	0	2.974	0.611	0.658
	2010	1.024	0.958	0	3.200	0.589	0.575
	2011	1.069	1.019	0	3.000	0.562	0.526
	2012	1.065	1.020	0	2.683	0.560	0.526
	2013	1.078	1.033	0	2.598	0.559	0.519
	2014	1.074	0.987	0	2.561	0.556	0.517
	2015	1.083	0.995	0	2.909	0.575	0.530
	2016	1.088	0.991	0	2.667	0.571	0.525
	2017	1.090	1.043	0	2.667	0.555	0.509
	2018	1.089	1.029	0	2.662	0.554	0.508

Table 2 Descriptive analysis: biology

		Arithmetic mean	Median	Minimum	Maximum	Standard deviation	Variation coefficient
Citations 2007-	2009	1967	698	0	3763	893.579	0.454
	2010	5446	1925	3	11,343	2504.199	0.460
	2011	10,648	3808	7	23,383	4897.748	0.460
	2012	16,403	6177	12	37,296	7598.687	0.463
	2013	22,029	8332	25	52,112	10,245.213	0.465
	2014	27,414	10,426	27	66,760	12,791.395	0.467
	2015	22,029	8332	25	52,112	10,245.213	0.465
	2016	37,559	14,224	34	96,952	17,751.230	0.473
	2017	42,248	16,048	40	111,674	20,072.216	0.475
	2018	43,515	16,507	41	115,919	20,713.527	0.476
Citations per Paper 2007-	2009	1.489	1.463	0.000	2.837	0.526	0.353
	2010	4.115	4.005	1.000	7.654	1.276	0.310
	2011	8.056	7.924	2.333	14.504	2.407	0.299
	2012	12.383	12.384	4.000	22.354	3.749	0.303
	2013	16.691	16.695	5.625	29.821	4.932	0.295
	2014	20.757	20.874	7.375	38.837	6.273	0.302
	2015	16.691	16.695	5.625	29.821	4.932	0.295
	2016	28.385	28.338	9.958	56.964	8.810	0.310
	2017	31.939	31.952	11.333	65.870	10.009	0.313
	2018	32.887	32.868	11.792	68.107	10.349	0.315
h-Index 2007-	2009	11	11	2	24	4.873	0.431
	2010	19	18	1	41	8.405	0.449
	2011	26	25	2	55	11.484	0.437
	2012	34	32	3	72	14.989	0.443
	2013	41	39	3	86	18.045	0.444
	2014	47	44	3	97	20.915	0.449
	2015	41	39	3	86	18.045	0.444
	2016	56	53	3	123	25.396	0.455
	2017	60	57	3	134	27.387	0.458
	2018	61	57	3	137	27.829	0.458
J-Factor 2007-	2009	0.835	0.788	0.000	2.152	0.315	0.378
	2010	1.089	1.101	0.353	1.545	0.167	0.153
	2011	1.100	1.115	0.654	1.765	0.157	0.143
	2012	1.082	1.098	0.641	1.620	0.147	0.136
	2013	1.082	1.101	0.665	1.502	0.125	0.115
	2014	1.071	1.079	0.702	1.423	0.121	0.113
	2015	1.060	1.070	0.651	1.359	0.127	0.120
	2016	1.056	1.075	0.612	1.323	0.126	0.120
	2017	1.054	1.069	0.611	1.297	0.124	0.117
	2018	1.054	1.070	0.602	1.296	0.123	0.116

Table 3 Descriptive analysis: medicine

		Arithmetic mean	Median	Minimum	Maximum	Standard deviation	Variation coefficient
Citations 2007-	2009	5886	5282	810	16,934	3807.862	0.647
	2010	16,420	14,559	2260	47,878	10,662.994	0.649
	2011	32,121	28,174	4619	93,899	20,968.791	0.653
	2012	49,345	42,794	7291	143,844	32,323.542	0.655
	2013	66,090	56,788	9618	192,697	43,379.871	0.656
	2014	81,834	69,773	11,839	238,885	53,916.250	0.659
	2015	96,691	81,866	13,905	282,714	63,957.474	0.661
	2016	110,884	93,206	15,813	325,730	73,668.724	0.664
	2017	123,946	103,502	17,516	365,260	82,664.699	0.667
	2018	127,239	106,184	17,984	375,037	84,919.709	0.667
Citations per Paper 2007-	2009	1.185	1.171	0.728	1.511	0.186	0.157
	2010	3.289	3.267	2.118	4.229	0.519	0.158
	2011	6.418	6.321	4.119	8.293	1.054	0.164
	2012	9.849	9.744	6.227	12.813	1.671	0.170
	2013	13.187	13.002	8.384	17.227	2.273	0.172
	2014	16.314	16.081	10.362	21.463	2.862	0.175
	2015	19.255	18.988	12.199	25.476	3.444	0.179
	2016	22.059	21.688	13.897	29.576	4.013	0.182
	2017	24.637	24.137	15.477	33.431	4.547	0.185
	2018	25.286	24.773	15.875	34.413	4.684	0.185
h-Index 2007-	2009	24	24	11	42	6.938	0.293
	2010	40	39	19	69	11.686	0.295
	2011	57	55	25	99	17.096	0.301
	2012	73	72	32	127	22.324	0.305
	2013	87	86	40	151	26.327	0.301
	2014	100	99	45	174	30.475	0.304
	2015	112	110	50	189	33.396	0.298
	2016	122	120	55	203	36.318	0.298
	2017	131	130	59	216	38.952	0.298
	2018	133	132	61	220	39.629	0.299
J-Factor 2007-	2009	0.864	0.841	0.657	1.748	0.183	0.212
	2010	1.157	1.147	0.992	1.406	0.073	0.063
	2011	1.156	1.158	1.013	1.317	0.060	0.052
	2012	1.146	1.149	1.043	1.275	0.054	0.047
	2013	1.142	1.141	1.032	1.284	0.053	0.047
	2014	1.136	1.133	1.030	1.283	0.053	0.047
	2015	1.131	1.128	1.028	1.282	0.053	0.047
	2016	1.128	1.122	1.017	1.281	0.054	0.048
	2017	1.125	1.120	1.010	1.275	0.055	0.049
	2018	1.123	1.118	1.008	1.274	0.055	0.049

per publication” increase strongly in the first 4 years. For business administration, the increase between the citation periods 2007 to 2009 and 2007 to 2010 is 109% (biology and medicine: approx. 178%), between 2007 to 2010 and 2007 to 2011 65% (biology and medicine: approx. 95%), between 2007 to 2011 and 2007 to 2012 44% (biology and medicine: approx. 54%) and between 2007 to 2012 and 2007 to 2013 33% (biology and medicine: approx. 34%). From the following periods onwards, the growth levelled off at approx. 20% and fell below 20% afterwards. In the last analysed citation period (from 2007 to 2018), there is only a growth of about 3% for all regarded disciplines.

Similar tendencies can be seen for the h-index. However, as discussed in our literature overview, not being on the same scale as the citations and citations per paper. Thus, the metrics for business administration increase between the citation periods from 2007 to 2009 and 2007 to 2010 by 47% (Biology and Medicine: approx. 65%), between 2007 to 2010 and 2007 to 2011 by 27% (Biology and Medicine: approx. 40%), between 2007 to 2011 and 2007 to 2012 by 18% (Biology and Medicine: approx. 28%), between 2007 to 2012 and 2007 to 2013 by 10% (Biology and Medicine: approx. 20%). From the following citation period on, the growth is below 10% (biology and medicine: approx. below 15%) and falls for all three disciplines to 1% in the following periods. The smaller rise for the h-index can be attributed to the fact that this metric, related to the length of the citation periods, exhibits a certain level of robustness and the fact that only a certain subset of the citations (h-core) is taken into account when calculating the h-index.

The results regarding the arithmetic mean of the *J*-factor paint a different picture. It shows that the *J*-factor between 2007 to 2009 and 2007 to 2010 for biology and medicine increase around 30%, since for business administration it is smaller (around 10%). For the rest of the citation periods, the *J*-factor is stagnant around – 1% and + 1% for all three disciplines.

If we take the median into account, we see that this diverges considerably for the absolute number of citations. Furthermore, the characteristics of the median are far below those of the arithmetic mean. This leads us to conclude that the distribution of the citation numbers has been right-skewed, in line with; Saam and Reiter (1999); Seglen (1992); Price (1965). This result is also confirmed by the high standard deviations and variation coefficients of all three disciplines, which indicate a high degree of distribution. To analyse this, we refer to the Lorenz curve, which displays the relative concentration of the citations’ frequency distributions, i.e. it shows the unevenness of a distribution. In Fig. 1 the cumulative percentage of the universities is shown on the abscissa and the cumulative percentage of the number of citations is shown on the ordinate. Thus, each point on the curve shows what percentage of the universities obtains what proportion of the overall number of citations. The (drawn) bisector is a hypothetical construct as a standard of comparison for the citation distributions of the three disciplines, it cannot be achieved in reality. It can be interpreted as an equal distribution of citations, i.e. each university receives the exact same number of citations.

For business administration, it can be seen that 80% of the universities receive around just 30% of the citations for their publications, while the remaining 20% receive almost 70% of the citations. The citations of biology and medicine are more uniform distributed than the citations of business administration, concerning biology 80% of the universities receives 50% of the citation (60% in medicine), while 20% gets the other 50% (40% in medicine).

In order to specify the uneven distribution more accurately, we calculate the “Gini coefficient”. It provides information about the degree of unevenness of the citation

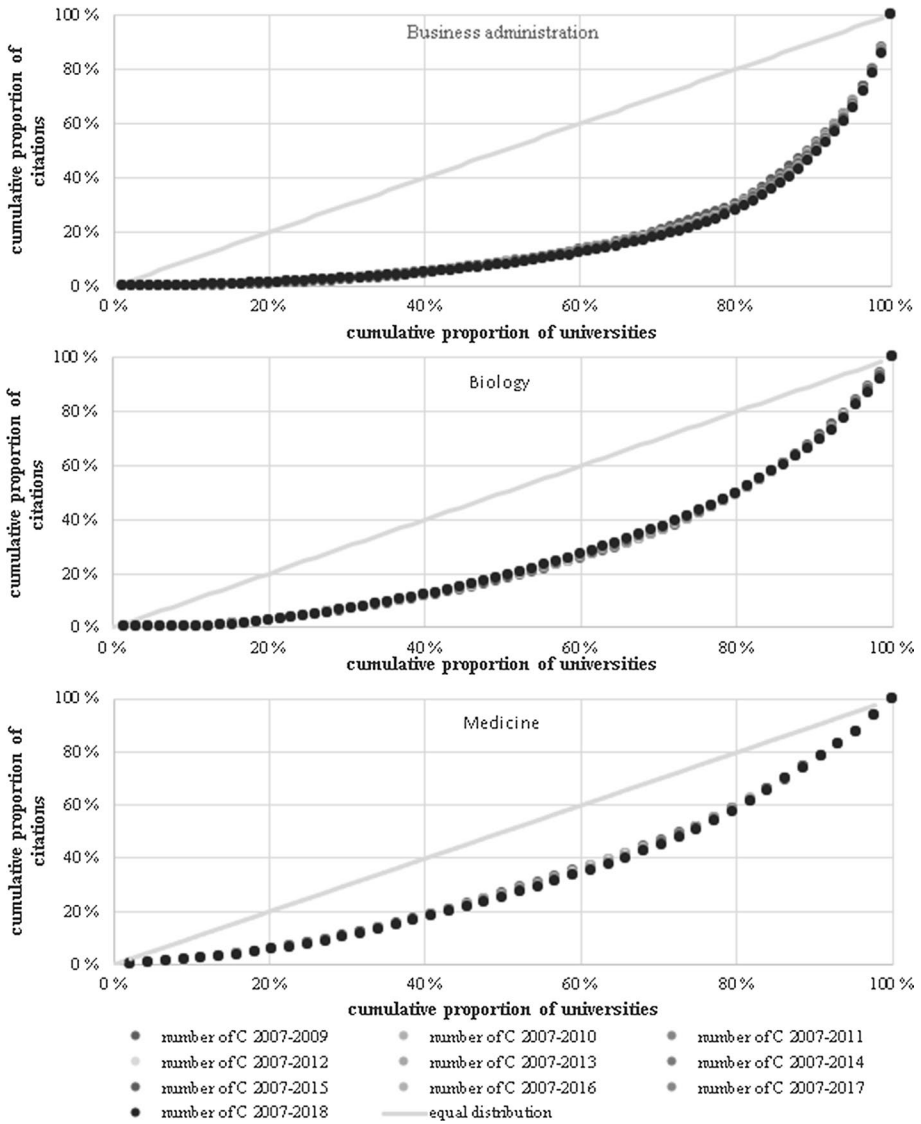


Fig. 1 Lorenz curves of citation distributions for all citation periods and disciplines

distributions and is determined based on the ratio of the area bounded by the 45° line and the Lorenz curve to the overall area (Pyatt 1976):
$$= \frac{2 \sum_{i=1}^n ix_i}{n \sum_{i=1}^n x_i} - \frac{n+1}{n}.$$

A Gini coefficient of 0 implies that all the citations are evenly distributed among the universities, while a Gini-coefficient of 1 means that a single university has received all the citations. Table 4 lists the Gini coefficients found for the three disciplines and citation periods. These once again confirm the statements made based on the Lorenz curves.

The disciplines’ different publication and citation behaviours can be attributed to different citation habits and the differing ratio of coverage in WoS. For example, business

Table 4 Gini coefficients

Discipline/citation period	2007–2009	2007–2010	2007–2011	2007–2012	2007–2013	2007–2014	2007–2015	2007–2016	2007–2017	2007–2018
Business administration	0.657	0.662	0.658	0.662	0.660	0.665	0.668	0.670	0.673	0.674
Biology	0.472	0.471	0.467	0.468	0.467	0.466	0.467	0.467	0.467	0.467
Medicine	0.344	0.348	0.350	0.352	0.353	0.355	0.356	0.357	0.358	0.359

administration has the lowest number of publications and citations in the present comparison of the disciplines and, based on the Lorenz curve, has the most uneven citation distribution. On the other hand, biology already shows 19-fold the number of publications and 35-fold the number of citations of business administration and the Lorenz curve runs more equally than for business administration. Medicine publishes 5-fold (105-fold) and cites 4-fold (150-fold) as much as biology (business administration) and the Lorenz curve has the most equal distribution of the three disciplines. Based on our dataset, we are able to identify discipline-dependent differences in the distributional skewness. With respect to the observation of the uneven distribution over time, the Gini coefficient shows minimal tendencies towards uniformity of the citation distribution. While the Gini coefficient for business administration and medicine only increases by 3% and 4%, the Gini coefficient for biology even decreases by 1%. These findings are confirmed by the course of the Lorenz curves over time.

Correlation analyses

An extract of the values from the correlation analyses for the citation indicators being studied and the included citation periods with regard to the individual disciplines can be found in Tables 5, 6 and 7. In the following, we look in particular at the correlations of the individual citation periods with the last period for the years 2007 to 2018. Resulting correlations of over 90% give sufficient indication that further survey periods only allow for lower additional information. Therefore, we assume that an appropriate citation period—especially from the point of view of a ranking—exists precisely when the ranking correlation coefficient exceeds this value of 90%.

Intertemporal evaluation of the correlations of citation indicators from the ten consecutive citation periods confirms the previously established results. A rank correlation of over 90% for the number of citations for business administration between the citation periods of 2007 to 2012 and 2007 to 2013 can be registered, for biology between 2007 to 2010 and 2007 to 2011 and for medicine between 2007 to 2009 and 2007 to 2010. While there are already rank correlation coefficients of more than 80% for biology and medicine from the first citation period on, there are more changes in business administration between 2009 and 2011. We have already seen in the descriptive analysis that the number of citations rises stronger for business administration during the first years. It seems that the time between publication and citing is longer in business administration than in biology and medicine.

Looking at the citations per paper, it can be seen that there are rank correlation coefficients for business administration and biology of over 90% from 2014 onwards. For medicine, however, rank correlations of over 90% are already achieved in the citation period 2012.

For the h-index it turns out that for all three disciplines, rank correlations of over 90% exist from 2012 onwards. In most cases there is also a high correlation of over 90% between the h-index and the absolute number of citations from all the citation periods, as well as between the h-index for each citation period. The rank correlation coefficients also have high values. Accordingly, there is no great change between rankings for the h-index depending on what citation period is looked at. The high rank correlations between the absolute number of citations and the h-index indicate that the rankings resulting from both indicators are similar (see also the results in Clermont et al. 2017).

Table 5 Extract of linear correlations and ranking correlations for the citation indicators in business administration

2007–2009	Pearson correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.988***	0.956***	0.916***	0.698***	0.620***	0.501***	0.948***	0.924***	0.900***	0.539***
2007–2010	Kendall correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.882***	0.789***	0.742***	0.602***	0.585***	0.554***	0.861***	0.807***	0.790***	0.490***
2007–2010	Pearson correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.986***	0.986***	0.958***	0.870***	0.815***	0.721***	0.966***	0.945***	0.837***	0.778***
2007–2011	Kendall correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.879***	0.879***	0.826***	0.811***	0.783***	0.747***	0.893***	0.856***	0.634***	0.579***
2007–2011	Pearson correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.995***	0.995***	0.972***	0.934***	0.892***	0.812***	0.983***	0.962***	0.925***	0.891***
2007–2012	Kendall correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.936***	0.936***	0.880***	0.889***	0.857***	0.825***	0.930***	0.887***	0.771***	0.712***
2007–2012	Pearson correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.999***	0.999***	0.982***	0.968***	0.936***	0.864***	0.992***	0.979***	0.963***	0.924***
2007–2013	Kendall correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.966***	0.966***	0.904***	0.908***	0.887***	0.857***	0.949***	0.906***	0.859***	0.765***
2007–2013	Pearson correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.989***	0.989***	0.989***	0.995***	0.980***	0.931***	0.985***	0.985***	0.961***	0.958***
2007–2013	Kendall correlation	2007–2010	2007–2013	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
		0.928***	0.928***	0.928***	0.949***	0.929***	0.884***	0.929***	0.929***	0.825***	0.819***

Table 5 (continued)

	2007–2010	2007–2013	2007–2018	2007–2014	2007–2015	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
2007–2014 Pearson correlation			0.993***	0.992***	0.958***				0.993***	0.970***	0.968***
Kendall correlation			0.944***	0.949***	0.909***				0.962***	0.862***	0.860***
2007–2015 Pearson correlation			0.997***		0.982***				0.995***	0.986***	0.982***
Kendall correlation			0.960***		0.937***				0.967***	0.901***	0.897***
2007–2016 Pearson correlation			0.999***		0.999***				0.997***	0.995***	0.993***
Kendall correlation			0.982***		0.966**				0.980***	0.936***	0.931***
2007–2017 Pearson correlation			0.999***		0.999***				0.999***		0.999***
Kendall correlation			0.994***		0.987***				0.996***		0.977***
				(b) Citations per Paper			(c) h-Index				(d) J-Factor

*Significant at 10% level

**Significant at 5% level

***Significant at 1% level

Table 6 Extract of linear correlations and ranking correlations for the citation indicators in biology

2007–2009	Pearson correlation	0.996***	0.990***	0.962***	0.881***	0.840***	0.772***	0.981***	0.964***	0.950***	0.210*	0.298**	0.277**
2007–2010	Kendall correlation	0.949***	0.924***	0.880***	0.698***	0.652***	0.588***	0.902***	0.850***	0.830***	0.152*	0.141*	0.126*
2007–2010	Pearson correlation	0.998***	0.998***	0.979***	0.960***	0.933***	0.873***	0.979***	0.965***	0.965***	0.745***	0.772***	0.723***
2007–2011	Kendall correlation	0.973***	0.973***	0.919***	0.811***	0.764***	0.691***	0.885***	0.861***	0.861***	0.586***	0.585***	0.569***
2007–2011	Pearson correlation	0.988***	0.988***	0.993***	0.993***	0.977***	0.931***	0.991***	0.979***	0.979***	0.917***	0.904***	0.866***
2007–2012	Kendall correlation	0.940***	0.940***	0.932***	0.932***	0.870***	0.787***	0.939***	0.899***	0.899***	0.818***	0.768***	0.735***
2007–2012	Pearson correlation	0.992***	0.992***	0.994***	0.994***	0.961***	0.961***	0.997***	0.989***	0.989***	0.928***	0.941***	0.907***
2007–2013	Kendall correlation	0.961***	0.961***	0.932***	0.932***	0.838***	0.838***	0.964***	0.922***	0.922***	0.890***	0.840***	0.805***
2007–2013	Pearson correlation	0.995***	0.995***	0.996***	0.996***	0.980***	0.980***	0.993***	0.993***	0.993***	0.944***	0.944***	0.941***
2007–2018	Kendall correlation	0.976***	0.976***	0.954***	0.954***	0.895***	0.895***	0.944***	0.944***	0.944***	0.855***	0.855***	0.824***

Table 6 (continued)

	2007– 2010	2007– 2011	2007– 2018	2007– 2012	2007– 2015	2007– 2018	2007– 2010	2007– 2013	2007– 2016	2007–2018
2007– 2014 Pearson correla- tion	0.997***	0.998***	0.989***	0.996***	0.984***	0.972***				
Kendall correla- tion	0.981***	0.961***	0.925***	0.963***	0.917***	0.878***				
2007– 2015 Pearson correla- tion	0.995***	0.980***	0.980***	0.993***	0.992***	0.977***				
Kendall correla- tion	0.976***	0.895***	0.895***	0.944***	0.944***	0.900***				
2007– 2016 Pearson correla- tion	0.999***	0.999***	0.999***	0.999***	0.999***	0.993***				
Kendall correla- tion	0.996***	0.973***	0.973***	0.969***	0.969***	0.949***				
2007– 2017 Pearson correla- tion	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***				
Kendall correla- tion	0.998***	0.997***	0.997***	0.991***	0.991***	0.985***				
	(a) Citations	(b) Citations per Paper	(c) h-Index	(d) J-Factor						

*Significant at 10% level

**Significant at 5% level

***Significant at 1% level

Table 7 Extract of linear correlations and ranking correlations for the citation indicators in medicine

	2007–2010	2007–2018	2007–2011	2007–2013	2007–2018	2007–2010	2007–2013	2007–2018	2007–2013	2007–2017	2007–2018
2007–2009 Pearson correlation	0.998***	0.986***	0.929***	0.879***	0.829***	0.989***	0.974***	0.965***	0.483***	0.471***	0.471***
2007–2009 Kendall correlation	0.956***	0.909***	0.776*	0.710***	0.641***	0.936***	0.878***	0.864***	0.243**	0.199*	0.184*
2007–2010 Pearson correlation	0.993***	0.993***	0.988***	0.960**	0.925***	0.986***	0.986***	0.980***	0.846***	0.748***	0.748***
2007–2010 Kendall correlation	0.949***	0.949***	0.924*	0.850*	0.780*	0.916***	0.916***	0.896***	0.613***	0.438***	0.448***
2007–2011 Pearson correlation	0.996***	0.996***	0.999***	0.999***	0.965***	0.991***	0.991***	0.986***	0.957***	0.886***	0.887***
2007–2011 Kendall correlation	0.966***	0.966***	0.922*	0.922*	0.852*	0.917***	0.917***	0.898***	0.854***	0.658***	0.668***
2007–2012 Pearson correlation	0.998***	0.998***	0.998***	0.998***	0.982***	0.998***	0.998***	0.994***	0.988***	0.935***	0.935***
2007–2012 Kendall correlation	0.981***	0.981***	0.977*	0.977*	0.903*	0.964***	0.964***	0.948***	0.922***	0.721***	0.740***
2007–2013 Pearson correlation	0.999***	0.999***	0.991***	0.991***	0.991***	0.997***	0.997***	0.997***	0.964***	0.964***	0.966***
2007–2013 Kendall correlation	0.987***	0.987***	0.926*	0.926*	0.926*	0.958***	0.958***	0.958***	0.778***	0.778***	0.793***
2007–2014 Pearson correlation	0.999***	0.999***	0.996***	0.996***	0.996***	0.998***	0.998***	0.998***	0.980***	0.980***	0.981***
2007–2014 Kendall correlation	0.992***	0.992***	0.941*	0.941*	0.941*	0.977***	0.977***	0.977***	0.829***	0.829***	0.848***
2007–2015 Pearson correlation	0.999***	0.999***	0.998***	0.998***	0.998***	0.999***	0.999***	0.999***	0.991***	0.991***	0.992***
2007–2015 Kendall correlation	0.998***	0.998***	0.960*	0.960*	0.960*	0.978***	0.978***	0.978***	0.890***	0.890***	0.901***
2007–2016 Pearson correlation	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.996***	0.996***	0.998***

Table 7 (continued)

	2007–2010	2007–2018	2007–2011	2007–2013	2007–2018	2007–2010	2007–2013	2007–2018	2007–2017	2007–2018
Kendall correlation		0.998***		0.987*				0.991***	0.911***	0.935***
2007–2017 Pearson correlation		0.999***		0.999***				0.999***		0.998***
Kendall correlation		1***		0.998*				0.997***		0.951***
	(a) Citations	(b) Citations per Paper	(c) h-Index	(d) J-Factor						

*Significant at 10% level
 **Significant at 5% level
 ***Significant at 1% level

The *J*-factor reveals another pattern here. The correlation coefficients for the *J*-factor primarily indicate a positive linear correlation. However, it is apparent that the concrete properties of the rank correlations for the *J*-factor diverge with regard to the different citation periods. Rank correlations of over 80% are achieved in business administration between the *J*-factors for 2013 and 2014, for biology between 2012 and 2013 and for medicine between 2014 and 2015. For business administration and medicine, *J*-factor rank correlations exceed 90% between the citation periods 2007 to 2016 and 2007 to 2017. For biology, however, rank correlations of 90% are already available between 2007 to 2015 and 2007 to 2016.

Implications

In this section, we answer our formulated research questions based on the results presented above. In order to do this, we will in each case begin by repeating the research question.

(R1) Does the informative content of citation indicators increase for longer citation periods?

As already explained in our literature overview, a distinction should be drawn between additional informative value from an increase in citations and thus the impact of a paper, as well as a change in the rankings. With regard to the general impact of a paper, our analyses show that the citations received increase sharply between the first four citation periods up to 2012. From then on, a moderate increase in citations runs up to a stagnation. This produces a larger population of received citations, which leads us to anticipate that generally speaking, this is accompanied by a greater informative value between a short and medium citation period length in terms of increased validity. Regarding the issue of the rankings, it has emerged that high rank correlations of more than 90% result between all the citation periods for the indicators “number of citations” in medicine, for biology one period later. For business administration 90% are reached for the citation period 2007–2012. This shows that, regarding the formation of a ranking, although the absolute numbers of citations increase over time, an additional analysis of the citation metric at a later point in the survey only provides redundant information and is not necessary. For this study, this means that considering citation periods from 2012 on is obsolete concerning the formation of a ranking.

(R2) Is there a citation period for which the long-term impact of citations can be determined and after which a longer period of analysis of the citation rate does not provide any additional informative value?

The above findings support this research question. It exists a citation period in which the trend of the citation rate can be detected early. Other extended periods do not necessarily provide any additional informative value. The position parameters illustrate that the citation numbers no longer increase sharply between the citation periods of 2013 and 2018. This allows us to conclude that the transition between the maturing phase and decline phase of the publications took place near 2013. From a ranking perspective, this period is earlier (when regarding the absolute number of citations). This result is mainly in line with the results in Glänzel (2008), who stated: “A 3-year citation window suffices at both the national and the institutional level if properly standardised and normalised citation

indicators are used.” Finally, in a sound bibliometric evaluation, the same rules of the game are applied to all units of assessment. However, as we will discuss later, the first part of this statement is not valid for the standardised and normalised citation indicator which we used here.

(R3) To what extent does the uneven distribution of citation numbers differ for the three disciplines under consideration and does it change over time?

In the course of the descriptive analysis of the citation numbers over time and for all three disciplines, our results confirm the current state of the art on research concerning citation distribution, since it shows a skewed distribution independent of the discipline and independent from time. This means that the results of all the disciplines are in line with Lotka’s Law: a small share of authors generate a large proportion of the total citations (Ruiz-Castillo and Costas 2018). However, differences in the characteristic of the citations unequal distribution have occurred between the disciplines. However, a trend appears to be emerging that more uniform distributions of citations occur as the number of publications and citations increases. One explanation could be that outliers in the form of, for example, highly cited publications, can distort the citation distribution. On the other hand, a larger number of publications leads to a greater diversity of data sets, which results in a balancing effect in the sum. However, this does not mean that a uniform distribution could be reached. This would only occur in the almost impossible case that all the citations were homogeneously distributed among the publications. Our investigations show that even if the time horizon of citation generation is extended, the citation distribution remains largely constant. However, Katz (2016) finds out that longer citation periods are associated with an increase in the skewness of the citation distribution. To state that means, whether a publication receives a high or low impact and thus high or low citation numbers is already crystallized in an ‘early stage’ of citation accumulation. It seems that publications that were already heavily cited at the beginning are also heavily cited in the maturing process up to the peak of citation, and publications that are cited infrequently receive relatively little attention in later periods. We should be aware of the fact that by citing a scientific publication, the author is making a decision about its relevance that can be reflected in performance indicators. It could not be explained otherwise that there is, e.g., a correlation between publications with intellectual relevance assessment and bibliometric indicators (Breuer et al. 2020). This is made all the more apparent by the fact that the underlying publication set was not changed at any time during this study. This means that all changes in distribution patterns that we can observe here are based solely on the perception of the underlying publications. The perception of a publication, whether it makes an essential contribution to a research topic or not, does not really change during the period of perception.

Consequently, a kind of citation pattern emerges for each publication that hardly changes over time. Only “sleeping beauties” change their patterns over time, these are however, as already noted, very rare (Glänzel et al. 2003), so they do not influence the general citation pattern.

(R4) Does the added informative value of the h-index remain limited in comparison with the absolute number of citations and the citations per publication when the citation period is extended?

The findings regarding the informative content of research question R1 can almost be transferred to the h-index. We can see similar tendencies, and there are no specific differences between the disciplines. While the rankings based on the number of citations are slightly

influenced by subject-specific citation behaviour, the rankings of the h-index hardly change and remain stable. This can be attributed to the construction of the h-index. Although it is clear that the fundamental absolute dimensions increase by less for the h-index, this is precisely what we expected. This is because the metric is designed in such a way that it does not uniformly increase with each additional citation. Therefore, a general increase in the added informative value can also be assumed for the h-index, while the ranking proves to be stable after a period of just 6 years for all three disciplines.

(R5) Does the informative value of the J -factor increase as the citation period grows longer?

Regarding the J -factor, in contrast with the other citation indicators, we have established differences in the properties between the citation periods. Only small correlations of between 20% and 60% can be registered between the first three citation periods between 2007 to 2009 and 2007 to 2011 for all three disciplines. The low correlations imply that the informative value of the J -factor diverges when the stated citation periods are used. As a consequence, considering extended citation periods provide additional informative value. This is due to the fact that numerous publications have not yet been cited in a relatively short citation period. If the citation periods are extended, the number of publications that have not been cited decreases, with the result that the J -factors for business administration and medicine for the 2016 period and biology for the 2015 period can be viewed as valid and informative. Thus, at least in relation to the ranking, the extension no longer provides any significant additional value.

(R6) Do the results concerning the length of the citation periods with regard to the effect of the research differ depending on the discipline being studied?

Over the course of the analyses of the position parameters and correlations, a nearly homogeneous pattern for all the disciplines being studied was noted. Specifically, a sharp increase in the citations is registered for the position parameters between the citation periods 2007 and 2009 up to 2007 and 2012, while only small increases result between 2007 to 2013 onwards. The distribution parameters show great differences. This can be attributed to the fact that the disciplines have different citation volumes, i.e. the more citations there are, the more the effect of outliers is reduced and the more uniform the distribution is. Because the citation distributions are slightly affected by the length of the citation periods, except for business administration, they can be ignored hereafter. Looking at the correlation coefficients shows that there are high correlations between the ten consecutive citation periods for all three disciplines and the difference in the degree of correlation is negligible. Based on the empirical findings, we can draw the conclusion that the length of the citation periods with regard to the effect of the research does not differ between the disciplines.

Limitations and outlook

With regard to collecting real data sets, it should be noted that only a portion of all the available publications can be included in the analysis. This can be attributed to the fact that coverage of the publications in the existing literature databases is not complete. The existing literature databases differ in the extent of the coverage and also show great differences for the various disciplines. According to a study by Craig et al. (2014), coverage in WoS is at 47% in economics, between 69% and 90% for biology, and at 84% for medicine.

According to Heinze et al. (2019), coverage of economics is at 53,8%, biology is at 83% and medicine is at 80,5%. Ultimately, there always needs to be a check for each discipline whether the section used is a sufficiently representative sample.

All WoS publications from the examined disciplines of the “article” or “review” document types have been included in this study. Other document types, particularly little-cited ones such as proceedings papers or letters, can cause fluctuations, especially in the J -factor, because relatively low absolute publication figures on the part of the benchmark could relatively quickly cause ratios well above the mean.

Finally, it should be noted that this study and its findings are based on a general approach. As a consequence, correlations are determined for complete rankings, based on the correlation and rank correlation analyses. However, it should be noted that major changes may still occur for a specific university in a particular discipline, even if the correlation coefficients are high.

Regarding further research, the extent to which our findings related to the J -factor are also valid for other normalizing citation indicators should also be studied. For this purpose, the analysis could be expanded to include another journal-normalized indicator and a field-normalized indicator. Looking at just one further journal-normalized indicator would not be sufficient, as it would not be possible to distinguish whether the findings are due simply to the method of normalization per se or are only caused by the method of journal-normalization. With regard to the internationalization of research, it would be of interest to compare the effects of research in Germany with other countries (Haustein and Tunger 2013).

Acknowledgements The project on which this paper is based was funded by the German Federal Ministry of Education and Research under the grant number 01PU17001. The responsibility of the content lies at the authors.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abramo, G., Cicero, T., & D’Angelo, C. A. (2011). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics*, 5(4), 659–667.
- Abramo, G., D’angelo, C. A., & Felici, G. (2019). Predicting publication long-term impact through a combination of early citations and journal impact factor. *Journal of Informetrics*, 13(1), 32–49.
- Adams, J. (2005). Early citation counts correlate with accumulated impact. *Scientometrics*, 63(3), 567–581.
- Albarran, P., Crespo, J. A., Ortuno, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, 88(2), 385–397.
- Albarran, P., & Ruiz-Castillo, J. (2011). References made and citations received by scientific articles. *Journal of the American Society for Information Science and Technology*, 62(1), 40–49.
- Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). h-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, 3(4), 273–289.

- Ball, R., Mittermaier, B., & Tunger, D. (2009). Creation of journal-based publication profiles of scientific institutions: A methodology for the interdisciplinary comparison of scientific research based on the *J*-factor. *Scientometrics*, *81*(2), 381–392.
- Bar-Ilan, J. (2008). Which h-index? A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, *74*(2), 257–271.
- Berghoff, S., Brandenburg, U., Carr, D., Hachmeister, C.-D., Leichsenring, H., & Ziegele, F. (2009). *Identifying the best: The CHE ExcellenceRanking for Natural Sciences, Economics, Political Science and Psychology in Europe*. Gütersloh: CHE Centre for Higher Education Development gGmbH.
- Bornmann, L., & Daniel, H. D. (2009). The state of h index research Is the h index the ideal way to measure research performance? *EMBO Reports*, *10*(1), 2–6.
- Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. D. (2011). A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics*, *5*(3), 346–359.
- Breuer, T., Schaer, P., & Tunger, D. (2020). Relations between relevance assessments, bibliometrics and altmetrics. In G. Cabanac, I. Frommholz, & P. Mayr (Eds.), *In Proceedings of the 10th international workshop on bibliometric-enhanced information retrieval* (pp. 101–112). <http://ceur-ws.org/Vol-2591/>. Accessed 28 Sept 2020.
- Clermont, M., & Dirksen, A. (2016). The measurement, evaluation, and publication of performance in higher education: An analysis of the research ranking of business schools in Germany from an accounting perspective. *Public Administration Quarterly*, *40*(2), 133–178.
- Clermont, M., Dirksen, A., Scheidt, B., & Tunger, D. (2017). Citation metrics as an additional indicator for evaluating research performance? An analysis of their correlations and validity. *Business Research*, *10*(2), 249–279.
- Costas, R., Van Leeuwen, T. N., & Van Raan, A. F. J. (2010). Is scientific literature subject to a ‘sell-by-date’? A general methodology to analyze the ‘durability’ of scientific documents. *Journal of the American Society for Information Science and Technology*, *61*(2), 329–339.
- Craig, I. D., Ferguson, L., & Finch, A. T. (2014). Journals ranking and impact factors: How the performance of journals is measured. In B. Cope & A. Phillips (Eds.), *The future of the academic journal* (2nd ed., pp. 259–298). Sawston: Chandos Publications.
- De Bellis, N. (2009). The Mathematical Foundations of Bibliometrie. In N. De Bellis (Ed.), *Bibliometrics and citation analysis: From the science citation index to cybermetrics* (pp. 75–140). Lanham, Maryland: The Scarecrow Press Inc.
- Dilger, A. (2010). Rankings von Zeitschriften und Personen in der BWL. *Zeitschrift für Management*, *5*(9), 91–102.
- Dorta-Gonzalez, P., & Dorta-Gonzalez, M. I. (2013). Impact maturity times and citation time windows: The 2-year maximum journal impact factor. *Journal of Informetrics*, *7*(3), 593–602.
- Dyckhoff, H., & Schmitz, C. (2007). Forschungsleistungsmessung mittels SSCI oder SCI-X? Internationale Sichtbarkeit und Wahrnehmung der Betriebswirtschaftslehre von 1990 bis 2004. *Die Betriebswirtschaft*, *67*(6), 640–664.
- El Aichouchi, A., & Gorry, P. (2018). Delayed recognition of Judah Folkman’s hypothesis on tumor angiogenesis: When a Prince awakens a Sleeping Beauty by self-citation. *Scientometrics*, *116*(1), 385–399.
- Fiala, D. (2014). Current index: A proposal for a dynamic rating system for researchers. *Journal of the Association for Information Science and Technology*, *65*(4), 850–855.
- Glänzel, W. (2008). Seven Myths in Bibliometrics: About facts and fiction in quantitative science studies. *Collnet Journal of Scientometrics and Information Management*, *2*(1), 9–17.
- Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, *53*(2), 171–193.
- Glänzel, W., Schlemmer, B., & Thijs, B. (2003). Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. *Scientometrics*, *58*(3), 571–586.
- Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on aging and reception processes of scientific literature. *Journal of Information Science*, *21*(1), 37–53.
- Glänzel, W., Thijs, B., & Schubert, A. (2008). Methodological aspects of a new generation of relational charts for comparative assessment of citation impact. *IV Seminario Internacional sobre estudios cuantitativos y cualitativos de la ciencia y la tecnología*, *56*(6), 373–379.
- Gonzalez, M. I. D., & Gonzalez, P. D. (2016). Do fixed citation windows match impact maturation rates of scientific journals? *Investigacion Bibliotecologica*, *30*(68), 73–89.
- Gross, P. L. K., & Gross, E. M. (1927). College libraries and chemical education. *Science*, *66*(1713), 385–389.

- Haustein, S., & Tunger, D. (2013). Sziento und bibliometrische Verfahren. In R. Kuhlen, W. Semar, & D. Straucht (Eds.), *Grundlagen der praktischen Information und Dokumentation: Handbuch zur Einführung in die Informationswissenschaft und -praxis* (pp. 479–492). Munich: De Gruyter.
- Heinze, T., Tunger, D., Fuchs, J. E., Jappe, A., & Eberhardt, P. (2019). *Research and teaching profiles of public universities in Germany: A mapping of selected fields*. Wuppertal: BUW.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572.
- Hirsch, J. E. (2007). Does the h index have predictive power? *Proceedings of the National Academy of Sciences of the United States of America*, *104*(49), 19193–19198.
- Honekopp, J., & Khan, J. (2012). Future publication success in science is better predicted by traditional measures than by the h index. *Scientometrics*, *90*(3), 843–853.
- Hu, Z. W., & Wu, Y. S. (2014). Regularity in the time-dependent distribution of the percentage of never-cited papers: An empirical pilot study based on the six journals. *Journal of Informetrics*, *8*(1), 136–146.
- Jensen, P., Rouquier, J. B., & Croissant, Y. (2009). Testing bibliometric indicators by their prediction of scientists promotions. *Scientometrics*, *78*(3), 467–479.
- Katz, J. S. (2016). What is a complex innovation system? *PLoS ONE*, *11*(6), e0156150.
- Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), 81–93.
- Ketzler, R., & Zimmermann, K. F. (2013). A citation-analysis of economic research institutes. *Scientometrics*, *95*(3), 1095–1112.
- Lehmann, S., Jackson, A. D., & Lautrup, B. E. (2006). Measures for measures. *Nature*, *444*(7122), 1003–1004.
- Leydesdorff, L., Zhou, P., & Bornmann, L. (2013). How can journal impact factors be normalized across fields of science? An assessment in terms of percentile ranks and fractional counts. *Journal of the American Society for Information Science and Technology*, *64*(1), 96–107.
- Li, Y. R., Radicchi, F., Castellano, C., & Ruiz-Castillo, J. (2013). Quantitative evaluation of alternative field normalization procedures. *Journal of Informetrics*, *7*(3), 746–755.
- Liu, X. L., Gai, S. S., Zhang, S. L., & Wang, P. (2015). An analysis of peer-reviewed scores and impact factors with different citation time windows: A case study of 28 ophthalmologic journals. *PLoS ONE*, *10*(8), 11.
- Lotka, A. J. (1926). The frequency-distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, *16*(12), 317–323.
- Malesios, C. C., & Psarakis, S. (2014). Comparison of the h-index for different fields of research using bootstrap methodology. *Quality & Quantity*, *48*(1), 521–545.
- Nicolaisen, J., & Frandsen, T. F. (2019). Zero impact: A large-scale study of uncitedness. *Scientometrics*, *119*(2), 1227–1254.
- Pan, R. K., & Fortunato, S. (2014). Author Impact Factor: Tracking the dynamics of individual scientific impact. *Scientific Reports*, *4*, 4880.
- Price, D. J. D. (1965). Networks of scientific papers. *Science*, *149*(3683), 510–515.
- Pyatt, G. (1976). On the interpretation and disaggregation of gini coefficients. *The Economic Journal*, *86*(342), 243–255.
- Radicchi, F., & Castellano, C. (2012). A reverse engineering approach to the suppression of citation biases reveals universal properties of citation distributions. *PLoS ONE*, *7*(3), e33833.
- Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17268–17272.
- Research Evaluation and Policy Project. (2005). *Qualitative indicators for research assessment: A literature overview*. Canberra: The Australian National University.
- Rovira-Esteva, S., Aixela, J. F., & Olalla-Soler, C. (2019). Citation patterns in translation studies: A format-dependent bibliometric analysis. *Translation & Interpreting—The International Journal of Translation and Interpreting*, *11*(1), 147–171.
- Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, *8*(4), 917–934.
- Ruiz-Castillo, J., & Costas, R. (2018). Individual and field citation distributions in 29 broad scientific fields. *Journal of Informetrics*, *12*(3), 868–892.
- Saam, N. J., & Reiter, L. (1999). Lotka's law reconsidered: The evolution of publication and citation distributions in scientific fields. *Scientometrics*, *44*(2), 135–155.
- Schendera, C. F. G. (2004). *Datenmanagement und Datenanalyse mit dem SAS-System*. München: Oldenbourg Wissenschaftsverlag.

- Schreiber, M. (2008). An empirical investigation of the g-index for 26 physicists in comparison with the h-Index, the A-index, and the R-index. *Journal of the American Society for Information Science and Technology*, 59(9), 1513–1522.
- Schreiber, M. (2015). Restricting the h-index to a publication and citation time window: A case study of a timed Hirsch index. *Journal of Informetrics*, 9(1), 150–155.
- Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628–638.
- Sharma, B., Boet, S., Grantcharov, T., Shin, E., Barrowman, N. J., & Bould, M. D. (2013). The h-index outperforms other bibliometrics in the assessment of research performance in general surgery: A province-wide study. *Surgery*, 153(4), 493–501.
- Song, Y., Situ, F. L., Zhu, H. J., & Lei, J. Z. (2018). To be the Prince to wake up Sleeping Beauty: The rediscovery of the delayed recognition studies. *Scientometrics*, 117(3479), 9–24.
- Spearman, C. (1904). “General intelligence” objectively determined and measured. *American Journal of Psychology*, 15(2), 201–292.
- Teixeira, A. A. C., Vieira, P. C., & Abreu, A. P. (2017). Sleeping Beauties and their princes in innovation studies. *Scientometrics*, 110(2), 541–580.
- Van Dalen, H. P., & Henkens, K. (2005). Signals in science—On the importance of signaling in gaining attention in science. *Scientometrics*, 64(2), 209–233.
- Van Raan, A. F. J. (2004a). Sleeping Beauties in science. *Scientometrics*, 59(3), 467–472.
- Van Raan, A. F. J. (2004b). Measuring Science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 19–50). Dordrecht: Springer.
- Van Raan, A. F. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, 67(3), 491–502.
- Van Raan, A. F. J., & Winnink, J. (2018). Do younger Sleeping Beauties prefer a technological prince? *Scientometrics*, 114(3), 701–717.
- Waltman, L., Van Eck, N. J., Van Leeuwen, T. N., Visser, M. S., & Van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47.
- Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, 94(3), 851–872.
- Wang, M. Y., Jiao, S. J., Chai, K. H., & Chen, G. S. (2019). Building journal’s long-term impact: Using indicators detected from the sustained active articles. *Scientometrics*, 121(1), 261–283.
- Xu, W. C., Hou, Y. H., Hung, Y. S., & Zou, Y. X. (2013). A comparative analysis of Spearman’s rho and Kendall’s tau in normal and contaminated normal models. *Signal Processing*, 93(1), 261–276.