

Previous Year's Cite Score Strongly Predicts the Next Year's Score: Ten Years of Evidence for the Top 400 Scopus-indexed Journals of 2021

Atul Kumar¹, Jaiprakash M Paliwal², Vinaydeep Brar³, Mahesh Singh⁴, Prashant R Tambe Patil⁵, Shirish S Raibagkar^{6,*}

¹Dr. D. Y. Patil B-School, Pune, Maharashtra, INDIA.

²Symbiosis Center For Skill Development and Symbiosis International (Deemed University), Nagpur, Maharashtra, INDIA.

³S.N.G. Institute of Management and Research, Pune, Maharashtra, INDIA.

⁴Kebri Dehar University, Kebri Dehar, ETHIOPIA.

⁵Akole Taluka Education Society's Technical Campus, Akole, Ahmednagar, Gujarat, INDIA.

⁶Institute of Cost Accountants of India, Kolkata, West Bengal, INDIA.

ABSTRACT

Over the last few years, CiteScore has emerged as a popular metric to measure the performance of Journals. In this paper, we analyze CiteScores of the top 400 Scopus-indexed journals of 2021 for years from 2011 to 2021. Some interesting observations emerged from the analysis. The average CiteScore of the top 400 journals doubled from 16.48 in 2011 to 31.83 in 2021. At the same time, the standard deviation has almost trebled from 13.53 in 2011 to 38.18 in 2021. The CiteScores also show sizable increases for skewness and kurtosis, implying major variations in the CiteScores of the journals for a year. Importantly, the previous year's CiteScores strongly predict the next year's scores. This has been observed consistently for the last ten years. The average Pearson correlation coefficient between the preceding and succeeding years' CiteScores for the ten years is 0.98. We also show that it is easily possible for even people with just basic knowledge of computers to forecast the CiteScore. Researchers can predict CiteScores based on the past year's CiteScores and decide better about publishing their current research in a journal with an idea about its likely CiteScore. Such a forecast can be useful to publishers, editorial staff, indexing services, university authorities, and funding agencies.

Keywords: CiteScore, Scopus, Journals, Research, Pearson's correlation coefficient, Forecast.

Correspondence:

Shirish S Raibagkar

Institute of Cost Accountants of India,
Kolkata, West Bengal, INDIA.

Email: srrsr696@gmail.com

ORCID: 0000-0002-8017-7740

Received: 11-02-2023;

Revised: 18-06-2023;

Accepted: 31-07-2023.

INTRODUCTION

Of late, CiteScore, a journal performance measure launched by Elsevier (Scopus), has emerged as a popular metric amongst researchers. It has been assigned to more journals than Clarivate Analytics' Journal Impact Factor (JIF), including journals that are indexed in Scopus but do not carry a JIF (Teixeira da Silva, 2020).^[1] Launched as an alternative to the JIF, in December 2016, CiteScore provided a much broader coverage for evaluating journals thanks to the Scopus database. While JIF evaluated around 11,000 journals, CiteScore was assigned to 22,000 journals (Van Noorden, 2016).^[2] As of the date of writing this article (23 November 2022), the Scopus database includes CiteScores for more than 44,000 journals (Scopus, 2022).^[3] The CiteScore methodology has changed since its launch. The 2021

methodology considers a four-year publication window covering citations and publications for 2018-2021. This methodology has been applied retroactively for earlier year CiteScores as well. Scopus claims that the new methodology ensures a stable, robust, and comprehensive metric indicating the impact factor of the journals (Scopus, 2022).^[3]

An example is given to understand the calculation of CiteScore quickly. The journal Ca-A Cancer Journal for Clinicians has a 2021 CiteScore of 716.2. It has been worked out based on the number of citations for the journal's documents during the period 2018-21 and the total number of documents published by the journal during the same period. The citations for publications of the Ca-A Cancer Journal for Clinicians for the period 2018-21 are 76,632, whereas the total number of documents published by the journal during the same period is 107. Thus, the CiteScore is 716.2 (76,632/107). The same journal had a CiteScore of 463.2 in 2020, where the citations for 2017-20 were 50,948, and the total documents published during 2017-20 were 110. (50,948/110 = 463.2) (Scopus, 2022).^[3]



DOI: 10.5530/jscires.12.2.020

Copyright Information :

Copyright Author (s) 2023 Distributed under
Creative Commons CC-BY 4.0

Publishing Partner : EManuscript Tech. [www.emanuscript.in]

However, the new methodology has drawn some criticism. Because the formula used to calculate the new CiteScore contains more items representing early citations, it is biased in favor of journals with a high proportion of early citations within four years after publication (Fang, 2021).^[4] Notwithstanding some such criticisms, the CiteScore has been hailed for its transparency, comprehensiveness, up-to-datedness, and free access (James *et al.*, 2018).^[5] Authors believe that the CiteScore journal ranking metrics data can strengthen strategic initiatives for librarians to assist faculties and university libraries with collective decision-making processes (Torres, 2022).^[6]

In this article, we demonstrate, with the help of examples, how the projection of the CiteScore for a Journal for a full year is possible based on its past CiteScores. We have consciously used a very simple methodology so that it can be used by even those who might not be better versed with sophisticated techniques like machine learning and others. We use a simple MS Excel formula that can be used even by a person with only basic knowledge of using an electronic spreadsheet. This article has two objectives:

To establish based on sizable empirical evidence that prediction of a CiteScore is possible based on past year's CiteScores, and,

To share a simple methodology that can be used by people who are not experts in computers to predict the CiteScore.

Our article makes a realistic assumption about the computer proficiency of users of CiteScores. They cannot be expected to be fluent in sophisticated techniques like machine learning and others. Rather an ordinary researcher or a librarian usually would be expected to have only basic knowledge like dealing with simple spreadsheets and formulae. We aim to empower the most common and ordinary users of CiteScore with a very simple formula that will enable them to forecast the CiteScore based on past CiteScores.

Croft and Sack (2022)^[7] have enumerated users and situations where a projected CiteScore can be used.

Journal editorial boards are directly responsible for ensuring consistent quality in the journals they manage. Decisions regarding the journal's management and the editorial board's composition are made with explicit consideration of the journal's future performance.

Indexing services must ensure that only journals of sufficient quality are accepted for indexing. Therefore, knowledge of the future performance of journals helps in deciding whether new journals should be accepted and whether currently indexed journals should be removed.

Publishers need to monitor the performance of journals (both their own and those of other publishers) to inform strategic decision-making. Knowledge of the trajectory of a journal's performance and its projection into the future can assist in

decision-making regarding acquiring existing journals and launching new journals.

Some grant agencies maintain lists of journals sorted by strength. These lists are regularly updated; early information on which journals to move up or down can be useful to grant agencies and applicants.

For academic institutions and countries, the quality of relevant journals can be an indicator in evaluating their scientific output. Information about journals' expected performance can help make budget allocation decisions.

For research groups and authors, information about the journal's quality is often considered when deciding where to submit their work. Given the lengthy duration of most review and publication processes, the future performance of the journal is highly relevant to this decision.

For librarians building journal collections, a journal's performance can help assess its selection value. Information on future performance helps select journals leading to a sustainable collection (Croft and Sack, 2022).^[7]

Thus, there are major and multiple uses of a projected CiteScore. What is required is the knowledge of producing the forecast in the easiest of ways possible after establishing that such forecasting is possible in the first place and can be expected to produce reasonably reliable results.

The remainder of the paper is divided into four sections – literature review; methods; data analysis, results, discussion; and conclusion.

LITERATURE REVIEW

Literature related to CiteScore, including its comparison with other metrics, was reviewed.

Based on the development of several years, Elsevier in December 2016 launched CiteScore, a transparent, comprehensive, updated, and free-access metric to evaluate the impact factors of journals. Transparency, comprehensiveness, being updated, and free access is among the desirable characteristics of a journal citation metric (James *et al.*, 2018).^[5]

The JIF has faced criticism from the research community for the potential for manipulation (Martin, 2016;^[8] Matthews, 2015;^[9] Vanclay, 2012^[10]) and the absence of transparency in the metric (Archambault and Larivière, 2009).^[11]

The year Y CiteScore calculation formula contains more entries representing citations received by eligible documents (E.D.s) published by the journal in Y – 3. Therefore, the new year Y CiteScore is more influenced by the impact of E.D.s published in Y – 3 than by the impact of E.D.s published in Y – 2 than in Y – 1 and Y. Moreover, the impact of E.D.s published in Y – 3

more strongly affects CiteScore Y than the traditional JIF Y, with the citation window extended to cover $Y - 3$ to Y (Fang, 2021).^[14]

CiteScores have been researched for a specific publisher group. For example, a study explored the CiteScores of 180 Hindawi group journals (Okagbue *et al.*, 2018).^[12] Similarly, studies have been conducted researching CiteScores for journals belonging to specific domains. A study has explored CiteScores of 105 Computer Science, Theory, and Methods journals (Okagbue *et al.*, 2020).^[13] A correlation between the CiteScore and JIF of top-notch libraries and information sciences journals has been studied. The study found a strong positive correlation between the CiteScore and the JIF ($r=0.787$; $r_s=0.828$) (Okagbue and Teixeira da Silva, 2020).^[14]

A change in the CiteScore methodology was announced in June 2020. The effect of these changes was examined in respect of 40 journals chosen from the middle five and top five ranked journals (CiteScore) in Social Sciences, Medicine, Material Science, and General Physics and Astronomy. A comparison was made with their impact factors. It was found that the new methodology was less prone to the impact of the proportion of editorial material in a journal but tended to favor journals that had quick citations (Trapp, 2020).^[15] More research is available comparing Clarivate Journal Citation Reports and Scopus CiteScore. The similarities and differences in the methods have been highlighted (Salisbury, 2020).^[16] Studies evaluating the performance of a single journal (Ambiente and Água Journal) based on CiteScores are seen. Improved CiteScores have been taken as a measure of the success of the strategy adopted by the journal to increase its visibility within the international scientific community (Dias, 2021).^[17]

CiteScores have been used to compare the impact of open-access and subscription journals. CiteScores of 2542 OA sources and 15,040 SB sources indexed in Scopus from 2014 to 2016 were presented and analyzed based on five inclusion criteria (Atayero *et al.*, 2018).^[18] From the analysis of 133 journals CiteScores for 2014 to 2016, it was found that journals from medicine; arts and humanities; administration and accounting; business; immunology and microbiology; and economics, econometrics, and finance have the highest impact (Henao-Rodríguez, 2019).^[19]

Researchers have demanded more clarity on “N/A” values instead of CiteScores that were assigned to journals from “Library and Information Science” (Krauskopf, 2020).^[20]

There are differences in the outcomes due to differences in the classification of journals in CiteScore and JIF methodology. While Pharmacy has been classified as an independent subject area in CiteScore, it has been merged with Pharmacology in the Journal Citation Reports. This merger leads to a need for more clarity compared to more true CiteScore results (Fernandez-Llimos, 2018).^[21]

CiteScore should only be used to evaluate the citation impact of titles in the same field. The CiteScore Percentile compares the citation impact of titles in different fields. The basket of metrics supports valuable and responsible input into decision-making (Colledge, 2017).^[22]

Elsevier (Scopus) CiteScore is an increasingly popular journal-based metric (JBM) that is rapidly gaining popularity over the once decade-dominating JBM, Clarivate Analytics’ Journal Impact Factor (JIF). CiteScore, which is currently assigned to more than 41,000 journals or other sources indexed by Scopus, faces a risk that does not seem to have been discussed yet, namely that it could be “hijacked” to create a copycat or misleading metric. JIF already famously suffers from this phenomenon in “predatory” open-access publishing, but academic publishing predators are constantly looking for ways to expand their “prey” base, i.e., authors (Teixeira da Silva, 2021).^[23]

Findings for the top six occupational therapy journals in English about JCR IF, EFS, AIS, SNIP, Citescore, SJR, and SJR IIF scores suggest that a range of available bibliometric indicators should be used to obtain a more comprehensive assessment of journal and article rankings rather than the singular use of IF scores currently and frequently found in many jurisdictions (Brown and Gutman, 2018).^[24]

The CiteScore method, as implemented to evaluate the quality of computer conferences, is highly effective as a benchmark for evaluating and comparing publication sites in computer science. However, Scopus needs to improve several indexing practices as the database, and the CiteScore method has become a standard tool for assessing conference quality (Meho, 2019).^[25]

A conference rating system could assess the quality of major conferences in many disciplines. One of the well-known evaluation metrics is CiteScore from Scopus. For example, computer science could only cover about 180 out of thousands of conferences across all industry categories, which is very limited in practice (Wahakit, 2021).^[26]

A review of 150 second-language journal articles revealed several prevalent statistical violations, including incomplete reporting of reliability, validity, nonsignificant results, effect sizes, and assumption checks, as well as concluding descriptive statistics and failure to correct for multiple comparisons. Scopus citation analysis metrics and SSCI indexes of the journal were predictors of journal statistical quality. No clear evidence was obtained favoring the newly introduced CiteScore over SNIP or SJR (Al-Hoorie and Vitta, 2019).^[27]

In the last decade, several journal editors have decided to publish alternative bibliometric indices parallel to the impact factor (IF): Scimago Journal Rank (SJR), Source Normalized Impact per Paper (SNIP), Eigenfactor Score (ES) and CiteScore; however, little is known about the correlations between them.

Findings support the hypothesis that IF does not show the best correlation among other metrics. Radiologists, interventional radiologists, or nuclear medicine physicians should clearly understand the relationships between journal bibliometrics for their decision-making during the manuscript submission phase (Villaseñor-Almaraz *et al.*, 2019).^[28]

Using the bibliometric metrics of two-year and five-year Journal Impact Factors, the H-index, and the newly revised CiteScore, the study examined the relationships between these metrics in a bibliometric study of forty-four representative family studies journals. Citation data was drawn from Journal Citation Reports, Scopus, and Google Scholar. Correlation analysis found strong positive relationships on the metrics. Despite strong correlations, inconsistencies in journal rankings were found (Liu, 2021).^[29]

O.A. journals accounted for approximately 17 percent of the total journals indexed by Scopus in 2015. The results revealed an uneven distribution of O.A. journals across disciplines, ranging from 5.5 to 28.7%. A study of journal quality as measured by CiteScore, SJR SNIP leads us to find that in all areas of research, except health professions and nursing, non-OA journals achieve statistically significantly higher average quality than O.A. journals (Erfanmanesh, 2017).^[30]

CiteScore, a Scopus/Elsevier open journal metric, is an attractive alternative to the Clarivate Analytics impact factor. In mid-2020, the equation used to calculate CiteScore changed to reflect a four-year data window from the previous three-year data set. By extrapolating CiteScore data from Scopus for the 1000 top-ranked journals, the authors wanted to appreciate the evolution of CiteScore over time. It was found that, on average, the CiteScore increased each year steadily between 2015 and 2019, from 13.877 to 16.536. This generally reflects a greater number of citations per publication over time, so a steady increase in citations. Academics should recognize this rise for a higher level of quality (Okagbue *et al.*, 2021).^[31]

A journal's CiteScore is positively correlated with the following variables or parameters: coverage of PubMed, Web of Science, and EMBASE ($p < 0.001$), articles in English ($p < 0.001$), age of the journal ($p = 0.001$), publication of review articles ($p = 0.23$), H-Index ($p < 0.001$) and Scimago Journal Rank ($p < 0.001$). Coverage of the journal in international databases, especially in PubMed, Web of Science, and EMBASE, is essential to increase its visibility. Publishing review articles that tend to be cited more often because they serve as comprehensive sources of information can increase a journal's CiteScore. Also, publishing more articles in English contributes to the number of journal article citations (Zolfaghari *et al.*, 2022).^[32]

CiteScore is a better way to measure the citation impact of sources such as journals. CiteScore is a journal metric product from Elsevier that uses citation data from the Scopus database to rank journals. CiteScore metrics are comprehensive, up-to-date,

and free metrics for resource titles in Scopus. In addition to the Impact factor, CiteScore is increasingly important in evaluating metrics for all journals (Rajkumar *et al.*, 2018).^[33]

Journal impact factor and CiteScore are known to be positively correlated with journal percentile, but the use of the latter to predict the former has yet to be debated, especially for journals in a subject-specific classification based on the Science Network. Significant positive correlations were obtained between the impact factor and CiteScore of journals (Okagbue *et al.*, 2019).^[34]

Until now, the Journal Impact Factor (JIF), owned by Thomson Reuters (now Clarivate Analytics), has been the dominant metric in scholarly publishing. Hated or loved, JIF has dominated academic publishing for over six decades. However, the rise of non-scholarly journals, academic corruption, and fraud have also led to a parallel universe of competing metrics, some of which may be predatory, misleading, or fraudulent, while others may be valid. On December 8, 2016, Elsevier B.V. launched a direct competitor metric to JIF, CiteScore (C.S.) (Teixeira da Silva and Memon, 2017).^[35]

The best indicator that can be used with IF is CiteScore. To measure the scientific quality of LIS journals, all stakeholders should consider correlations between different indicators. Furthermore, they can rely on CiteScore as an adequate alternative to IF (Ali, 2021).^[36]

Researchers have suggested a modified version of CiteScore to factor in self-citation impact (Okagbue *et al.*, 2019).^[37]

The impact of Open Access on the journal CiteScores was studied. The overall effect was positive but not uniform across different types of journals. Specifically, two types of heterogeneous treatment effects were examined: (1) differential treatment effect among journals grouped by academic field, publisher, and level; and (2) differential effects of open access as a function of treatment propensity (Li *et al.*, 2018).^[38]

Quite a few studies have looked into different dimensions of the CiteScore in isolation or in comparison with other metrics. However, other than a study by Crodt and Sack (2022),^[7] we could not find any research into predicting the CiteScore. As stated in the introduction, a projected CiteScore can be useful to researchers, publishers, editorial staff, indexing services, university authorities, and funding agencies. Given the value of the prediction and almost non-existent research in this area, we set up two research questions for this study:

RQ1: Does the previous year's CiteScore predict the next year's score?

RQ2: How can a common user of CiteScore with little computer expertise predict CiteScore?

METHODOLOGY

The number of journals indexed in Scopus as of 23rd November 2022 was 44,034. We aimed to establish that the previous year's CiteScore predicts the next year's score based on sizable sample size. Reference to standard sample size tables like Krejcie and Morgan (1970)^[39] returned a sample size of 381 for a population of 44,034 at a 95 percent confidence level and a 5 percent confidence interval. The sample size was rounded off to 400. The top 400 Scopus-indexed journals based on their 2021 CiteScore ranks were selected for the study. It was expected that the CiteScores for the top 2021 400 journals would also be available for past years up to 2011. A dataset of the top-ranked 400 Scopus-indexed journals was compiled based on the 2021 CiteScores. Excel lists of the top 1000 journals were extracted from the Scopus database for each of the past ten years, 2011 to 2020, expecting that one would be able to find all or maximum out of the 2021 top 400 journals in the previous year list so that the 2021 dataset can be extended to years up to 2011. However, every year CiteScores of some of the journals from the 2021 400 list were missing. The number of 2021 400 journals whose score was missing for the years 2020 onwards is shown in Table 1.

Thus, for 2020, CiteScores for 14 journals from the top 2021 400 lists were unavailable from the list of 1000 top CiteScores of 2020. An easy way out of the problem of missing scores was to use the 2021 CiteScore for these missing 14 journals in 2020, assuming that the previous year's scores would be the same as that of the current year. However, this assumption needed to be more logical as there was a clear downward trend of the CiteScores while moving backward from 2021 to 2011. Hence, taking the succeeding year's score for the previous year's CiteScore for the missing journals, for the time being, a year average CiteScore for each year was worked out. Based on the average scores, a factor was calculated for the past ten years from 2020 to 2011 by dividing the previous year's average CiteScore by the succeeding year's average CiteScore. These factors are shown in Table 2.

Thus, the average CiteScore for the 400 2021 top journals for 2020 was 85 percent of the 2021 average score. Similarly, the average CiteScore for the same 400 journals in 2019 was 91 percent of the 2020 average score. The not available CiteScore values were calculated for the respective years using these factors. For instance, the CiteScore for the journal EnergyChem for 2020 was unavailable. So, the same was worked out by multiplying the 2021 score of 33.4 with the 2020 factor of 0.85 and was taken as 28.3 for 2020.

In the same way, the journal EnergyChem's score was also not available for 2019. In this case, the 2020 score of 28.3 was multiplied by the 2019 factor of 0.91 to arrive at the 2019 score of 25.7. Thus, it was ensured that the unavailable scores were assumed on a logical basis of decreasing trend in the scores moving backward from 2021 to 2011. As explained earlier, the final calculations were made on the data that consisted of the duly-adjusted, not available scores. To the extent of adjustments made in the scores of the not-available CiteScores, the accuracy of the analysis was affected. However, the analysis was more of a directional nature, and overall results mattered to establish the predictability of the succeeding years' CiteScore based on the previous year's CiteScore.

The final dataset of the CiteScores including lists up to 2011 has been deposited with a repository and can be accessed from <https://www.openicpsr.org/openicpsr/project/183201/version/V1/view>.

Regression analysis was run based on the final data for pairs of years. The first regression analysis took 2021 CiteScores as the dependent variable and the 2020 CiteScores as the independent variable. The second regression analysis took 2020 CiteScores as the dependent variable and the 2019 CiteScores as the independent variable. The tenth regression analysis took the 2012 CiteScores as the dependent variable and the 2011 CiteScores as the independent variable. Thus, reports of such ten regression analyses were generated. The hypothesis put to the test was as under:

Table 1: Missing CiteScores in our analysis of top 400 journals for past years.

Year	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
No. of N/A 2021 journals CiteScores	14	34	21	66	84	104	112	118	128	146

Source: Authors compilation.
N/A = Not Applicable.

Table 2: Factors for discounting the not available CiteScores.

Year	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
Average score of previous year divided by average score of succeeding year	0.85	0.91	0.93	0.97	0.95	0.96	0.96	0.98	0.96	0.95

Source: Authors compilation.

Table 3: CiteScores of 11 years for the top 20 journals.

Source title	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
Ca-A Cancer Journal for Clinicians	716.2	463.2	435.4	387.2	290.8	237.4	189.1	210.9	196.9	157.3	150.5
Nature Reviews Molecular Cell Biology	140.9	99.7	73.4	62.6	76.0	61.7	64.3	58.9	48.9	42.2	36.0
The Lancet	115.3	91.5	73.4	64.9	57.9	51.4	55.2	48.2	49.4	45.8	45.0
New England Journal of Medicine	110.5	80.6	66.1	73.1	67.2	61.1	57.4	57.3	58.0	60.2	56.0
Reviews of Modern Physics	102.0	86.5	75.8	67.7	63.2	62.3	59.8	76.8	80.7	90.2	76.6
Chemical Reviews	98.8	96.9	100.5	94.5	85.8	66.6	72.5	67.1	78.6	65.8	56.3
Nature Reviews Materials	96.7	115.7	123.7	74.4	34.2	32.4	31.2	29.9	29.2	28.1	26.8
Nature Medicine	91.9	62.4	45.9	48.3	47.5	41.2	38.5	31.2	26.1	27.2	29.0
Living Reviews in Relativity	84.8	67.4	54.1	33.6	65.7	77.1	49.9	36.7	34.9	25.3	26.3
Nature Energy	78.0	68.7	71.2	43.7	28.4	26.9	25.9	24.8	24.3	23.3	22.2
Nature Reviews Cancer	77.1	78.3	70.4	70.3	64.7	61.9	66.6	64.7	47.1	40.0	31.3
Cell	77.0	63.4	58.7	56.2	54.9	53.5	54.6	53.9	52.4	55.5	56.8
Chemical Society Reviews	75.9	72.4	67.1	70.7	72.7	66.4	61.6	51.9	42.8	43.9	38.1
Progress in Materials Science	70.9	61.7	47.1	39.6	45.6	44.2	54.5	47.0	37.8	33.7	29.9
Nature Reviews Immunology	70.8	53.9	62.9	68.7	69.9	72.2	62.6	61.8	47.4	35.9	31.6
Nature	70.2	56.9	51.0	55.7	53.7	49.2	51.6	49.9	50.9	51.0	53.1
Nature Reviews Genetics	69.7	62.4	73.5	67.3	76.6	68.5	71.3	68.5	58.2	44.6	32.8
IEEE Communications Surveys and Tutorials	69.4	62.1	52.6	45.8	40.4	36.2	25.1	17.6	14.1	20.6	15.6
Physiological Reviews	68.2	48.9	36.1	37.1	49.5	54.0	48.2	55.3	54.8	49.2	50.8
Nature Reviews Drug Discovery	65.9	48.8	35.5	40.1	47.9	46.3	45.6	39.3	38.1	32.3	28.1

Source: Scopus database and authors compilation.

Ho: Previous year's CiteScore does not predict the succeeding years' CiteScore.

Ha: Previous year's CiteScore predicts the succeeding years' CiteScore.

Results are presented and discussed in the next section.

Data analysis, Results, and Discussion

A glimpse of the compilation of the dataset is shown in Table 3, which contains CiteScores of 11 years for the top 20 journals.

Table 4 gives the descriptive statistics of the 400 2021 top journals for 11 years.

All four measures show an increase from 2011 to 2021. The mean (average) CiteScore has almost doubled and has risen from 16.48 in 2011 to 31.83 in 2021. At the same time, the standard deviation has almost trebled and has risen from 13.53 in 2011 to 38.18 in 2021. Similarly, the skewness and kurtosis coefficients have shown a significant increase indicating that the distribution is becoming more and more asymmetrical.

If we observe Table 3, it can be seen that for all the years, the number one journal in the ranking, Ca-A Cancer Journal for Clinicians, has abnormally high CiteScores than the other

journals. So, we reworked the descriptive statistics for the dataset removing CiteScores for Ca-A Cancer Journal for Clinicians. The results of the revised work are given in Table 5.

When we compare results from Table 5 (399 journals – excluding Ca-A Cancer Journal for Clinicians) with Table 4 (400 journals – including Ca-A Cancer Journal for Clinicians), we do not see a major change in the average CiteScore. However, there is a significant change in the standard deviation, skewness, and kurtosis measures. While the standard deviation for the distribution of 400 journals for 2021 was 38.18, it dropped to 16.79 for the same year, 2021, excluding the Ca-A Cancer Journal for Clinicians. Similarly, while the skewness coefficient 2021 for the distribution of 400 journals for 2021 was 14.70, it dropped drastically to 2.65 for the same year 2021 with the exclusion of Ca-A Cancer Journal for Clinicians. The same is the case with the kurtosis coefficient, which was 259.93 for 2021, and the distribution of 400 journals in 2021. It also drops drastically to 9.42 for the same year, 2021, with the exclusion of the Ca-A Cancer Journal for Clinicians.

In Table 6, we present the summary of the ten regression runs, taking the succeeding year as the dependent variable and the preceding year as the independent variable.

Table 4: Descriptive statistics of the sample of top 400 journals for 11 years.

Measure	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
Average	31.83	26.88	24.31	22.64	21.91	20.61	19.75	18.82	18.31	17.46	16.48
SD	38.18	26.33	24.96	22.82	19.35	17.06	15.60	15.87	15.31	14.23	13.53
Skewness	14.70	11.85	11.69	10.65	7.38	5.91	4.21	5.36	4.95	3.80	3.72
Kurtosis	259.93	189.87	185.41	163.73	93.48	65.40	35.31	54.41	47.18	26.52	26.10

Source: Authors compilation.
SD = Standard Deviation.

Table 5: Descriptive statistics of the sample of top 399 journals for 11 years.

Measure	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
Average	30.11	25.79	23.28	21.73	21.24	20.07	19.32	18.33	17.86	17.11	16.15
SD	16.79	14.68	14.11	13.68	13.91	13.17	13.11	12.63	12.43	12.40	11.76
Skewness	2.65	2.49	2.66	1.96	1.84	1.75	1.77	1.89	1.89	2.22	2.03
Kurtosis	9.42	8.20	10.28	4.81	3.76	3.23	3.30	4.20	4.31	7.34	5.35

Source: Authors compilation.
SD = Standard Deviation.

Table 6: Summary of ten regression results (n=400).

Years	r	R ²	Equation of the model	F	P
2011-12	0.975	0.95	2012 = 0.56+1.03*2011	7563.41	<0.0001
2012-13	0.969	0.939	2013 = 0.11+1.04*2012	6113.434	<0.0001
2013-14	0.98	0.961	2014 = 0.22+1.02*2013	9790.514	<0.0001
2014-15	0.98	0.961	2015 = 1.61+0.96*2014	9723.034	<0.0001
2015-16	0.973	0.947	2016 = -0.40+1.06*2015	7147.663	<0.0001
2016-17	0.981	0.963	2017 = -1.04+1.11*2016	10451.182	<0.0001
2017-18	0.966	0.9334	2018 = -2.32+1.14*2017	5597.228	<0.0001
2018-19	0.974	0.949	2019 = 0.17+1.07*2018	7433.909	<0.0001
2019-20	0.984	0.968	2020 = 1.65+1.04*2019	12152.051	<0.0001
2020-21	0.978	0.956	2021 = -6.29+1.42*2020	8677.323	<0.0001

Source: Authors regression calculations.
r = Pearson's correlation coefficient, R² = Measure of variability.
F = F-statistic, p = Probability.

In the first case, 2012 CiteScores were taken as the dependent variable, whereas 2011 CiteScores were taken as the independent variable. Pearson's correlation coefficient r is 0.975. R² is 95%. The regression model equation is 2012 = 0.56+1.03*2011, and the results are statistically significant (F=7563.41; p<0.0001). Similar results are observed for the rest of the year pairs with very high r and R² with p-values of <0.0001. Based on these results, we reject the null hypothesis that the previous year's CiteScore does not predict the succeeding years' CiteScore; in favor of the alternate, the previous year's CiteScore predicts the succeeding years' CiteScore.

DISCUSSION

The average CiteScore for the sample of the top 2021 400 journals shows a consistent rise from 2011 to 2021. There is a clear upward trend in the scores for the period under consideration. This means

that, by and large, the documents are getting more citations every year, which speaks of an increase in the quality of the published documents. Or in other words, there is an increasing impact of the published documents, as indicated by the rising CiteScores. However, there are journals like the Ca-A Cancer Journal for Clinicians, which has abnormally high CiteScores than the other journals, making the overall data quite asymmetrical. For analysis purposes, it is better to exclude such cases (outliers) to have a normalized scenario. Our attempt on similar lines shows that with the exclusion of the Ca-A Cancer Journal for Clinicians, there is a significant variation in the standard deviation, skewness, and kurtosis measures for the dataset.

The regression analysis shows a strong positive correlation between a preceding and a succeeding year CiteScores. The average Pearson's correlation coefficient (r) works out to 0.98, which indicates a very high positive correlation. Similarly, the

Table 7: Comparison of regression results with actuals for top 10 journals for 2020-21.

Source title	2021(A)	2020(A)	2021(P)	Accuracy
Ca-A Cancer Journal for Clinicians	716.2	463.2	651.5	0.91
Nature Reviews Molecular Cell Biology	140.9	99.7	135.3	0.96
The Lancet	115.3	91.5	123.6	1.07
New England Journal of Medicine	110.5	80.6	108.2	0.98
Reviews of Modern Physics	102.0	86.5	116.5	1.14
Chemical Reviews	98.8	96.9	131.3	1.33
Nature Reviews Materials	96.7	115.7	158.0	1.63
Nature Medicine	91.9	62.4	82.3	0.90
Living Reviews in Relativity	84.8	67.4	89.4	1.05
Nature Energy	78.0	68.7	91.3	1.17

Source: Authors calculations.

(A) = Actuals, (P) = Projected.

average R^2 for the ten regressions is 95.3%, indicating that the previous year's CiteScore explains a sizable variability of the succeeding years' CiteScore. These results establish our claim that the preceding year's scores strongly predict the succeeding years' CiteScores.

We tested our results for the years 2020-21 based on the equation $2021 = -6.29 + (1.42 * 2020)$ for the top 10 journals of 2021. We derived CiteScores for these ten journals using the regression equation and compared the projected results with the actual 10 CiteScores. This comparison is shown in Table 7.

2021(A) and 2020(A) are the ten journals' actual 2021 and 2020 CiteScores. 2021(P) is the projected CiteScore based on the equation $2021 = -6.29 + (1.42 * 2020)$. Accuracy has been calculated by dividing the 2021(P) CiteScore by the 2021(A) CiteScore. For instance, 2021(P), that is, the projected score for Ca-A Cancer Journal for Clinicians, was 651.5, while 2021(A), that is, the actual score was 716.2 giving us an accuracy of 0.91 ($651.5/716.2$). We extended this exercise for all 400 journals, and the overall average accuracy between 2021(P) and 2021(A) was 0.97, a fairly reasonable accuracy for a statistical projection.

The method to do all this is relatively easy. A non-specialist using MS Excel can produce forecasts with little expertise and in-depth knowledge. We summarize the steps for the method with the help of our 2021-20 exercise, based on which we show a forecast of a 2022 CiteScore.

Select the top 400 journals from the Scopus Source website (Scopus, 2022) by ticking the "All" option available at the top of the list, as shown in the screenshot in Figure 1.

The red color appearing in the box indicates that the journals displayed are selected.

Create an Excel list of the CiteScores for 2021 by clicking the button "Export to Excel" (next to the "All" button").

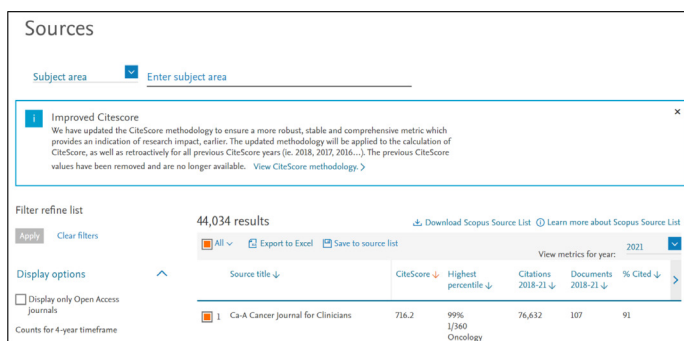


Figure 1: Screenshot of the Scopus source list with the "All" option selected.

(Source: Scopus, 2022).^[3]

Repeat the same procedure in 1 and 2 above by changing the year from 2021 to 2020, 2019 up to 2011, and creating lists of 1000 top journals for the years 2020 up to 2011.

Import the 2020-11 data in the 2021 file.

Using a =vlookup command, extract the 2020-11 scores for the 2021 top 400 journals from the list of 1000 2020-11 journals.

Manually enter scores for any "N/A" entries in the 2020-11 columns.

After these steps, a dataset, as shown in Figure 2, will be created.

Arrange the data for a given journal in ascending order. For example, we have chosen Chemical Reviews (a journal with some up-and-down scores).

Use the =FORECAST function to predict the score for the year 2022. The arrangement, forecast, and formula are shown in Figure 3.

The MS Excel =FORECAST function is a relatively simple method of forecasting based on past data. Its syntax is =FORECAST(x,known_ys,known_xs). X is the period for which

	A	M	N	O	P	Q	R	S	T	U	V	W
1	Source titl	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
2	Ca-A Canc	716.2	463.2	435.4	387.2	290.8	237.4	189.1	210.9	196.9	157.3	150.5
3	Nature Re	140.9	99.7	73.4	62.6	76.0	61.7	64.3	58.9	48.9	42.2	36.0
4	The Lancet	115.3	91.5	73.4	64.9	57.9	51.4	55.2	48.2	49.4	45.8	45.0
5	New Engla	110.5	80.6	66.1	73.1	67.2	61.1	57.4	57.3	58.0	60.2	56.0
6	Reviews of	102.0	86.5	75.8	67.7	63.2	62.3	59.8	76.8	80.7	90.2	76.6
7	Chemical f	98.8	96.9	100.5	94.5	85.8	66.6	72.5	67.1	78.6	65.8	56.3
8	Nature Re	96.7	115.7	123.7	74.4	34.2	32.4	31.2	29.9	29.2	28.1	26.8
9	Nature Me	91.9	62.4	45.9	48.3	47.5	41.2	38.5	31.2	26.1	27.2	29.0
10	Living Rev	84.8	67.4	54.1	33.6	65.7	77.1	49.9	36.7	34.9	25.3	26.3
11	Nature En	78.0	68.7	71.2	43.7	28.4	26.9	25.9	24.8	24.3	23.3	22.2
12	Nature Re	77.1	78.3	70.4	70.3	64.7	61.9	66.6	64.7	47.1	40.0	31.3
13	Cell	77.0	63.4	58.7	56.2	54.9	53.5	54.6	53.9	52.4	55.5	56.8
14	Chemical S	75.9	72.4	67.1	70.7	72.7	66.4	61.6	51.9	42.8	43.9	38.1
15	Progress in	70.9	61.7	47.1	39.6	45.6	44.2	54.5	47.0	37.8	33.7	29.9

Figure 2: Dataset snapshot.

Source: Authors compilation.

	A	B	C	D	E
5	Year CiteScores				
6	2011	56.3			
7	2012	65.8			
8	2013	78.6			
9	2014	67.1			
10	2015	72.5			
11	2016	66.6			
12	2017	85.8			
13	2018	94.5			
14	2019	100.5			
15	2020	96.9			
16	2021	98.8			
17	2022	106.0	=FORECAST(A17,B6:B16,A6:A16)		
18					

Figure 3: Forecasting of 2022 CiteScore for the journal Chemical Reviews.

the value is to be forecasted, known_ys are the past values of the variable, and known_xs is the past period for which the values are known. Using this formula, we have a projected CiteScore for the journal Chemical Reviews for 2022 as 106. Using the same function, we can also forecast the 2022 CiteScores for other journals based on their past years' CiteScores. The forecasting at an individual journal level is backed up by the overall ten years regression calculations which establish that the CiteScores of succeeding years are strongly predicted by those of the preceding years with an average Pearson correlation coefficient (r) of 0.98 and R^2 of 95%.

Our study's overall findings align with those of Croft and Sack (2022).^[7]

LIMITATIONS

Our entire analysis is subject to some serious limitations which ought to be mentioned. First, we have used a sample of 400 journals, so sampling limitations^[40] apply to our study. Second, we manually entered the "not available" values in the dataset for the number of journals, as stated in Table 1. Even though we

have applied a factor keeping in mind the descending trend of the overall CiteScore in moving backward from 2021 to 2011, the overall accuracy has been compromised to some extent due to these estimates. The third and most important thing to be noted is that statistical forecasting suffers from limitations (Gordon, 2010).^[41] As we have shown in Table 7, the predicted and actual results can vary. While on an overall basis, the results for the 400 top 2021 journals predicted and actual 2020 CiteScores have an average accuracy index of 0.97, the variation can be much more on a case-to-case basis. Users must keep this limitation in mind and exercise their judgment in interpreting the forecasts. The forecasting faces problems with abnormal cases like a Ca-A Cancer Journal for Clinicians, which has very high CiteScores with wide variations. Such cases of outliers also adversely affect the normality of the datasets.

CONCLUSION

CiteScore has emerged as an important journal impact factor metric over the past few years. It has gained popularity due to its wide coverage of journal rankings, transparency in methods, comprehensiveness in approach, and free user access. Based on the preceding years' CiteScores, it is possible to predict the succeeding years' CiteScores. While there are some obvious limitations in doing this, including the possibility of an inaccurate prediction, on an overall basis, it has been found that there is a strong positive correlation between the CiteScores of preceding and succeeding years. Academic spaces are strongly competed. Researchers want that their research should get published in a better journal. Towards this, the forecasting of the CiteScores can be useful. Apart from the researcher, the publishers, editorial staff, indexing services, university authorities, and funding agencies can be interested in a projected CiteScore. The dataset created by us for the top 2021 400 Scopus-indexed journals and the subsequent analysis reveals a strong positive correlation between two consecutive years of CiteScores. As a result, it is possible to predict the next year's CiteScore based on historical data of CiteScores. The forecasting technique can be simple if we use a formula like = FORECAST () in MS Excel. It does not require knowledge of advanced computing skills like machine learning and others. Common users of CiteScores can comfortably do the forecasting. However, they should exercise due caution while doing this. They should remember that a forecast is only a prediction, and actual results may vary case-to-case basis. However, the forecast has a directional value and can help the researchers and others make better-informed decisions. The forecast scientifically captures the trend based on past data and is any time better than a wild guess.

More research is invited in this area of predicting CiteScores in a simple manner that people with even basic computing skills can easily use. For future studies it is recommended to select journals by area of knowledge, since journals have a citation rate depending on the area to which they belong.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

- Teixeira da Silva JA. CiteScore: Advances, evolution, applications, and limitations. *Publishing Research Quarterly*. 2020;36(3):459-68. <https://doi.org/10.1007/s12109-020-09736-y>
- Van Noorden, R. Impact factor gets heavyweight rival. *Journal Citation Reports*. 2016;30:20. <https://www.nature.com/articles/nature.2016.21131.pdf?origin=ppub> Accessed 23 Nov 2022.
- Scopus. Sources. *Scopus Preview*; Available from: <https://www.scopus.com/sources.uri?zone=TopNavBar&origin=searchbasic>. 2022.
- Fang H. Analysis of the new scopus CiteScore. *Scientometrics*. 2021;126(6):5321-31. <https://doi.org/10.1007/s11192-021-03964-5>
- James, C., Colledge, L., Meester, W., Azoulay, N., Plume, A. CiteScore metrics: Creating journal metrics from the Scopus citation index. 2018. *arXiv preprint arXiv:1812.06871*. <https://doi.org/10.48550/arXiv.1812.06871>
- Torres J. An Innovative Approach to Bridging Open Access, Collection Development, and Faculty: Altmetric and CiteScore Analyses at a Large Public University. *Qualitative and Quantitative Methods in Libraries*. 2022;11(2):385-411. <https://www.qqml-journal.net/index.php/qqml/article/view/769>
- Croft WL, Sack JR. Predicting the citation count and CiteScore of journals one year in advance. *Journal of Informetrics*. 2022;16(4):101349. <https://doi.org/10.1016/j.joi.2022.101349>
- Martin BR. Editors' JIF-boosting stratagems – Which are appropriate and which not? *Research Policy*. 2016;45(1):1-7. <https://doi.org/10.1016/j.respol.2015.09.001>
- Matthews D. Journal impact factors 'no longer credible'. 2015. <https://www.timeshighereducation.com/news/journal-impact-factors-no-longer-credible>. Accessed 23 Nov 2022.
- Vanclay JK. Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*. 2012;92(2). <https://doi.org/10.1007/s11192-011-0561-0>
- Archambault É, Larivière V. History of the journal impact factor: Contingencies and consequences. *Scientometrics*. 2009;79(3):635-49. <https://doi.org/10.1007/s11192-007-2036-x>
- Okagbue HI, Atayero AA, Adamu MO, Bishop SA, Oguntunde PE, Opanuga AA. Exploration of editorial board composition, Citescore and percentiles of Hindawi journals indexed in Scopus. *Data in Brief*. 2018;19:743-52. <https://doi.org/10.1016/j.dib.2018.05.066>
- Okagbue HI, Bishop SA, Adamu PI, Opanuga AA, Obasi EC. Analysis of percentiles of computer science, theory and methods journals: CiteScore versus impact factor. *DESIDOC Journal of Library and Information Technology*. 2020;40(1):359-65. DOI: 10.14429/djlit.40.1.14866
- Okagbue HI, Teixeira da Silva JA. Correlation between the CiteScore and Journal Impact Factor of top-ranked library and information science journals. *Scientometrics*. 2020;124(1):797-801. <https://doi.org/10.1007/s11192-020-03457-x>
- Trapp JV. The new Scopus CiteScore formula and the Journal Impact Factor: a look at top ranking journals and middle ranking journals in the Scopus categories of General Physics and Astronomy, Materials Science, General Medicine and Social Sciences. *Physical and Engineering Sciences in Medicine*. 2020;43(3):739-48. <https://doi.org/10.1007/s13246-020-00903-1>
- Salisbury L. Scopus CiteScore and Clarivate Journal Citation Reports. *The Charleston Advisor*. 2020;21(4):5-15. <https://doi.org/10.5260/chara.21.4.5>
- Dias NW. The growing international relevance of Ambiente and Água according to Scopus CiteScore results. *Revista Ambiente and Água*. 2021;16. <https://doi.org/10.4136/ambi-agua.2670>
- Atayero AA, Popoola SI, Egeonu J, Oludayo O. Citation analytics: Data exploration and comparative analyses of CiteScores of Open Access and Subscription-Based publications indexed in Scopus (2014–2016). *Data in brief*. 2018;19:198-213. <https://doi.org/10.1016/j.dib.2018.05.005>
- Henao-Rodríguez C, Lis-Gutiérrez JP, Bouza C, Gaitán-Angulo M, Viloria A. Citescore of publications indexed in Scopus: an implementation of panel data. In *International Conference on Data Mining and Big Data*, 2019:53-60. Springer, Singapore. DOI: 10.1007/978-981-32-9563-6_6
- Krauskopf E. Sources without a CiteScore value: more clarity is required. *Scientometrics*. 2020;122(3):1801-12. <https://doi.org/10.1007/s11192-020-03350-7>
- Fernandez-Llimos F. Differences and similarities between journal impact factor and citescore. *Pharmacy Practice (Granada)*. 2018;16(2). <https://dx.doi.org/10.18549/pharmpract.2018.02.1282>
- Colledge L, James C, Azoulay N, Meester W, Plume A. CiteScore metrics are suitable to address different situations—A case study. *Euro. Sci. Edit*. 43(2):27-31. DOI:10.20316/ESE.2017.43.003
- Teixeira da Silva JA. CiteScore: risk of copy-cat, fake and misleading metrics. *Scientometrics*. 2021;126(2):1859-62. <https://doi.org/10.1007/s11192-020-03791-0>
- Brown T, Gutman SA. Impact factor, eigenfactor, article influence, scopus SNIP, and SCImage journal rank of occupational therapy journals. *Scandinavian Journal of Occupational Therapy*. 2018;26(7):475-483. <https://doi.org/10.1080/11038128.2018.1473489>
- Meho LI. Using Scopus's CiteScore for assessing the quality of computer science conferences. *Journal of Informetrics*. 2019;13(1):419-33. <https://doi.org/10.1016/j.joi.2019.02.006>
- Wahakit S, Boonsom N, Kusakunniran W, Thongkanchorn K. Construction of CiteScore based metric for Conferences on a subject area of Computer Science in Scopus. In *25th International Computer Science and Engineering Conference (ICSEC)*. 2021:289-94. IEEE. <https://doi.org/10.1109/ICSEC53205.2021.9684581>
- Al-Hoorie AH, Vitta JP. The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research*. 2019;23(6):727-44. <https://doi.org/10.1177/1362168818767191>
- Villaseñor-Almaraz M, Islas-Serrano J, Murata C, Roldan-Valadez E. Impact factor correlations with Scimago journal rank, source normalized impact per paper, Eigenfactor score, and the CiteScore in radiology, nuclear medicine and medical imaging journals. *La radiologia medica*. 2019;124(6):495-504. <https://doi.org/10.1007/s11547-019-00996-z>
- Liu Z. A bibliometric study of family studies journals using journal impact factors, CiteScore and H-index. *International Journal of Librarianship*. 6(1):1-2. <https://doi.org/10.23974/ijol.2021.vol6.1.174>
- Erfanmanesh M. Status and quality of open access journals in Scopus. *Collection building*. 2017;36(4):155-162. <https://doi.org/10.1108/CB-02-2017-0007>
- Okagbue HI, Akhmetshin EM, Teixeira da Silva JA. Distinct clusters of CiteScore and percentiles in top 1000 journals in Scopus. *COLLNET Journal of Scientometrics and Information Management*. 2021;15(1):133-43. <https://doi.org/10.1080/09737766.2021.1934604>
- Zolfaghari Z, Shokrpour N, Ghahramani L, Sarveravan P. CiteScores of cardiology and cardiovascular journals indexed in Scopus in 2019: A bibliometric analysis. *European Science Editing*. 2022;48:e73949. <https://doi.org/10.3897/ese.2022.e73949>
- Rajkumar KV, Adimulam Y, Subrahmanyam K. A critical study and analysis of journal metric 'CiteScore' cluster and regression analysis. *International Journal of Engineering and Technology*. 2018;7(2.7):28-32.
- Okagbue H, Adamu P, Bishop S, Obasi E, Akinola A. Curve estimation models for estimation and prediction of impact factor and citescore using the journal percentiles: A case study of telecommunication journals. 2019;31-41. <https://www.learntechlib.org/p/218057/>
- Teixeira da Silva JA, Memon AR. CiteScore: A cite for sore eyes, or a valuable, transparent metric?. *Scientometrics*. 2017;111(1):553-6. <https://doi.org/10.1007/s11192-017-2250-0>
- Ali MF. Evaluating the correlation between different impact indicators for library and information science journals: Comparing the journal citation reports and scopus. *Learned Publishing*. 2021;34(3):315-30. <https://doi.org/10.1002/leap.1353>
- Okagbue HI, Bishop SA, Oguntunde PE, Adamu PI, Opanuga AA, Akhmetshin EM. Modified CiteScore metric for reducing the effect of self-citations. *Telkomnika (Telecommunication Computing Electronics and Control)*. 2019;17(6):3044-9. <http://doi.org/10.12928/telkomnika.v17i6.12292>
- Li Y, Wu C, Yan E, Li K. Will open access increase journal CiteScores? An empirical investigation over multiple disciplines. *PLoS one*. 2018;13(8):e0201885. <https://doi.org/10.1371/journal.pone.0201885>
- Krejcie RV, Morgan DW. Determining sample size for research activities. *Educational and psychological measurement*. 1970;30(3):607-10. <https://doi.org/10.1177/001316447003000308>
- Chaudhuri S, Das G, Datar M, Motwani R, Narasayya V. Overcoming limitations of sampling for aggregation queries. In *Proceedings 17th International Conference on Data Engineering*. 2001;534-42. IEEE. <https://doi.org/10.1109/ICDE.2001.914867>
- Gordon A. The boundaries of quantitative forecasting methods: respecting the limits of determinism. *Foresight: The International Journal of Applied Forecasting*. 2010;19:9-15.

Cite this article: Kumar A, Paliwal JM, Brar V, Singh M, Patil PRT, Raibagkar SS, et al. Previous Year's Cite Score Strongly Predicts the Next Year's Score: Ten Years of Evidence for the Top 400 Scopus-indexed Journals of 2021. *J Scientometric Res*. 2023;12(2):254-63.