

# Analysis of the Characteristics and Influence Factors of High-Impact Papers that Common Concern to Researchers and the Public

Qilong Wu\*, Sumei Chen

School of Management, Wuhan Textile University, Wuhan 430200, Hubei Province, China

\*Corresponding author: Qilong Wu, 15607123225m@sina.cn

---

**Abstract:** The analysis of the evolution of the distribution characteristics of “high-impact papers,” which are highly concerned by researchers and the public, and their influence on academic and social factors, provides a reference for the combination of Altmetrics index and traditional literature measurement index in scientific research evaluation. Selecting 2015 to 2019 Altmetrics TOP100 papers, which are also highly cited papers in Web of Science as the “high-impact papers” data set. This paper explores the basic characteristics of high-impact papers and the influencing factors of Altmetrics indicators on academia and society by using basic statistical methods and SPSS. The results show that high-impact papers mainly come from international authoritative journals in natural science and biomedical categories with higher influence factors; there are more researches on medical and health sciences in high-impact papers, followed by earth and environmental sciences and human society studies; paid acquisition is the main way to obtain high-impact papers, but the proportion of open or free access to papers is increasing; News, Twitter and Blogs are the main ways to spread high-impact papers on social media, Mendeley, Blogs and Facebook have a certain correlation with the academic influence of papers, and Mendeley is the main factor influencing the academic influence. Combining traditional measurement index with Altmetrics index can evaluate the influence of the paper in multiple dimensions more objectively.

**Keywords:** Academic influence; Social influence; High impact paper; Feature analysis; Literature measurement

---

**Publication date:** October 2021; **Online publication:** October 29, 2021

## 1. Introduction

In 2010, J. Priem proposed the term Altmetrics, aiming to develop some new evaluation indicators, such as social media reports and comments, to make up for the shortcomings of the traditional scientific research results evaluation system <sup>[1]</sup>. Altmetrics has been hotly discussed by scholars in the field of scientometrics since it was put forward, and it is considered as an effective supplement to the traditional citation measurement index <sup>[2]</sup>. The rise of new online scientific communication mode has brought new data sources and research perspectives for the academic and social impact evaluation of papers <sup>[3]</sup>, therefore, exploring and analyzing the academic and social influence of the thesis based on Altmetrics indicators has become a research hotspot in this field.

Since Altmetrics was formally put forward, its theory and practical applications have been rapidly enriched and perfected. Because it can gather information about related scientific research activities through the Internet and capture social and academic influences, it is considered to be able to more comprehensively reflect the influence of scientific research activities <sup>[4]</sup>. At present, based on the perspective of Altmetrics,

scholars have carried out a series of studies on the characteristics of high Altmetrics index papers or highly cited papers in different research fields, as well as the relationship between the Altmetrics index and the relationship between the Altmetrics index and other indexes <sup>[5-7]</sup>. However, there is little research on academic papers that are of common concern between the scientific community and the public under the two index systems.

Based on this, this paper takes the high-impact papers that researchers and the public pay attention to from 2015 to 2019 as the research objects, analyzes the evolution of time distribution, subject areas and acquisition methods of high-impact papers in recent five years. And based on Altmetrics, SPSS is used to analyze the contribution of Altmetrics indicators to Altmetrics scores and cited frequency, which fully reveals the evolution of characteristics of high-impact papers and the influence of social media networks on them in academia and society, providing reference for further understanding the relationship between Altmetrics and traditional cited indicators and using Altmetrics indicators.

## **2. Data and methods**

### **2.1. Data sources**

Choose Altmetric.com as the source website for obtaining data. It was founded by Altmetrics, a tool for analyzing and evaluating the Altmetrics index of a single paper. It is currently one of the more mature Altmetrics analysis tools. Since 2013, Altmetric.com has published on its website each year the top 100 academic papers with the highest Altmetric scores in the previous year—Altmetric TOP100, as well as the Altmetrics scores, source journals, discipline distribution, author list, access methods, and differences of these academic papers Media attention and other information <sup>[8]</sup>. This paper selects Altmetrics TOP100 papers published in 2015-2019 as the data sample, and uses the DOI of these papers to retrieve and download the corresponding literature data in the WOS core collection database, and obtains 472 valid data retrieved by WOS, with a matching rate of 94.4%. The data collection time is October 17, 2020.

### **2.2. Data processing**

As the published journals of Altmetrics TOP100 include pre-release platforms such as arXiv, bioRxiv, SSRN, etc., the papers in these platforms are not collected by WOS, some papers cannot obtain relevant information through DOI. The cited times of high Altmetrics index papers from 2015 to 2019 were obtained through WOS, and the highly cited papers were selected by year as the “high-impact papers” data set. After data cleaning and sorting, a total of 336 high-impact papers were obtained from 2015 to 2019, accounting for 72%. Subsequently, descriptive statistical analysis is made on high-impact papers. Secondly, SPSS and other analytical tools are used to analyze the evolution of the correlation coefficients of different Altmetrics indicators on Altmetrics scores and cited frequency in high-impact papers over time, and explore the influence factors of high-impact papers on academia and society.

## **3. Analysis on the evolution of the characteristics of high-impact papers**

### **3.1. Evolution of the distribution of high-impact papers and journals**

**Table 1.** lists the evolution of the distribution of the main source journals for high-impact papers. It can be seen from the table that the papers mainly come from six journals: Nature and its sub-journals, Science and its sub-journals, JAMA and its sub-journals, Lancet and its sub-journals, PNAS and New England Journal of Medicine, the papers from these journals account for about 80% of all papers. These journals are international authoritative journals of natural sciences and biomedicine, they have high impact factors. This shows to a certain extent that the high Altmetrics score paper is consistent with the high impact of the source journal. Among them, medical journals account for a high proportion, indicating that research on medical and health topics can attract public attention and is more likely to be spread by social media.

**Table 1.** The evolution of the distribution of the main source journals for high-impact papers

| Journals                        | 2015 | 2016 | 2017 | 2018 | 2019 |
|---------------------------------|------|------|------|------|------|
| Nature/sub-journals of Nature   | 13/3 | 13/2 | 14/5 | 8/5  | 11/7 |
| Science/sub-journals of Science | 11/5 | 7/2  | 5/2  | 9/5  | 8/4  |
| PNAS                            | 8    | 4    | 3    | 8    | 3    |
| JAMA/sub-journals of JAMA       | 2/4  | 7/7  | 4/5  | 4/2  | 4/3  |
| Lancet/sub-journals of Lancet   | 5/1  | 3/2  | 7/1  | 6/4  | 7/2  |
| British Medical Journal         | 5    | 2    | 4    | 0    | 3    |
| New England Journal of Medicine | 1    | 6    | 3    | 8    | 3    |
| Annals of Internal Medicine     | 1    | 1    | 4    | 0    | 2    |
| Total number                    | 59   | 56   | 57   | 59   | 57   |

### 3.2. The evolution of the subject distribution of high-impact papers

**Table 2.** shows that among the high-impact papers, medicine and health science have been the most concerned research fields of the public for five consecutive years, and their proportion has increased year by year, with a total of 171 related papers. It shows that people pay more and more attention to the field of medicine and health. The second is biological sciences, earth and environmental sciences, and human society studies, which are closely related to human life. The range of disciplines involved in the thesis has also increased in recent years. Although the disciplines of mathematics, chemistry, and materials with strong theoretical expertise have attracted public attention in recent years, their dissemination is still weak. It can be seen that the public's focus can be captured by studying some issues that may involve their own interests or affect them (such as health, environment, etc.).

**Table 2.** The evolution of the main subjects of high-impact papers

| Subjects                         | 2015 | 2016 | 2017 | 2018 | 2019 |
|----------------------------------|------|------|------|------|------|
| Medical and Health Sciences      | 25   | 32   | 41   | 35   | 38   |
| Biological Sciences              | 10   | 7    | 14   | 3    | 4    |
| Earth and Environmental Sciences | 10   | 5    | 7    | 13   | 11   |
| Studies in Human Society         | 7    | 8    | 5    | 3    | 3    |

### 3.3. Evolution of the distribution of access to high-impact papers

**Table 3.** shows that from 2015 to 2017, each year about 50% to 60% of the papers need to be paid for, but in 2018 and 2019, there are fewer and fewer papers that are paid for, and it has dropped to 36% in 2019. In the past five years, the literature that can be obtained in an open way has basically maintained between 30% and 42%; in 2019, the proportion of freely obtained literature has been greatly improved. It can be seen from the changes in the distribution of access methods that in recent years, the public has become more and more accessible to acquire periodical papers, which makes it easier to disseminate papers widely, and at the same time it has gained more audiences, and the attention has been increased accordingly.

High-impact papers are mainly obtained through paid access, but the proportion of open or free access to papers continues to increase. It shows that open or free access to papers can effectively promote the dissemination of papers, making it easier for these academic papers to attract a lot of attention in various media or networks. For high-impact papers, changes in access methods make them more likely to attract attention in various social media and networks. It is easy to attract attention on various social media and

networks, and their Altmetric scores are higher.

**Table 3.** Evolution of the distribution of access to high-impact papers

| Acquisition method | 2015   | 2016   | 2017   | 2018   | 2019    |
|--------------------|--------|--------|--------|--------|---------|
| Open access        | 24     | 29     | 36     | 30     | 35      |
| Paywalled          | 42     | 35     | 38     | 37     | 30      |
| Ratio              | 57.14% | 82.86% | 94.74% | 81.08% | 116.67% |

#### 4. Analysis of influential factors of high-impact papers

SPSS statistical analysis tool is used to calculate the correlation coefficient between Altmetrics index and Altmetrics score and cited frequency, and analyze the influencing factors of high-impact papers in social media and academia. Before choosing the correlation analysis method, SPSS was used to test the statistical data through K-S single sample normality test, and it was found that the sample data did not obey the normal distribution. Therefore, this paper chooses the Spearman correlation analysis method to test the correlation between the indicators.

##### 4.1. Analysis of social influence factors of high-impact papers

Altmetric scores quantitatively reflect the degree to which academic papers receive public attention in social and news media. In the process of papers being paid attention to, scientific and technological knowledge in academic papers is spread and understood, known or discussed by the public, thus reflecting the social influence of academic papers. **Table 4.** shows the correlation analysis between Altmetrics metrics and Altmetric scores of high-impact papers from 2015 to 2019. It can be seen from the table that News, Blogs and Twitter have the most obvious influence on Altmetrics scores, which are significantly correlated with Altmetrics scores for five consecutive years, indicating that News, Twitter and Blogs are the main influencing factors of high-influence papers on social media.

**Table 4.** Correlation between different Altmetrics indicators and Altmetric scores

| Indicators | 2015   | 2016   | 2017   | 2018   | 2019   |
|------------|--------|--------|--------|--------|--------|
| News       | .460** | .579** | .438** | .443** | .485** |
| Blogs      | .594** | .471** | .483** | .301*  | .570** |
| Twitter    | .509** | .435** | .452** | .739** | .620** |
| Facebook   | .259*  | .276*  | .284*  | .380** | 0.234  |
| Google+    | .431** | .310*  | .344** | 0.093  | -0.163 |
| Reddit     | 0.075  | .280*  | .322** | .304*  | .371** |
| Video      | 0.185  | 0.138  | 0.127  | -0.056 | .310*  |
| Wikipedia  | .354** | 0.208  | -0.075 | 0.08   | 0.206  |
| Mendeley   | .598** | 0.149  | .276*  | .369** | .337** |

Note: \*\* means when the confidence level (double test) is 0.01, the correlation is significant; \* means when the confidence level (double test) is 0.05, the correlation is significant.

##### 4.2. Analysis of academic influence factors of high-impact papers

Altmetrics, as an econometric analysis of network indicators, is consistent with the characteristics of data papers only obtained from the network and communicated, and can be used as an effective supplementary

measurement method in addition to citation analysis. Altmetrics can obtain and analyze the utilization behavior of data papers by researchers at any time through the network, thus improving the time-lag problem of citation analysis to some extent. This paper analyzes the correlation between different Altmetrics indicators and citation frequency in high-impact papers, and explores the main factors of academic influence in high-impact papers. **Table 5.** shows the correlation analysis between different Altmetrics indicators and citation frequency of high-impact papers from 2015 to 2019.

It can be seen that Blogs, Mendeley and Facebook have the most obvious influence on citation frequency, and Mendeley has a significant positive correlation with citation frequency for five consecutive years, and the correlation coefficient is the highest among all indicators, indicating that Mendeley is the main influential factor of high-impact papers in academic.

**Table 5.** Correlation between different Altmetrics indicators and cited frequency

| Indicators | 2015   | 2016   | 2017   | 2018   | 2019   |
|------------|--------|--------|--------|--------|--------|
| News       | .347** | -0.108 | 0.127  | 0.181  | 0.039  |
| Blogs      | .406** | .300*  | .255*  | .267*  | 0.127  |
| Twitter    | 0.167  | .414** | 0.22   | .247*  | 0.24   |
| Facebook   | .257*  | .375** | .259*  | 0.188  | 0.186  |
| Google+    | .246*  | .348** | 0.14   | 0.116  | .367** |
| Reddit     | -0.042 | 0.176  | 0.123  | 0.216  | 0.028  |
| Video      | 0.037  | 0.009  | -0.043 | -0.096 | 0.073  |
| Wikipedia  | .461** | 0.177  | 0.17   | 0.016  | 0.041  |
| Mendeley   | .812** | .750** | .750** | .640** | .674** |

Note: \*\* means when the confidence level (double test) is 0.01, the correlation is significant; \* means when the confidence level (double test) is 0.05, the correlation is significant.

## 5. Conclusion and prospect

This research analyzes the characteristics and influence factors of highly cited papers in Altmetrics TOP 100 papers from 2015 to 2019, and explores the characteristics and influence factors of high-impact papers that are of common concern to researchers and the public. Objective conclusions are expected to be helpful to related research.

The new way of online communication not only improves the efficiency of academic communication, but also promotes the further deepening of academic openness. In addition to scholars and researchers, the public also pays extensive attention to the research results that are closely related to social life or life health. Academic influence is not limited to the dissemination among scholars, but has been extended to the dissemination of social influence. Altmetrics, which has developed in response to this trend, has features and advantages that traditional measurement indicators can't match. Altmetrics indicators are not like traditional citation indicators that require time to accumulate. They are obtained instantly, and open data can be obtained for free through API tools. This kind of openness and transparency is worthy of recognition. These advantages enable Altmetrics to open up a new path for the dissemination and impact evaluation of academic achievements. It more reflects the degree of attention of academic papers in social media, and is an effective supplement to traditional citation measurement indicators. Therefore, in the future evaluation system, traditional measurement indicators should be combined with Altmetrics indicators to more objectively evaluate the influence of papers in multiple dimensions.

Of course, due to the limitations of research data and methods, this paper also has some shortcomings, which need to be explored in future research. With the vigorous development of big data technology,

artificial intelligence, and the Internet, social media has become increasingly popular, and the dissemination of scientific knowledge and technological innovation are being affected in more ways. Therefore, the characteristics and influencing factors of high-impact papers will become more and more complex, diversified and dynamic, and the research on the influencing factors of papers needs to be further improved.

### **Disclosure statement**

The author declares no conflict of interest.

### **Author contributions**

Sumei Chen. conceived the idea of the study and performed the experiments. Qilong Wu. analyzed the data and wrote the paper.

### **References**

- [1] Priem J, Costello KL, 2010, How and why scholars cite on twitter. *Proceedings of the American Society for Information Science and Technology*, 47(1): 1-4.
- [2] Vanti N, Sanz-Casado E, 2016, Altmetrics: social media metrics for a more democratic science. *Transinformação*, 28(3), 2016, 24(2): 358-349.
- [3] Sud P, Thelwall M, 2014, Evaluating altmetrics. *Scientometrics*, 98(2): 1131-1143.
- [4] Maricato JM, Vilan Filho JL, 2018, The potential for altmetrics to measure other types of impact in scientific production: academic and social impact dynamics in social media and networks.
- [5] Kim ES, Yoon DY, Kim HJ, et al., 2019, The most mentioned neuroimaging articles in online media: a bibliometric analysis of the top 100 articles with the highest Altmetric Attention Scores. *Acta Radiologica*, 60(12): 1680-1686.
- [6] Baek S, Yoon DY, Lim KJ, et al., 2020, Top-cited articles versus top Altmetric articles in nuclear medicine: A comparative bibliometric analysis. *Acta Radiologica*, 61(10): 1343-1349.
- [7] Cho J, 2021, Altmetrics of Highly Cited Research Papers in Social Science. *Serials Review*, 47(1): 17-27.
- [8] Thelwall M, Haustein S, Larivière V, et al., 2013, Do altmetrics work? Twitter and ten other social web services. *PloS one*, 8(5): e64841.