

Usage Count: A New Indicator to Detect Research Fronts

Guoqiang Liang, Haiyan Hou[†], Zhigang Hu[†], Fu Huang,
Yajie Wang & Shanshan Zhang

WISE Lab, Dalian University of Technology, Dalian 116023, China

Abstract

Purpose: Research fronts build on recent work, but using times cited as a traditional indicator to detect research fronts will inevitably result in a certain time lag. This study attempts to explore the effects of usage count as a new indicator to detect research fronts in shortening the time lag of classic indicators in research fronts detection.

Design/methodology/approach: An exploratory study was conducted where the new indicator “usage count” was compared to the traditional citation count, “times cited,” in detecting research fronts of the regenerative medicine domain. An initial topic search of the term “regenerative medicine” returned 10,553 records published between 2000 and 2015 in the Web of Science (WoS). We first ranked these records with usage count and times cited, respectively, and selected the top 2,000 records for each. We then performed a co-citation analysis in order to obtain the citing papers of the co-citation clusters as the research fronts. Finally, we compared the average publication year of the citing papers as well as the mean cited year of the co-citation clusters.

Findings: The citing articles detected by usage count tend to be published more recently compared with times cited within the same research front. Moreover, research fronts detected by usage count tend to be within the last two years, which presents a higher immediacy and real-time feature compared to times cited. There is approximately a three-year time span among the mean cited years (known as “intellectual base”) of all clusters generated by usage count and this figure is about four years in the network of times cited. In comparison to times cited, usage count is a dynamic and instant indicator.

Research limitations: We are trying to find the cutting-edge research fronts, but those generated based on co-citations may refer to the hot research fronts. The usage count of older highly cited papers was not taken into consideration, because the usage count indicator released by WoS only reflects usage logs after February 2013.

Practical implications: The article provides a new perspective on using usage count as a new indicator to detect research fronts.

Citation: Guoqiang Liang, Haiyan Hou, Zhigang Hu, Fu Huang, Yajie Wang & Shanshan Zhang (2017). Usage Count: A New Indicator to Detect Research Fronts.

Vol. 2 No. 1, 2017

pp 89–104

DOI: 10.1515/jdis-2017-0005

Received: Sept. 26, 2016

Revised: Nov. 27, 2016

Accepted: Dec. 4, 2016



[†] Corresponding authors: Haiyan Hou (E-mail: htieshan@dlut.edu.cn) and Zhigang Hu (E-mail: huzhigang@dlut.edu.cn).

Originality/value: Usage count can greatly shorten the time lag in research fronts detection, which would be a promising complementary indicator in detection of the latest research fronts.

Keywords Research front; Citation analysis; Usage count; Times cited

1 Introduction

Research fronts detection has become the focus of global scientific and technological competition. Through detecting and tracking research fronts timely and accurately, Japan (Kuwahara, 2007; Nagano, 2005) and the US (Porter, Guo, & Chiavatta, 2011) have made significant advancements in science and technology (S&T) policy-making and technological evaluation.

Research front was originally a term used by Price (1965). He concluded that a research front is a small part of earlier literature knitted together by the new year's crop of paper, and used the phrases "epidermal layer" and "growing tip" to describe research fronts. Since then, scholars have made efforts to identify research fronts from various points of view. Small and Griffith (1974) considered co-citation clusters as research fronts; Vlachý (1984) summarized prior research on scientometric studies on research fronts detection and pointed out that "science grows from a very thin skin of its research front" and "a core body of seminal literature" constitutes "a sort of epidermal layer, an active research front" (p. 95). Garfield (1994) pointed out that research fronts are co-citation clusters plus citing articles; Morris et al. (2003) applied bibliographic coupling methods to identify the research fronts; Shibata et al. (2008) proposed that research fronts are direct citation clusters. Presently, Chen (2006), Braam, Moed, and van Raan (1991), and Persson's (1994) views are the mainstream in research fronts detection. They concurred that research fronts are clusters of citing papers sharing a common intellectual base.

The research fronts detection method of this paper is in accordance with Chen, Braam et al., and Persson's points of view. We label groups of citing articles that cite clusters of co-cited references as research fronts, and they are labeled from the title of an article which cites the most references in the cluster. When beginning co-citation analysis, the usual method is first to set a threshold and find the representative highly cited papers by "times cited," and then make a co-cited matrix before clustering the networks and identifying the research fronts. However, due to the large time lag, there is a problem in using times cited as an indicator in research fronts identification. It might take up to two years for a paper to become highly cited (Shibata et al., 2008), and the situation varies among disciplines. Besides,



times cited is affected by authors' different citing motivations, articles accessibility (Bollen et al., 2005), etc. Therefore, as a traditional indicator, times cited cannot reflect the current interests of the research community. Faced with the fast paced development of S&T, new methods, tools, and indicators need to be developed to capture the research fronts more precisely in order to support S&T policy-making.

Open access journal publishers such as PLoS[®] took the lead to provide online usage data for published articles. When more and more publishers chose to disclose the usage data of academic articles, researchers proposed to use articles' usage data as potential complements, perhaps even alternatives, for research evaluation (Das & Mishra, 2014; Yan & Gerstein, 2011). They emphasized the superiority of usage data to citation data, such as ease to access and more convenient data collection. However, there are divergent opinions with regard to whether usage data such as times of views and downloads can be used as metrics of research evaluation. Some researchers reported significant correlation between specific usage types (Line & Sandison, 1975), especially downloads and citations, but others (Schloegl & Gorraiz, 2010; 2011) found only moderate or a rather low correlation between downloads and citations.

On September 26, 2015, Thomson Reuters started to provide article-level usage data, called "usage count" on the Web of Science (WoS) platform (Thomson Reuters, 2015). The new indicator, consisting of "U1" and "U2," reflects the user's level of interest by marking an article when read or downloaded by researchers. U1 is the count of the number of times the full text of a record has been accessed or saved within the last 180 days. U2 is the count of the number of times the full text of a record has been accessed or saved since February 1, 2013. The usage count is recorded every second day after users full-text request of an article, exports to bibliographic management tools or to formats for later import into bibliographic management tools, thus it does not need to wait for the tedious submitting and publishing process of times cited.

There is limited research on the effects of usage data on research evaluation. Presently, we have retrieved two relative articles (Martín-martín, 2016; Wang, Fang, & Sun, 2016) involving the relationship between WoS usage count and times cited. They discussed the usage patterns of articles and the correlation between usage count and times cited. In their research, Wang, Fang, and Sun (2016) discovered that citations play an important role in determining the usage count for old papers, and highly cited old papers are more likely to be used for even a long time after



[®] <http://blogs.plos.org/plos/2009/09/article-level-metrics-at-plos-addition-of-usage-data/>

publication. Following their study, we are curious whether usage count² can become a new and more effective indicator in detection of the latest research fronts, which are valuable for decision-makers to track the latest research trend and take the lead in scientific competition. To compare research fronts detected based on different indicators, we use “recentness” as a measure and we define recentness of a research front as the average publication year of citing papers of the co-citation cluster. Our research questions are:

- 1) *What is the difference in recentness of research fronts generated by usage count and times cited?*
- 2) *Are there any common research fronts detected by usage count and times cited, and if there are, what about the recentness?*
- 3) *What is the difference in recentness of the top 10 highly cited papers selected by usage count and times cited?*

2 Data and Methodology

2.1 Data Source

This article takes the regenerative medicine domain as an example to retrieve data. After using the topic search “regenerative medicine” in titles, abstracts or keywords (including keywords plus), and filtering out less representative record types, such as proceeding papers, meeting abstracts, news items, letters, etc., a total of 10,545 records dated between 2000 and 2015 were downloaded from WoS. We gathered the top 2,000 records sorted by both times cited and usage count. We regard these records as a representative dataset that can reflect the total records downloaded. The exported data format, search strategy, indexes, time span, and document types of the two indicators remained consistent.

2.2 Methodology

Co-citation analysis is the typical bibliometric method, initially proposed by Small and Griffith (1974). Articles were clustered together based on their co-occurrence in the references lists of papers. In other words, if articles A and B are both cited by article C, it is more likely that they belong to the same research field and share similar topics or methods.

To identify clusters which represent the intellectual bases, spectral clustering algorithms have been used in this paper. Spectral clustering is a clustering method that uses eigenvectors of an affinity matrix derived from the data (Dhillon, 2004),

² In this article, we mainly discuss U1. In the subsequent sections of this article, usage count refers to U1, the count of the number of times the full text of a record has been accessed or saved within the last 180 days.



and results derived by spectral clustering often outperform the traditional algorithms such as k -means or single linkage (von Luxburg, 2007). Besides, spectral clustering is easy to implement. This study uses CiteSpace 4.0.R5 (Chen, 2006) to identify the co-citation clusters and furthermore to find the citing articles of each cluster as well as the research front.

There are 84,315 and 113,339 references in dataset “times cited” and dataset “usage count,” respectively. In order to eliminate records that have little relationship with our research and pick out the most frequently used references, we select the top 1% most-cited records within the datasets to be further analyzed. In order to reduce disparities caused by absolute frequencies, we construct the co-citation matrix in terms of cosine coefficients. Additionally, a minimum spanning tree network is used in the software for network pruning in order to hide relatively weak citation links between item i and item j in the matrix, and improve the pruning efficiency as well. We compare the average publication year of citing articles (the recentness) as well as the mean cited year of the co-citation clusters. In this paper, average publication year of citing articles is equal to the mean publication year of the citing publications, and does not take the number of citations into account. Figure 1 presents organization of this study.

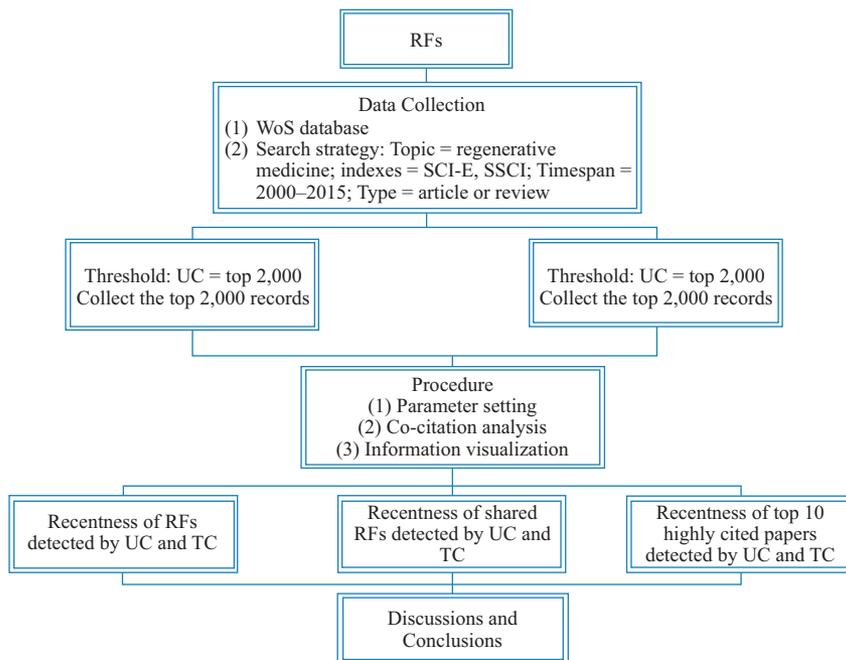


Figure 1. Organization and procedure of the present study. RF is the abbreviation for research front and TC for times cited. UC refers to usage count.



3 Results

3.1 Basic Statistic Comparison of Times Cited and Usage Count

We investigate the distribution of citing articles and cited references within the two datasets. Figure 2 shows that the citing articles of usage count have experienced an exponential growth, because researchers prefer to use newly published literature, while Figure 3 indicates a right-skewed distribution for citing articles of times cited. This phenomenon is an indication that it will take several years for a paper to become highly cited. The distribution of cited references in Figures 4 and 5 presents a right-skewed distribution, but the data collected by usage count present a relatively high real-time property when compared with that collected by times cited. Many cited references were gathered from 2010 to 2013 by usage count in comparison with 2005–2008 by times cited.

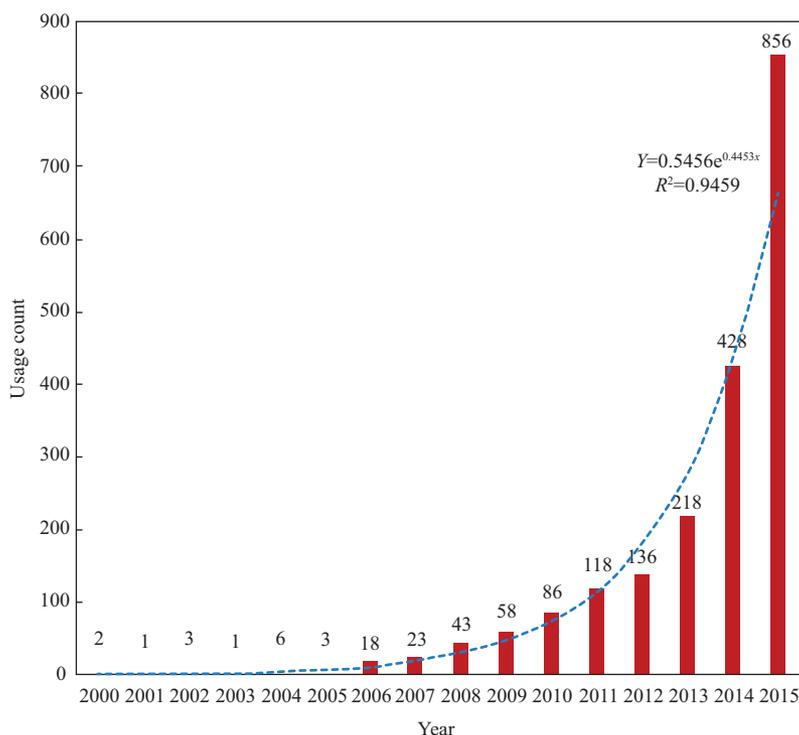


Figure 2. Yearly frequency distribution of citing articles with threshold of usage count. Y represents the usage count and R^2 reflects the goodness-of-fit of this equation. A value closer to 1 indicates a satisfactory goodness-of-fit.

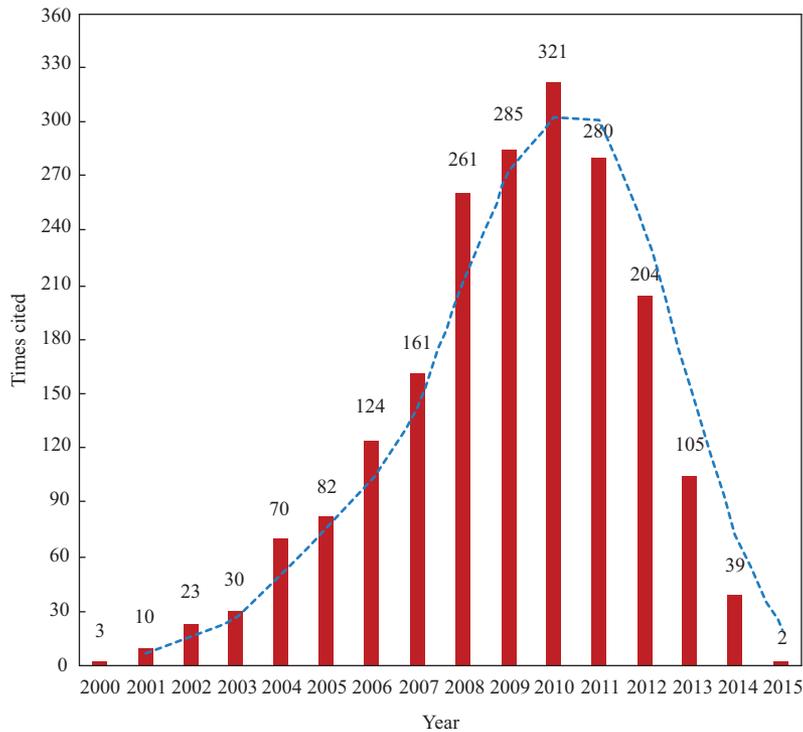


Figure 3. Yearly frequency distribution of citing articles with threshold of times cited.

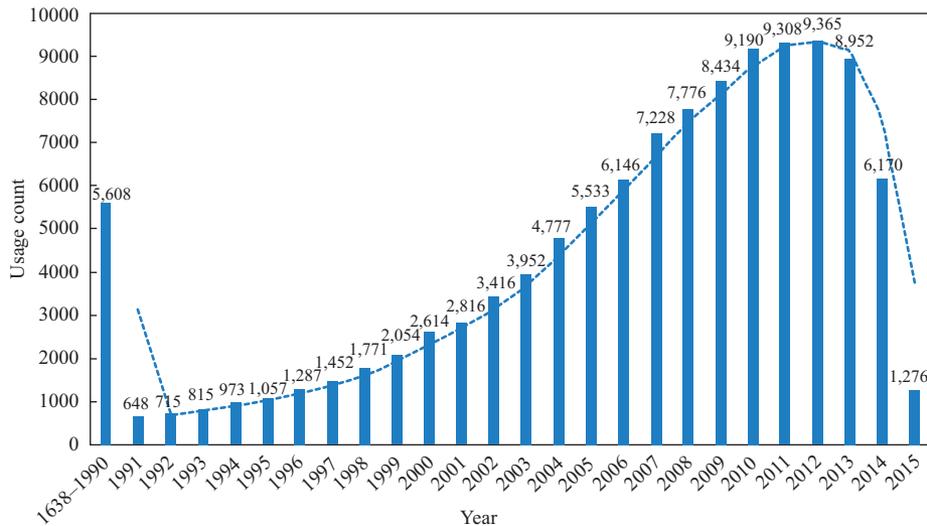


Figure 4. Yearly frequency distribution of cited references with threshold of usage count. The time span of cited references is 1638–2015, and the study integrated the data of 1638–1990 to facilitate reading.



Research Paper

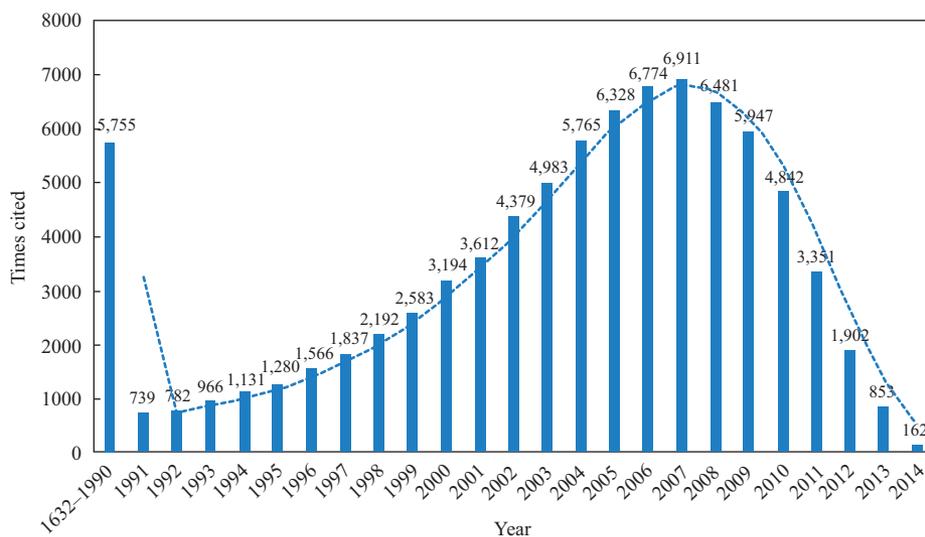


Figure 5. Yearly frequency distribution of cited references with threshold of times cited. The time span of cited references is 1632–2014, and the study integrated the data of 1632–1990 to facilitate reading.

There is approximately a four-year time lag in the mean publication year of citing articles, as well as a three-year time lag in the mean publication year of cited references collected by times cited. As is shown in Table 1, the recentness in times cited is 2009 and 2013.3 in usage count. As for the mean publication year of cited references, the figure is 2002.6 in times cited and 2005.7 in usage count.

Table 1. Overview of the dataset.

	Times cited	Usage count
No. of cited references	84,315	113,339
Recentness	2009.0	2013.3
Mean publication year of cited references	2002.6	2005.7

Note. Recentness means the average publication year of citing papers.

3.2 Comparison of the Recentness of Research Fronts

In this section, we compute the recentness of each research front. Table 2 lists the details of the two networks in which the number of articles in each cluster is above five, and the study ranks each cluster by recentness. The method to label the cluster is based on word profiles derived from the title of an article which cites the most co-cited articles within the cluster (Chen, 2006).

Tables 2 and 3 show there are 20 research fronts detected by the times cited, and 26 detected by usage count. The recentness in a majority of clusters detected by usage count tends to be published within the last four years, while this figure is



relatively less recent in those detected by times cited. The recentness is 2011.59 detected by usage count and 2009.07 by times cited.

Table 2. Recentness of the research fronts detected by times cited.

Clusters	No. of references	Mean cited year	No. of citing articles	Recentness
Emerging peptide nanomedicine	29	2008	54	2010.44
Pluripotent stem cell	25	2007	56	2008.89
Adipose-derived stem cell	25	2003	83	2010.14
Somatic cell	22	2008	52	2009.33
Mesenchymal stem cell	22	2003	55	2010.07
Induced pluripotent stem cell	22	2008	41	2009.05
Embryonic stem cell	22	2004	60	2010.08
Organ level tissue engineering	21	2008	50	2009.80
Mesenchymal stromal cell	21	2005	52	2009.33
Synthetic hydrogels	21	2002	48	2010.45
Hippo pathway	20	2008	52	2010.69
Human induced pluripotent stem cell	19	2009	64	2010.36
Human embryonic stem cells	18	2007	40	2008.05
Regenerative biology	18	2001	51	2010.25
Marrow-derived mesenchymal cell	18	2001	38	2007.34
Human wharton	17	2006	27	2010.00
Genetic modification	17	2002	31	2007.81
Generation	16	2008	40	2007.40
Therapeutic application	16	2002	41	2010.00
Review	8	2000	8	2002.00
Mean value of all clusters	19.85	2005	47.15	2009.07

Note. Recentness means the average publication year of citing papers.

As indicated in Tables 2 and 3, a majority of research fronts generated by usage count tend to be more newly published compared to times cited. Because usage count is a reflection of researchers' interest level within the last 180 days, most researchers pay more attention to achievements published within the previous two-three years, in order to stay in step with their colleagues and keep abreast of what is going on across scholarly communities. Besides, usage count can capture users' full-text searching behaviors instantly instead of waiting for the tedious publishing process compared with times cited. Due to this, it is reasonable that a large proportion of citing articles in each cluster created by usage count are newly published. Meanwhile, we also observed that articles detected by usage count were published almost two years later than those detected based on times cited accordingly, because the newly published citing articles are prone to cite recent achievements due to rapid knowledge updates in the regenerative medicine field.

3.3 Comparison of the Recentness of the Common Research Fronts

The two indicators both detect the induced pluripotent stem cell (iPSc) as one of the research fronts in the regenerative medicine field. We calculated all the citing



Table 3. Recentness of the research fronts detected by usage count.

Clusters	No. of references	Mean cited year	No. of citing articles	Recentness
Clinic	5	2012	39	2014.35
Whole organ engineering	25	2011	39	2013.77
Expansion	22	2011	55	2013.46
Hydrogel	27	2010	52	2013.32
Overview	26	2010	36	2013.21
Extracellular vesicle	26	2010	49	2013.02
Regulating stem cell fate	23	2010	61	2013.00
Induction	23	2010	44	2013.00
Induced pluripotent stem cell differentiation	21	2010	38	2013.00
Carbon nanotube	19	2010	40	2012.72
Human pluripotent stem	19	2010	22	2012.60
Stem cell application	26	2009	39	2012.28
Peptide	24	2009	32	2012.05
Porous scaffold	21	2009	42	2011.81
Poly	5	2009	35	2011.50
Layer	27	2008	66	2011.49
Biomedicine	25	2008	57	2011.31
Induced pluripotent stem cell	25	2008	36	2011.13
Glycosaminoglycan-binding substratum	24	2008	36	2011.09
Pro-angiogenic properties	26	2007	27	2010.96
Nanotechnologies	16	2007	34	2010.23
Water filtration	31	2005	51	2010.20
Supramolecular design	25	2005	20	2010.02
Biodegradable hydrogel	24	2005	8	2009.90
Present status	21	2004	20	2009.37
Biological characterization	5	1998	3	2002.67
Mean value of all clusters	21.58	2008.19	37.73	2011.59

Note. Recentness means the average publication year of citing papers.

articles of the IPSc field detected by the two indicators (Table 4 only lists the top 10 citing articles), to compare the recentness of the common research fronts detected by usage count and times cited. There are 55 citing articles in the IPSc field detected by times cited and 22 by usage count. The recentness in the IPSc field created by usage count is 2011.09 and 2010.07 by times cited. Moreover, the two indicators found seven common papers in the IPSc field.

As illustrated in Table 4, the two indicators can both detect research fronts in the IPSc field, and the citing articles of the IPSc field detected by usage count tend to be published more recently than times cited. Takahashi and Yamanaka (2006) originally introduced IPSc in 2006, and showed that the introduction of four specific gene encoding transcription factors could convert adult cells into pluripotent stem cells, and was awarded the 2012 Nobel Prize. As the shortage of donor organs for treating end-stage organ failure highlights the need for generating organs from IPSc



Table 4. Comparison of the top 10 citing papers of the common research front.

Coverage (%)	Times cited		Publishing year	Usage count		Publishing year
	Citing articles	Citing articles		Coverage (%)	Citing articles	
55	Wang, Y. <i>A transcriptional roadmap to the induction of pluripotency in somatic cells.</i>	2010	40	Patel, M. <i>Advances in reprogramming somatic cells to induced pluripotent stem cells.</i>	2010	
50	Kiskinis, E. <i>Progress toward the clinical application of patient-specific pluripotent stem cells.</i>	2010	32	Warren, L. <i>Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA.</i>	2010	
50	Li, W.L. <i>Small molecules that modulate embryonic stem cell fate and somatic cell reprogramming.</i>	2010	24	Ben-David, U. <i>The tumorigenicity of human embryonic and induced pluripotent stem cells.</i>	2011	
50	Masip, M. <i>Reprogramming with defined factors: From induced pluripotency to induced transdifferentiation.</i>	2010	20	Lister, R. <i>Hotsspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells.</i>	2011	
45	Warren, L. <i>Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA.</i>	2010	20	Tsuji, O. <i>Therapeutic potential of appropriately evaluated safe-induced pluripotent stem cells for spinal cord injury.</i>	2010	
41	Cox, J.L. <i>Induced pluripotent stem cells: What lies beyond the paradigm shift.</i>	2010	20	Wu, S.M. <i>Harnessing the potential of induced pluripotent stem cells for regenerative medicine.</i>	2011	
41	Lengner, C.J. <i>IPS cell technology in regenerative medicine.</i>	2010	16	Zhao, T.B. <i>Immunogenicity of induced pluripotent stem cells.</i>	2011	
41	Tamaoki, N. <i>Dental pulp cells for induced pluripotent stem cell banking.</i>	2010	12	Young, R.A. <i>Control of the embryonic stem cell state.</i>	2011	
36	Chun, Y.S. <i>Applications of patient-specific induced pluripotent stem cells: focused on disease modeling, drug screening and therapeutic potentials for liver disease.</i>	2010	8	Burridge, P.W. <i>A universal system for highly efficient cardiac differentiation of human induced pluripotent stem cells that eliminates interline variability.</i>	2011	
36	Nakagawa, M. <i>Promotion of direct reprogramming by transformation-deficient myc.</i>	2010	8	Klim, J.R. <i>A defined glycosaminoglycan-binding substratum for human pluripotent stem cells.</i>	2010	
Recentness		2010.07	Recentness		2011.09	



(Takebe et al., 2013), we can expect that more and more researchers will retrieve and download classical articles in this field, and thus usage count will capture and accumulate the usage logs accordingly. Articles listed in the IPSc field by usage count are expected to be more recently published than those listed by times cited.

3.4 Comparison of the Recentness of the Top 10 Most Highly Cited Papers

In this section, we compare the top 10 most highly cited papers detected by usage count and times cited. Frequency refers to the times cited in local datasets. The results indicate that there are four papers in common listed in Table 5. Coincidentally, these papers rank within the top five results due to frequency. Moreover, there are three articles published before the year 2004 detected by times cited, while no article is published before 2004 in the top 10 detected by usage count.

From Table 5 we can see that the top 10 most highly cited papers selected by usage count and times cited are classical articles, but we find papers selected by usage count tend to be more recently published than those by times cited. The mean year of the top 10 most highly cited papers is 2004.6 detected by times cited, while this figure is 2007.2 when sorted by usage count. It indicates an approximate three-year time span among the mean cited years (known as “intellectual base”) of all clusters generated by usage count in the regenerative medicine domain.

4 Discussion and Conclusion

The study collects 2,000 records by both times cited and usage count to measure whether usage count can be a new indicator in detection of research fronts. We find both indicators can be used in detection of research fronts, but using usage count can detect the latest research fronts than using times cited. In comparing the effects of the two indicators, first, we note that the majority of research fronts generated by usage count tend to be newer than times cited. Second, we investigate the recentness of a common research front detected by usage count and times cited. Results indicate using usage count can detect the latest research fronts than using times cited. Third, we compare the top 10 most highly cited papers detected by usage count and times cited. We find the top 10 papers selected by usage count represent more recent research fronts than selected by times cited. Moreover, we draw the conclusion that research fronts detected by usage count tend to be within the last two years, and present a higher immediacy and real time accuracy compared with times cited. Usage count can greatly shorten the time lag in research fronts detection, which could become a complementary indicator in the recentness detection of research fronts.



Table 5. Top 10 highly cited papers detected by usage count and times cited.

Articles detected by times cited			Articles detected by usage count		
Frequency	Articles	Publishing year	Frequency	Articles	Publishing year
225	Takahashi, K., <i>Cell</i> , V126, P663	2006	168	Takahashi, K., <i>Cell</i> , V131, P861	2007
212	Takahashi, K., <i>Cell</i> , V131, P861	2007	155	Engler, A.J., <i>Cell</i> , V126, P677	2006
188	Yu, J.Y., <i>Science</i> , V318, P1917	2007	110	Slaughter, B.V., <i>Bvadv Mater</i> , V21, P3307	2009
115	Engler, A.J., <i>Cell</i> , V126, P677	2006	107	Yu, J.Y., <i>Science</i> , V318, P1917	2007
113	Domimici, M., <i>Mcytotherapy</i> , V8, P315	2006	104	Takahashi, K., <i>Cell</i> , V126, P663	2006
106	Jiang, Y.H., <i>Nature</i> , V418, P41	2002	77	Dalby, M.J., <i>Nat Mater</i> , V6, P997	2007
99	Okita, K., <i>Nature</i> , V448, P313	2007	74	Ott, H.C., <i>Nat Med</i> , V14, P213	2008
91	Park, I.H., <i>Nature</i> , V451, P141	2008	63	Lutolf, M.P., <i>Nat Biotechnol</i> , V23, P47	2005
89	Pittenger, M.F., <i>Science</i> , V284, P143	1999	61	Discher, D.E., <i>Science</i> , V324, P1673	2009
86	Thomson, J.A., <i>Science</i> , V282, P1145	1998	55	Macchiarini, P., <i>Lancet</i> , V372, P2023	2008
Mean year		2004.6	Mean year		2007.2

Note. Only the first author and the starting page of these articles are listed.



Usage count would be a new indicator in recentness detection of research fronts. If paper A is cited frequently within a period of time, the times cited will be added to WoS once the citing articles are published online. Usage count captures the researchers' preference on various publications within the last 180 days. Generally, researchers prefer to use newly published papers, and therefore the usage data from publications within the last three years will reach a peak with relatively few citations (Wang, Fang, & Sun, 2016). Therefore, the meta data collected by usage count are most likely to be recent publications. In the research front detecting process, cited references are clustered as intellectual base and the citing articles form the "footprints" of research fronts accordingly. Citation activity can lag behind the publication of an article and some research domains are slow to be cited. In this sense, there is a relatively larger time lag in the research front detection based on times cited.

This paper represents preliminary work on the study of usage count in research fronts detection. However, there are some limitations in the study. For instance, the research fronts generated based on co-citations may refer to the hot research fronts, while we are trying to identify the cutting-edge research fronts. The usage count of older highly cited papers were not taken into consideration, because the new usage count indicator released by WoS only reflects usage logs after February 2013. In comparison to times cited, usage count is a dynamic and instant indicator. However, the correlation between usage count and times cited needs to be further discussed in the future.

Acknowledgements

We would like to give special thanks to Nina Chang (Product Manager of Web of Science) for providing detailed information on "usage count." Our gratitude also extends to anonymous reviewers for their valuable comments on the first version of this article, which have greatly improved the readability and accuracy. This work is supported by the National Social Science Foundation of China (Grant No.:14BTQ030).

Author Contributions

H.Y. Hou (htieshan@dlut.edu.cn, corresponding author) and Z.G. Hu (huzhigang@dlut.edu.cn, corresponding author) planned and designed the outline, jointly discussed the findings, and contributed to the final draft. G.Q. Liang (Liang_1988@mail.dlut.edu.cn) proposed the research idea, carried out the data collection and data analysis, and wrote the first draft. F. Huang (hf206@163.com), Y.J. Wang (yjwang55@mail.dlut.edu.cn), and S.S. Zhang (shann1027@sina.com) joined discussion of the findings and put forward valuable suggestions.



References

- Bollen, J., van de Sompel, H., Smith, J.A., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing and Management*, 41(6), 1419–1440.
- Braam, R.R., Moed, H.F., & van Raan, A.F.J. (1991). Mapping of science by combined co-citation and word analysis. I: Structural aspects. *Journal of the American Society for Information Science*, 42(4), 233–251.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377.
- Das, A.K., & Mishra, S. (2014). The spread of scientific information: Insights from the web usage statistics in PLoS article-level metrics. *Journal of Scientometric Research*, 3(2), 1–16.
- Dhillon, I.S. (2004). Kernel k-means, Spectral clustering and normalized cuts. *Compute*, CI(78712), 551–556.
- Garfield, E. (1994). Research fronts. *Current Contents*, 41(10), 3–7.
- Kuwahara, A.S.Y. (2007). Benchmarking S&T capacity and future direction of S&T development in Japan (Policy analysis of the science and technology basic plans). *Journal of Science Policy & Research Management*, 21(1), 28–34.
- Line, M.B., & Sandison, A. (1975). Practical interpretation of citation and library use studies. *College and Research Libraries*, 36(5), 393–396.
- Martín-Martín, A. (2016). Thomson Reuters utiliza métricas: Usage counts para los artículos indexados en la Web of Science. *Anuario ThinkEPI*, 10, 209–221.
- Morris, S.A., Yen, G., Wu, Z., & Asnake, B. (2003). Time line visualization of research fronts. *Journal of the American Society for Information Science and Technology*, 54(5), 413–422.
- Nagano, H. (2005). Comprehensive analysis of science and technology benchmarking and foresight. Nistep Report. Retrieved on November 27, 2016, from <http://ci.nii.ac.jp/naid/40007130230>.
- Persson, O. (1994). The intellectual base and research fronts of JASIS 1986–1990. *Journal of the American Society for Information Science*, 45(1), 31–38.
- Porter, A.L., Guo, Y., & Chiavatta, D. (2011). Tech mining: Text mining and visualization tools, as applied to nanoenhanced solar cells. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(2), 172–181.
- Price, D.J. (1965). Networks of scientific papers. *Science*, 149(3683), 510.
- Schloegl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: The case of oncology journals. *Scientometrics*, 82(3), 567–580.
- Schloegl, C., & Gorraiz, J. (2011). Global usage versus global citation metrics: The case of pharmacology journals. *Journal of the Association for Information Science and Technology*, 62(1), 161–170.
- Shibata, N., Kajikawa, Y., Takeda, Y., & Matsushima, K. (2008). Detecting emerging research fronts based on topological measures in citation networks of scientific publications. *Technovation*, 28(11), 758–775.
- Small, H., & Griffith, B.C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Social Studies of Science*, 4(1), 17–40.



Research Paper

- Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4), 663–676.
- Takebe, T., Zhang, R.R., Koike, H., Kimura, M., Yoshizawa, E., Enomura, M., Koike, N., Sekine, K., & Taniguchi, H. (2013). Generation of a vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature Protocols*, 9(2), 481–484.
- Thomson Reuters. (2015). Usage count. Retrieved on September 10, 2016, from http://images.wbofknowledge.com/WOKRS521R5/help/WOK/hp_usage_score.html.
- Urashima, K. (2007). Comprehensive analysis of science and technology benchmarking and foresighting. *International Journal of Plasma Environment Science & Technology*, 1(1), 3–7.
- Vlachý. (1984). Priority choice and research front specialties in physics. *Czechoslovak Journal of Physics B*, 34(1), 95–98.
- von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4), 395–416.
- Wang, X., Fang, Z., & Sun, X. (2016). Usage patterns of scholarly articles on Web of Science: A study on Web of Science usage count. *Scientometrics*, 109(2), 917–926.
- Yan, K.K., & Gerstein, M. (2011). The spread of scientific information: Insights from the web usage statistics in PLoS article-level metrics. *PLoS ONE*, 6(5), 1–7.



This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

