

# Global Usage vs. Global Citation Metrics Using Science Direct Pharmacology Journals

Christian Schlögl<sup>1</sup> and Juan Gorraiz<sup>2</sup>

<sup>1</sup>*christian.schloegl@uni-graz.at*

University of Graz, Institute of Information Science and Information Systems, Universitaetsstr. 15/F3, A-8010  
Graz (Austria)

<sup>2</sup>*juan.gorraiz@univie.ac.at*

University of Vienna, Library and Archive Services, Bibliometrics Department, Boltzmanngasse 5,  
A-1090 Vienna (Austria)

## Abstract

In our study we investigate similarities and differences of usage and citation metrics for pharmacology journals using data from Science Direct, Web of Science and Journal Citation Reports. Our preliminary results show that the number of downloads have increased more than fivefold between 2001 and 2006. In our analyses at journal level we found a high correlation between citation and download frequencies but differences in obsolescence characteristics. Also the download patterns on a paper-by-paper basis for five journals reveal that articles published in recent years have the highest download requests immediately after they are available electronically. However, the age distribution is different for articles of older volumes which might be due to the strong increase in e-journal use during the investigation period. A more detailed analysis comparing monthly distributions of downloads and citations suggests that citations have an effect on downloads for highly cited papers even a few years after their publication.

## Introduction

After the introduction of e-journals in recent years it has become much easier to collect usage data. Since usage data offer several advantages compared to citation data (for instance, higher timeliness), the question arises whether usage metrics can be considered complementary to citation metrics as suggested by Bollen et al. (2005). However, since e-journal usage data have been available for several years only, there still is a need for further research. Kurz et al. (2005) even expect similarities and differences of reads and citations to become a central issue in bibliometric research.

Various e-journal usage studies have been published in recent years. While most of these analyses were conducted using local usage data (for instance, McDonald, 2007), there are much less studies which were performed at a global level. Usually these studies used either data from subject repositories like, for instance, RePEc (Chu & Krichel, 2007) and arXiv.org (Brody et al., 2006), or open access journals (e.g., Moed, 2005). In our research we use data from Science Direct (SD) which covers around one quarter of the world's full-text scientific, technical and medical literature from more than 2,000 peer-reviewed journals (Elsevier, n.d.).

## Methodology

Our contribution aims at investigating comprehensively the similarities and differences of usage and citation metrics by addressing both

- obsolescence characteristics and the
- relation between citation and download frequencies (absolute and relative indicators).

Both issues are analysed at journal level and at article level. For the analyses at journal level, we use the classical citation indicators from the Journal Citation Reports (JCR) (impact factor, immediacy index and cited half-life) and, in line with Rolands and Nicholoas (2007), calculate the corresponding usage indicators the same way. When computing the usage impact factor and the usage immediacy index, we also use the article counts from the JCR (which refer to citable documents only) in order to guarantee a correct comparison.

We were provided with the usage data for Science Direct journals from the subject category “Pharmacology, Toxicology and Pharmaceutical Science” both at journal level and on a paper-by-paper basis. At journal level, we received the download data for all journals between 2001 and 2006. For each journal the year-wise download counts (full-text article requests - FTAs) for articles from the particular year down to articles from 1995 (if the volumes from that year were already accessible online) were given. At article level, we were provided with the monthly downloads for five journals (“Biochemical Pharmacology”, “Drug Discovery Today”, “European Journal of Pharmacology”, “Life Sciences” and “Pain”) between January 2001 and December 2006. Thus, it is possible to investigate the relation between citations and downloads at article level and to perform obsolescence analyses on a monthly basis. Furthermore, we can examine if citations have an effect on downloads. The necessary citation data will be retrieved from the Web of Science (WOS).

### Preliminary results

In this section we present the results of a few selected analyses we have already conducted. Table 1 shows the increase in e-journal use in Science Direct (SD), subject category “Pharmacology, Toxicology and Pharmaceutical Science”, between 2001 and 2006. As can be seen, the number of downloads has increased more than fivefold in this period. In particular, the growth rate was very high in 2003 (+95%). This increase was not caused by the higher number of SD journals (86 in 2003) but might be mainly due to the general risen acceptance of e-journals at that time.

**Table 1. Growth of downloaded articles (from 1995 onwards) from pharmacology journals in Science Direct (2001-2006).**

	2001	2002	2003	2004	2005	2006
Total no of FTAs	4,582,661	6,449,412	12,593,727	17,960,640	21,542,223	25,712,253
Annual increase		41%	95%	43%	20%	19%
SD journals (n)	68	82	86	96	100	102

### Analyses at journal level

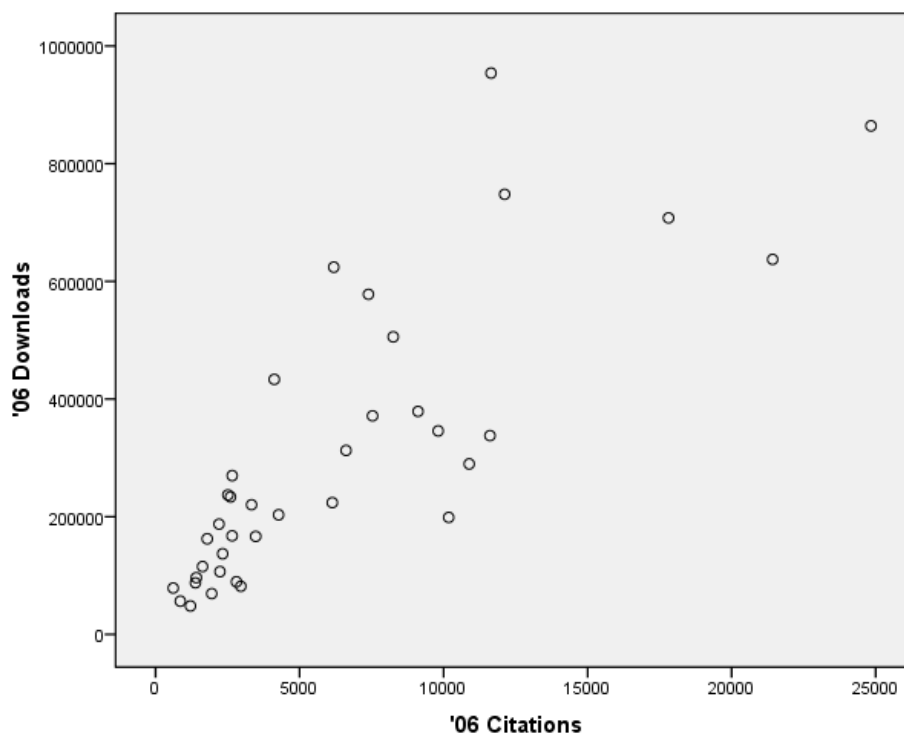
**Table 2. Obsolescence of pharmacology journals (with 10 years backfile data) in SD (2006 FTAs) (n=56).**

	All	2006	2005	2004	2003	2002	2001	2000	<2000
2006 FTAs to articles from particular years (in 1000)	23480	7185	4294	2799	2059	1639	1329	1111	3064
Relative (in %)	100%	31%	18%	12%	9%	7%	6%	5%	13%

Table 2 lists the obsolescence data for those 56 SD pharmacological journals for which access data were available for at least 10 years in 2006. The age distribution exhibits that, unlike citations, downloads relate primarily to recent articles: More than 30% of the articles downloaded in 2006 were published in the same year, half of the downloaded articles were not older than two years. Accordingly, we computed a median usage half-life of 1.9 years and an aggregate usage half-year of 2.1 (n=56). By contrast, the aggregate cited half-life of journals from the JCR, subject category “pharmacology & pharmacy”, is 6.1. Though this comparison is problematic to some degree since citations go back much further (our SD download data started with articles from 1995 only), we do not expect the usage half-life to increase considerable even if online data are available for longer time periods.

Though download counts are one magnitude higher than citation numbers (we computed an average ratio of 1:49 for those journals which are covered by both JCR and SD in 2006), there

is a high relation between those two indicators (see Figure 1 and Table 3). Table 3 displays the rank correlations (Spearman-Rho). Because of different obsolescence characteristics (see above), we related the downloads of a particular year also to the citations of subsequent years. The correlation between the impact factor and the usage impact factor is slightly lower (for instance, 0.66 in 2006, 0.81 in 2005 and 0.70 in 2004).



**Figure 1. Citation frequencies versus download frequencies in 2006 (n=37).**

**Table 3. Rank correlation (Spearman-Rho) between citations and downloads in various years.**

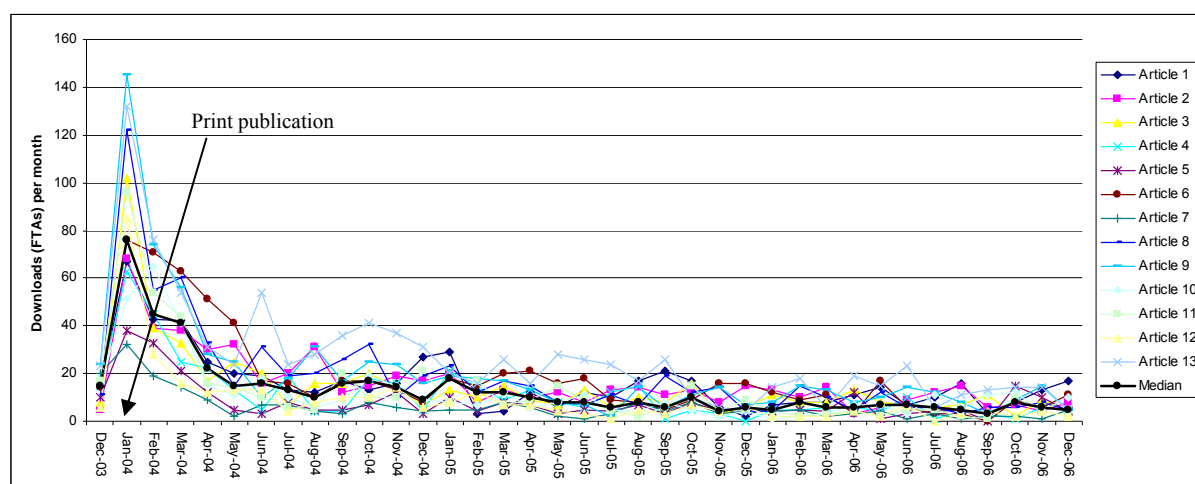
	'06 citations	'05 citations	'04 citations	'03 citations
'03 downloads	0.86 (n=35)	0.84 (n=35)	0.79 (n=36)	0.79 (n=35)
'04 downloads	0.89 (n=35)	0.85 (n=35)	0.80 (n=36)	
'05 downloads	0.87 (n=37)	0.84 (n=37)		
'06 downloads	0.86 (n=37)			

#### *Analyses at article level*

The download patterns on a paper-by-paper basis for the five considered journals (for an example, see Figure 2) show the following characteristics:

- in recent years (2005 and 2006) articles have the highest download requests immediately after they are available electronically; furthermore, as is the case with the journal in Figure 2, journals are put online and downloaded often before print publication;
- the age distribution is different for older volumes: download curves do not decline so quickly, some time is required in order to reach the download peak.

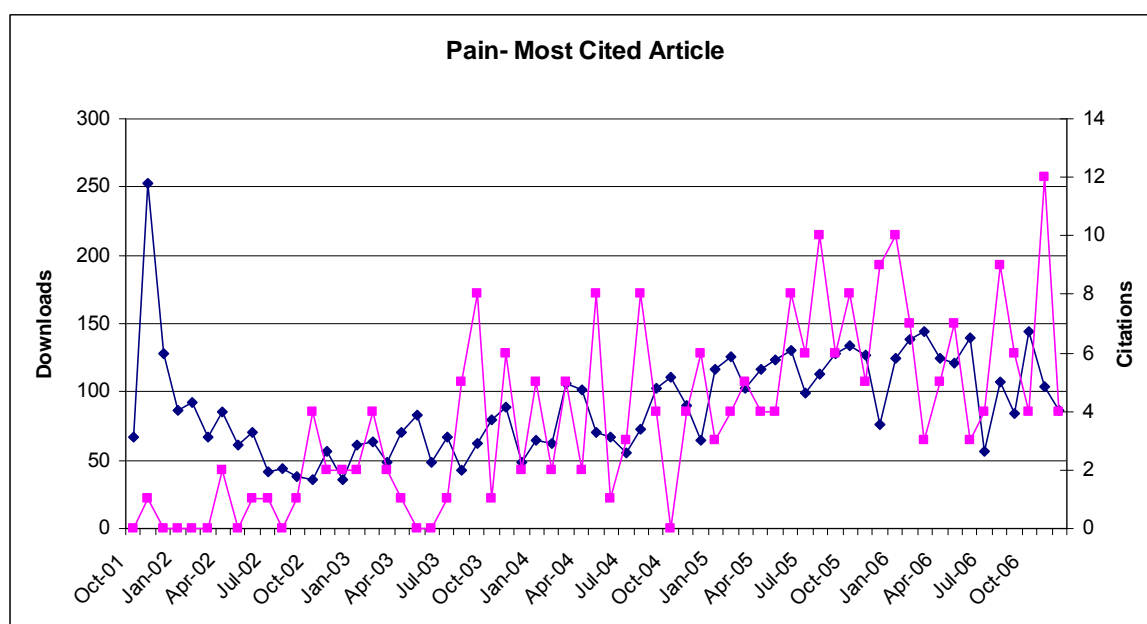
These results are in perfect agreement with our previous outcomes for the oncology journals and corroborate our assumption that there is some distortion in download obsolescence due to the strong increase in e-journal use between 2001 and 2004 (see Table 1) and other factors as suggested by Parker (1982).



**Figure 2. Monthly downloads of articles in vol. 484, iss. 1 (cover date: Jan. 1, 2004) of “European Journal of Pharmacology”.**

In order to get a first idea if citations have a direct influence on downloads, we opposed the monthly downloads to the monthly citations for a small sample of articles. The citation data were retrieved in Web of Science, hence the monthly assignment of the citations was not always possible because the publication date (month) was not specified for approximately 10% of all citing articles in this database. Another problem was that most of the articles were hardly cited on a monthly basis so that a well-founded analysis did not make sense.

The results of our first analyses suggest that citations can indeed have an effect on downloads for more highly cited papers and that a substantial number of citations can provoke a considerable increase of article requests even several years after their publication (see Figure 3 for an example). For the most cited articles we may even confirm the argument by Moed (2005) according to which downloads increase considerably one or two months after citation. The suggested growth of 25 percent might, however, be very dependent on the number of citations and the number of downloads for each article.



**Figure 3. Downloads vs. citations of most cited article in “Pain” (by Farrar, J.T. et al., published in vol. 94, iss. 2, cover date: Oct. 1, 2001, online date: Oct. 25, 2001).**

### Future research / prospects

Besides the analyses we have presented above and besides the topics we have already identified for further research, we want to investigate the following issues:

- Since our analyses have shown that most articles are downloaded in the year of publication, we agree with Wan et al. (2008) that the usage immediacy index might be a good predictor for later citation impact. In particular we want to investigate the relation between the usage immediacy index and the impact factor and h-index.
- We suppose that the availability of e-journals might have a medium and long-term influence on the citation indicators, especially on the immediacy index and the impact factor. As a consequence, we plan to investigate a possible change of these two indicators for journals included in SD and JCR in recent years.
- At article level, we want to continue in investigating comprehensively the relation between citations and downloads.
- And finally, since we have conducted many of the analyses presented in this contribution already for oncology journals, we want to explore possible differences in the usage characteristics between pharmacology and oncology journals.

### Acknowledgments

The authors would like to thank Mr. Niels Weertman from Science Direct for providing us with the necessary usage data.

### 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*. Retrieved January 17, 2009 from: <http://public.lanl.gov/herbertv/papers/ipm05jb-final.pdf>
- Brody, T.; Harnad, S. & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060-1072.
- Chu, H. & Krichel, T. (2007). Downloads vs. citations in economics: Relationships, contributing factors & beyond. In D. Torres-Salinas & H.F. Moed (Eds.), *Proceedings of the 11th International Society for Scientometrics and Informetrics Conference* (pp. 207-215). Madrid: CINDOC.
- Elsevier (n.d.). *ScienceDirect brochure*. Retrieved January 17, 2009 from [http://www.sciencedirect.info/about/SDPlatformBrochure\\_06.pdf](http://www.sciencedirect.info/about/SDPlatformBrochure_06.pdf)
- Kurtz, M.J.; Eichhorn, G.; Accomazzi, A.; Grant, C., Demleitner, M.; Murray, S.S.; Martimbeau, N. & Elwell, B. (2005). The bibliometric properties of article readership information. *Journal of the American Society for Information Science and Technology*, 56(2), 111-128.
- McDonald, J.D. (2007). Understanding journal usage: A statistical analysis of citation and use. *Journal of the American Society for Information Science and Technology*, 58(1), 39-50.
- Moed, H.F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. *Journal of the American Society for Information Science and Technology*, 56(10), 1088-1097.
- Parker, R.H. (1982). Bibliometric models for management of an information store. II. Use as a function of age of material. *Journal of the American Society for Information Science*, 33, 129-133.
- Rowlands, I. & Nicholas, D. (2006). The missing link: journal usage metrics. *Aslib Proceedings*, 59(3), 222-228.
- Wan, J.-K.; Hua, P.-H.; Rousseau, R. & Sun, X.-K. (2008). The download immediacy index (DII): Experiences using the CNKI full-text database, in press.