

ILDS 35.1 Bador et al rev 14.10.06

Cartographic analysis of the correlation between document supply and citations in pharmacology: a case study from INIST in France (part 2)

*Pascal Bador
Chérifa Boukacem-Zeghmouri
Thierry Lafouge
Hélène Prost
Joachim Schöpfel*

Abstract

Purpose: The article investigates the correlation between citations and the document supply of print serials.

Approach: The study is based on data from INIST-CNRS for document supply requests and from ISI for citations from 89 serials with JCR impact factors in pharmacology. Data were collected from 1992 to 2004.

Findings: The results distinguish four groups of serials with different relationships between document supply requests and citations. The characteristics of the serials of the four groups are described (year of creation, price, etc). The evolution of the correlation between document supply and citations is analysed.

Originality: This is the second of three articles on a longitudinal study over 13 years on different aspects of the relationship between document supply and impact factor (citations) in pharmacology, a scientific domain with a high use of scientific information.

Keywords

Document supply, Impact Factor, Citations, Serials, Electronic Resources, Longitudinal Study, INIST

Paper type: Case study

1. Introduction and issues

This paper is the follow-up of our study of the trends in document supply requests from print pharmaceutical journals in a digital environment. In our first article (Boukacem-Zeghmouri et al. 2006), we studied a sample of pharmaceutical serials supplied by INIST-CNRS and observed an overall downward trend in print document supply by INIST beginning in 1999. Four groups of serial titles could be distinguished each with a different evolution of document supply requests.

In our second article, we will focus on the link between document supply requests recorded at INIST for the serials sample¹ and the journal citations in the sample taking into account the digital environment and the online use of these journals. The objective of this paper is to confirm the hypothesis that document supply requests for print documents no longer correspond to the realities of the research activities of researchers and therefore can no longer be an index for journal citation (see Bador et al. 2003). Nowadays, it would seem that journals are essentially read in their digital version.

Then we will show the break in the link between document supply and citations in the context of electronic journals. The originality of our work resides in showing this break through empirical data.

2. The sample

We used the same serials sample as in our first paper, i.e. 95 journals listed under the “Pharmacology and Pharmacy” heading in *Journal Citation Reports (JCR)* published by the Institute for Scientific Information (ISI). We added one year to the initial study period of 12 years to cover 1992 to 2004. During this period, the number of journals went from 136 titles in 1992 to 187 titles in 2004 (see Figure 1).

Year	92	93	94	95	96	97	98	99	00	01	02	03	04	Average
Number of journals	136	142	146	149	152	156	176	175	181	186	187	185	187	166

Supprimé : Insert Figure 1

Figure 1: Evolution of serials titles under the “Pharmacology and Pharmacy” heading in ISI’s *JCR*.

Some journals were excluded from the sample because they were either included in *JCR* only after 1992 or disappeared from *JCR* before 2004. We further analyzed the following three periods:

(1) 1993-1994

New serial titles in *JCR*: 17 including 5 titles launched between 1986 and 1989, 11 titles launched after 1990 (5 following title merger or change) and one title launched in 1973.

Serial titles dropped from *JCR*: 4 discontinued titles including 3 title mergers/changes.

(2) 1997-1999

New serial titles in *JCR*: 39 including 23 titles launched after 1990 (12 title mergers/changes), 12 titles launched after 1995 (7 title mergers/changes).

¹ We studied document supply by photocopies, fax or ARIEL from the INIST print holdings. This means that all journals from our sample were available in print format during the whole period 1992-2004.

Serial titles dropped from JCR: 8 discontinued titles including 6 title mergers/changes.

(3) 2002

New serial titles in JCR: 5 titles launched in 1986, 1994 and 2001 including 2 title mergers/changes.

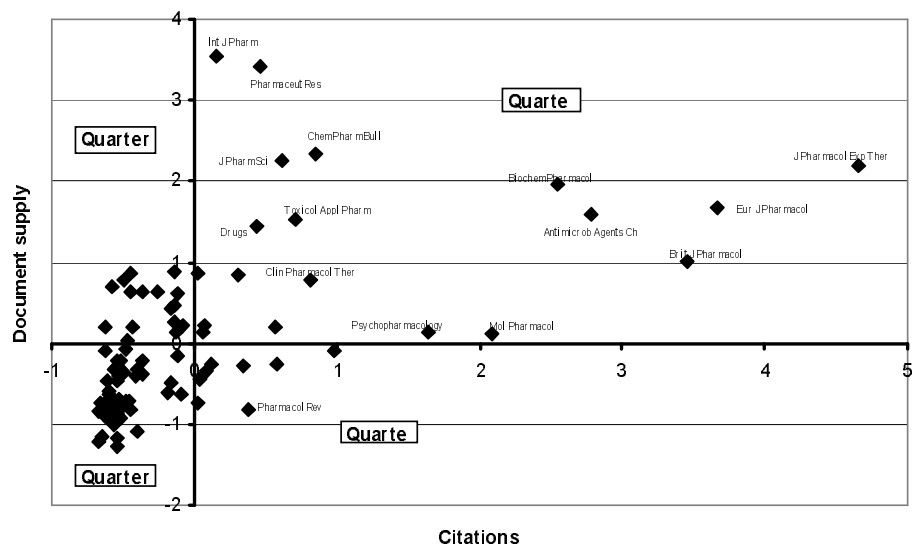
These three periods representing 6 out of the 13 years of our study, show that a large number of serials (79 titles) were excluded from the sample because they either disappeared or appeared during the study period. 36 titles out of 79 were either newly launched or discontinued; others appeared in or disappeared from JCR for reasons that only ISI could explain.

Therefore, our sample finally comprises 89 serial titles representing 54% of the average number of serial titles (166) for the 13-year period covered by our study.

3. Matching citations and document supply

In our first article, we processed raw data. In this study, we proceeded otherwise. First we transformed the citation and document supply values into standard scores (z-scores)², a common statistical tool when comparing two distributions with different value scales.

The mean of the standardized document supply and citation values is 0, their standard deviation is 1. In this way, we obtained for each year a “dot cloud” of journals centred at 0.0 and divided in four groups or “quarters”. Figure 2 shows the 1999 map.



² “A dimensionless quantity derived by subtracting the population mean from an individual (raw) score and then dividing the difference by the population standard deviation” (Wikipedia). In a normal distribution, about 68% of the scores are between -1 and +1.

Figure 2: Standard scores for document supply (requests) and citations (1999)

Each dot on the map corresponds to a specific journal, characterised by its standard scores for document supply and citations. To illustrate the result of the mapping, some journal titles are included in the map.

Some explanations may be necessary to understand the map and its reduced central values (dots). For instance, in 1999 INIST customers ordered 832 copies from *The Journal of Pharmacology and Experimental Therapeutics*, and it has been cited by 32,134 articles (ISI). In 1999, the average number of copy orders for the whole sample is 308.71 per journal (INIST), and the average number of citations 4,235.11 (ISI). We then calculated the standard deviation which is the most important measure of dispersion in statistics and allows to quantify the repartition of individual values in relationship with the average. The standard deviation for both values of *The Journal of Pharmacology and Experimental Therapeutics* in 1999 are 239.54 (copy orders) and 5,982.99 (citations). The high standard deviation for both values indicates that they are highly dispersed (distant) from the average values.

The Z or standard scores (reduced central values, see Carpenter & Vasu 1978) are directly derived from the standard deviation:

Axe x (citations 1999): $(32,134 - 4,225.11) / 5,982.19 = 4.66$.

Axe y (copy orders 1999): $(832 - 308.71) / 239.54 = 2.18$.

In statistics, the reduced central values have an average score of 0 and a standard deviation of 1. For our example, *The Journal of Pharmacology and Experimental Therapeutics*, this means that this journal figures in the group with highly cited journals and a copy order number above the average.

This statistical technique is often applied to classify objects characterised by two different numerical values into four groups around an average value. For instance, in scientometrics a map like Figure 2 is called a “strategic diagram”.

4. Comparison and typology of the journals

In a second step, we compared the 13 maps for each year. Since these maps were not easy to read, our approach was to focus on the stability of the journals in the quarters. To do so, we had to define comparison indicators:

(1) The first indicator was to identify in which quarter a journal most often appeared (this indicator is equivalent to the median). We did not assign an inclusion threshold to a quarter. The only parameter taken into account was each journal’s score (“dot”).

(2) The second indicator was to count the number of times a journal changed quarters during the 13 years studied.

This method led us to propose a classification of the serial titles into four groups as shown below:

- QUARTER 1 SERIALS: 19 titles most often found over 13 years in the upper right hand corner of the maps, i.e. those serial titles accounting for the highest number of document supply requests and the highest number of citations.
- QUARTER 2 SERIALS: 16 titles most often found over 13 years in the upper left hand corner of the maps, i.e. those serial titles accounting for the highest number of document supply requests but a lesser number of citations.
- QUARTER 3 SERIALS: 47 titles most often found over 13 years in the lower left hand corner of the maps, i.e. those serial titles accounting for the lowest number of document supply requests and the lowest number of citations.
- QUARTER 4 SERIALS: 7 titles most often found over 13 years in the lower right hand corner of the maps, i.e. those serial titles accounting for the highest number of citations but a lesser number of document supply requests.

The complete title lists for each group (quarter) are presented in the appendix.

To confirm the representative character of this clustering, we performed a second processing taking into account the average number of citations in *JCR* and the average number of document supply requests recorded at INIST between 1992 and 2004. Therefore, we transformed for each serial title the document supply request and citation averages over the 13 years into standard scores.

Thus, we obtained a new map distributing the 89 serial titles into four quarters (see Figure 3). We observed that, out of the highest quarters, only three journals showed up in different quarters than those of the preceding processing: *Pharmacology & Therapeutics* and *Drug Metabolism and Disposition* moved from Quarter 2 (highest) to Quarter 1 (average); the *Journal of Natural Products* moved from Quarter 3 (highest) to Quarter 4 (average). This somewhat surprising result showed that the data was very stable and confirmed our clustering.

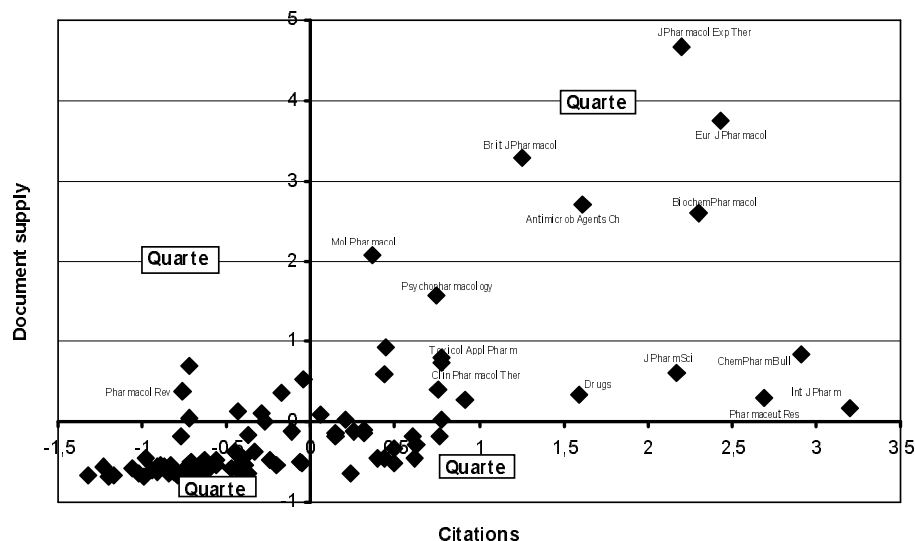


Figure 3: Standard scores for document supply (requests) and citations (averages 1992-2004)

Because of the stability and consistency of the two results, we analyzed the characteristics of the titles in the four groups obtained by the first processing.

4.1. First quarter: most cited, most requested

The 19 serials in this first group (21% of the sample) are the titles which over 13 years (1992-2004) are located for the most part in Quarter 1 i.e. the quarter for the titles most often requested and most often cited. Twelve of these titles (63%) never left this quarter.

This serials group most important for researchers in the pharmaceutical field therefore shows a strong stability over the period studied.

Furthermore, Quarter 1 serials on average account for the highest number of document supply requests at INIST (546 article requests per year) and the highest number of citations (11296 citations or over twice as much as the number of citations for Quarter 4 serials). Their average subscription price (€2313 according to the 2005 Ulrich catalogue) is higher than the prices for the other 3 groups. This suggests that a large number of research laboratories prefer to order articles from INIST rather than subscribe to these titles.

Another trait of the Quarter 1 serials is that they are the oldest titles, with 1963 as the average date of creation and *The Journal of Pharmacology and Experimental Therapeutics* the oldest title first published in 1909. As a corollary, these journals have been known and renowned for a great many years for their scientific quality, *a priori*

without title mergers or title changes. This long-term reputation partly explains, the high number of document supply requests and citations for these titles.

Quarter 1 serials are therefore the titles most often requested, most often cited, the most expensive and the oldest.

4.2. Second quarter: less cited, most requested

The 16 serials in this group (18% of the sample) are titles that over 13 years showed up most often in Quarter 2, the quarter for the most requested, least cited titles. These titles are less stable than those in Quarter 1 since only 4 of them (25%) never left Quarter 2, 5 titles (31%) left once, 2 titles (13,5%) left twice, 3 titles (19%) left 3 times, 1 title (6%) left 4 times and 1 title (6%) left 5 times. Even though their number of document supply requests at INIST is above average, they are less requested than Quarter 1 journals.

The average number of citations for Group 2 serials is below that of our sample. Their average subscription price (€1384) puts them in second position behind Quarter 1 serials. Therefore, there is a stronger correlation between the subscription price of the serial titles in our sample and the number of article requests at INIST than between the number of articles requested and the number of citations since the articles most often requested come from serial titles which on average are the most expensive.

The average creation date of Group 2 serials is 1971, which puts them in third position as far as title age is concerned (oldest titles: *Pharmazie* and *Therapie* first published in 1946).

4.3. Third quarter: less cited, less requested

This is the largest group with 47 serial titles i.e. 53% of our sample. These titles are by far the least requested (4 times less than Quadrant 1) and the least cited (11 times less than Quadrant 1). The serials in this group are rather stable over the period studied since 34 titles (72%) never left this group. Their average subscription price is €1 050 (most expensive titles *Drug Development Research* at €3900 and *Journal of Labelled Compounds and Radiopharmaceuticals* at € 3 100; least expensive titles *Yakugaku Zasshi* and *American Journal of Pharmaceutical Education* at € 80). On average, these titles are the least expensive along with Quarter 4 titles.

Furthermore, in this Quarter, the average creation year is 1977, which makes this serials group the most recent in our sample. Finally, Quarter 3 serials are not as highly rated by researchers probably because they are relatively recent.

4.4. Fourth quarter: most cited, less requested

Group 4 is the smallest group with only 7 serial titles (8% of the sample). Only one title never left the group while the six other titles left between one and four times. The number of document supply requests for these titles is below average (average 173 requests per year) but their number of citations is above average. Their €972 average subscription

price is the lowest of the groups. Therefore, it is in Quarter 3 and 4 that we find the titles that are least requested at INIST and cost the least. Their prices make them more affordable for researchers who are more liable to subscribe to these titles and therefore less likely to send document supply requests to INIST for articles from these title.

The average year of creation of Quarter 4 serials is 1968 (*Pharmacological Reviews* created in 1949) which makes these titles on average the oldest after those in Quarter 1. Therefore, we observed that Quarter 1 and Quarter 4 titles are the oldest and most often cited, that their average date of creation probably has an impact on the number of citations and therefore an impact on the popularity and the use of these titles in terms of citations.

4.5. Other characteristics

Based on 2005 data from Bowker's Ulrichsweb and the INIST catalogue, the sample and quarters can be described in more details.

Peer review: All titles are refereed journals.

Indexing in databases: All titles are referenced by at least three out of a selection of significant databases for pharmacology and pharmacy (abstracting and indexing services). 29 titles are indexed by all of these databases. Some differences can be observed between the databases:

Database	N indexed titles	% of sample
Chemical Abstracts	87	98%
Current Contents	87	98%
Embase	87	98%
Medline	86	97%
BIOSIS	85	96%
PASCAL	52	54%
International Pharmaceutical Abstracts	48	54%

Figure 4: Indexing in databases for the whole sample (89 titles)

Furthermore, the titles of Quarter 1 and 2 are clearly more often referenced in these databases than those of Quarter 3 and 4 (91% and 94% vs. 81% and 82%).

Annual number of issues: The journals of Quarter 1, 2 and 4 are more often published than those of Quarter 3:

Annual number of issues	Quarter 1	Quarter 2	Quarter 3	Quarter 4
> 12 issues per year	32%	25%	4%	14%
12 issues per year	68%	50%	30%	57%
< 12 issues per year	0%	25%	66%	29%

Figure 5: Annual number of issues in the four groups (in %)

While all journals of Quarter 1 are edited at least monthly, two thirds of the Quarter 3 are edited a maximum 10 times per year. This may explain some of the differences between the number of citations and document supply.

Online edition: All titles are published online, either by the publisher or by another vendor. In 2005, backfiles were available from at least 1999 for 82% of the titles. For two Elsevier journals, *Biochemical Pharmacology* and *Toxicology and Applied Pharmacology*, backfiles go back to 1958 and 1959. Nevertheless, for most of the titles one cannot determine clearly the year of the first online edition.

Subject: A comparison of the subject indexing by Bowker (Dewey and Ulrich's classification) and INIST (catalogue, PASCAL database) does not reveal any significant differences between the four groups. 72% of the sample are indexed as belonging to the Dewey code 615 (pharmacology and therapeutics).

Publishers: The 89 titles are edited by 35 publishers located in 13 countries (38% US, 21% UK, 11% Germany, 9% Netherlands). 86 titles are edited in English.

Matched against the Ulrich's Core Collection, our sample represents 27% of the refereed and abstracted/indexed journals in the domain of pharmacy and pharmacology with a digital version (online edition) and an ISI impact factor.

5. Global analysis

Once we had confirmed and explained our clustering, we approached this evolution from a broader perspective through document supply requests and citations.

5.1. Evolution of the document supply requests

The evolution of the average number of document supply requests by quarter and by year is shown in Figure 6 below. The curve shows that overall the evolution of all the quarters is similar and we can see three distinct evolution phases common to all quarters:

- First, a sharp increase between 1992 and 1995,
- Then, some kind of plateau between 1995 and 1999,

- Finally, a decline between 1999 and 2004.

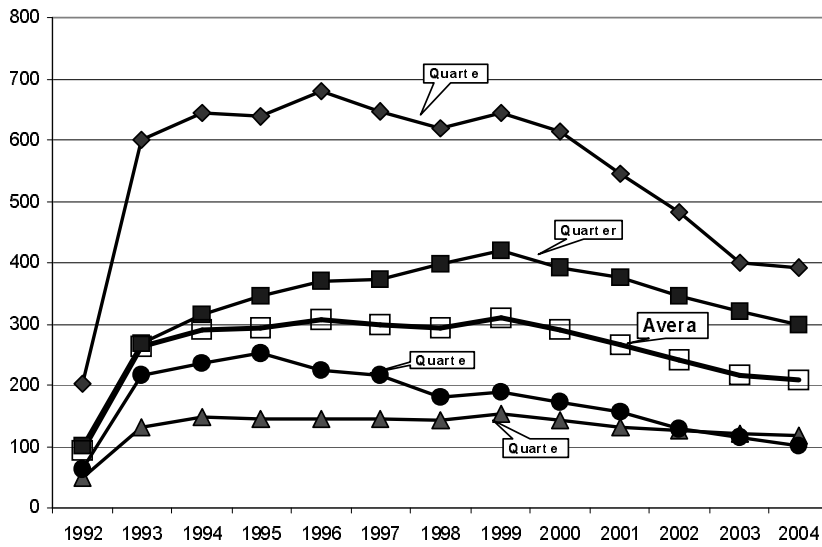


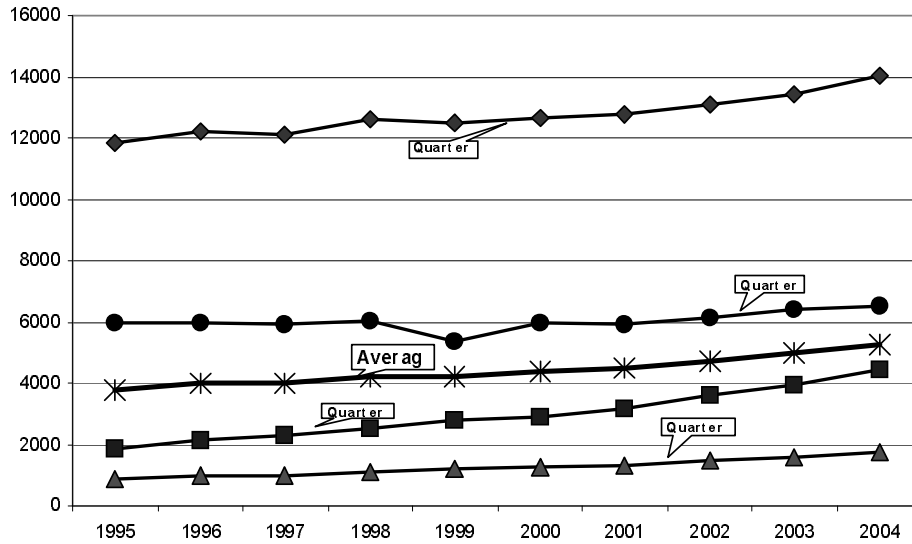
Figure 6: Evolution of document supply in the four quarters (1992-2004, requests in 1 000s)

5.2. Evolution of the number of citations

To show the evolution of the number of citations in our sample, we calculated the citation average per journal according to our clustering.

Figure 7 below shows the evolution of the average number of citations per each group and per year

Figure 7: Evolution of citations for the four quarters 1992-2004 (average number of citations per journal)



The curves in Figure 7 show a regular and significant evolution over the period considered. Quarter 1 is well above the other quarters with a large number of citations but stays within the overall increasing trend of the sample.

5.3. The relation between document supply and citations

One of the most significant points of our demonstration was the measurement of the correlation between the document supply request and the citation.

Figure 8 shows an overall decrease in the correlation between the number of document supply requests placed at INIST and the number of citations in *JCR* between 1992 and 2004. Indeed, we observed an overall increase in the average number of requests until 1998 followed by a regular decline until 2004 while overall, the average number of citations showed a steady increase over the 1992-2004 period.

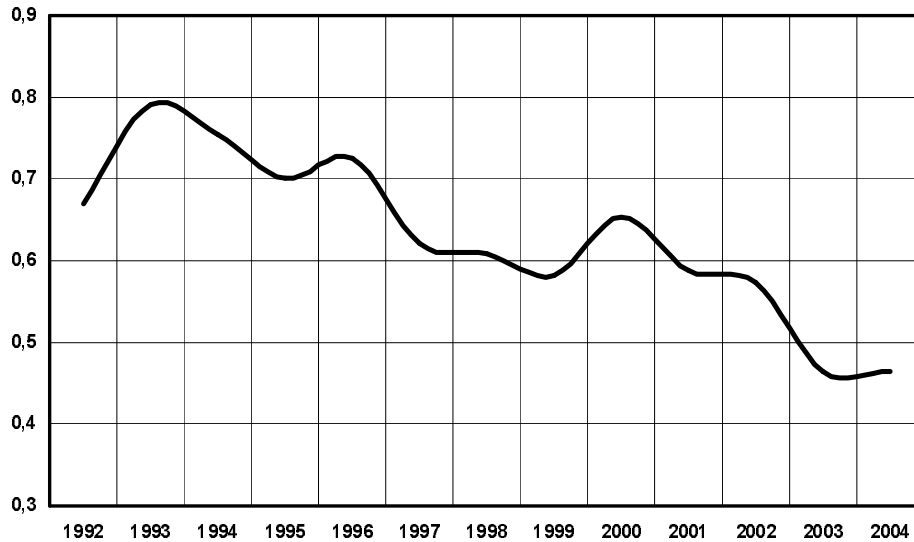


Figure 8: Evolution of the correlation between document supply and citations (1992-2004)

A closer analysis by quarters showed that Quarter 1 serials, i.e. titles generating the highest numbers of document supply requests and of citations, presented a slight increase in the number of document supply requests between 1993 and 1998 followed by a sharp decline between 1999 and 2004 while their number of citations showed a regular and strong increase.

All serials with the highest number of document supply requests (Quarter 1 and 2) showed overall a rather strong increase in the number of document supply requests between 1993 and 1998 followed by a rapid decline between 1999 and 2004 that can be explained by the boom in electronic access to serial titles in general and the titles in our sample in particular. Concurrently, the number of citations shows a moderate to rapid increase between 1993 and 1998 followed by a rapid increase between 1999 and 2004.

Titles with the lowest numbers of document supply requests (Quarter 3 and 4) showed respectively a stable level to mild decline in the number of requests between 1993 and 1998 followed by a moderate to high decline between 1999 and 2004 that can be accounted for by the advent of electronic versions of the titles. Concurrently, those titles showed overall a rather weak increase in their number of citations between 1992 and 2004.

Titles with the highest number of citations (Quarter 1 and 4) showed respectively an almost stable level of document supply requests between 1993 and 1998 followed by a rapid decline between 1999 and 2004.

Titles with the lowest number of citations (Quarter 2 and 3) showed overall a moderate decline in the number of document supply requests between 1993 and 1998 followed by a

major decline between 1999 and 2004. Conversely, their number of citations showed a regular increase between 1992 and 2004.

Two former studies reported a correlation index of .63 and .64 (Salaün et al. 2003, Bador et al. 2005). Obviously, this correlation between document supply and citations tends to become less significant.

6. Methodological limitations

By observing the four quarters, we can see that 47 serial titles of the sample never changed Quarters during the period studied. This lack of “movement” is significant since it shows a certain amount of stability in the sample.

What first comes to mind is that INIST’s collections are partly put together on the basis of the public they serve which is composed of public research organisations, higher education, and the corporate sector. This heterogeneity contrasts with the stability we observed in our sample.

However, we must draw attention to the fact that our sampling was biased by the limitation to journals with an Impact Factor, which means that all selected titles were cited several times. Thus, we introduced “artificial” homogeneity or stability into the sample. Therefore, it is no wonder that our sample did not reveal many differences.

7. Conclusion

With the boom in electronic publishing, searching and using electronic resources has become a fact of life in the scientific community. However, the reality is not uniform across all disciplines, even, as the work we report in this article shows, inside a given domain such as pharmacology.

We saw also that to understand the break of the link between citation and document supply request we must take into account elements related to the way in which scientific communication processes operate and particularly:

- The status of journals rooted in a discipline which due to their seniority in the field have historically contributed to the shaping of a discipline and which continue to do so.
- The guarantee of quality that they represent for researchers who see them as a “must” when they are looking for information.
- The steady identity of these journals despite the mergers which have been rampant in scientific publishing for over 30 years and despite the continuing inflation of journal prices.

In fact, the analysis by quarters which we proposed in this article and which stresses the impact of electronic documents on the requests-citations relationship, also shows that this impact seems limited or differed for serials deeply anchored in a discipline. -

The third and last article will focus on the characteristics of the requests (customers) in order to give empirical evidence of the underlying dynamics of the evolution of document supply.

Bibliography

- Bador, P., Ben Romdhane, M. and Lafouge, T. (2003), "European pharmaceutical journals: relationship between demand and indexation: the example of the main French document supplier", *The Canadian journal of information and library science*, Vol 27, pp. 33 – 55
- Boukacem-Zeghmouri, C., Bador, P., Lafouge, T., Prost, H. & Schöpfel, J. (2006): "Analysis of the downward trend in document delivery in pharmacology: a case study from INIST in France (part 1)". *Interlending & Document Supply*, Vol 34 No. 4.
- Carpenter, R.L. & Vasu, E.S. (1978): *Statistical methods for librarians*. Chapter 1. Chicago: American Library Association.
- Salaün, J. M., Lafouge, T. and Boukacem, C. (2000), "How far does demand for scientific articles reflect development in electronic publishing? An example from the Institut de l'Information Scientifique et Technique (France)", *Scientometrics*, Vol 47 No 3, pp. 561-588

APPENDIX

Quarter 1: 19 journals		
ANTIMICROB AGENTS CH BIOCHEM PHARMACOL BRIT J CLIN PHARMACO BRIT J PHARMACOL CHEM PHARM BULL CLIN PHARMACOL THER DRUGS	EUR J PHARMACOL INT J PHARM J ANTIMICROB CHEMOTH J CARDIOVASC PHARM J PHARM PHARMACOL J PHARM SCI J PHARMACOL EXP THER	MOL PHARMACOL PHARMACEUT RES PHARMACOL BIOCHEM BE PSYCHOPHARMACOLOGY TOXICOL APPL PHARM
Quarter 2: 16 journals		
ALIMENT PHARM THERAP ARZNEIMITTEL-FORSCH CLIN PHARMACOKINET CLIN THER DRUG DEV IND PHARM DRUG METAB DISPOS	J CLIN PHARMACOL J CLIN PSYCHOPHARM J CONTROL RELEASE J PHARMACEUT BIOMED PHARMACOL THERAPEUT PHARMACOTHERAPY	PHARMAZIE PLANTA MED THERAPIE TOXICOLOGY
Quarter 3: 47 journals		
ALCOHOL AM J PHARM EDUC ARCH PHARM BIOMED CHROMATOGR BIOMED PHARMACOTHER BIOPHARM DRUG DISPOS CANCER CHEMOTH PHARM CARDIOVASC DRUG REV CARDIOVASC DRUG THER CHEMOTHERAPY CHIRALITY CLIN EXP PHARMACOL P CLIN NEUROPHARMACOL CRIT REV THER DRUG CURR THER RES CLIN E DRUG CHEM TOXICOL	DRUG DEVELOP RES DRUG EXP CLIN RES DRUG METAB REV DRUG SAFETY EUR J DRUG METAB PH FARMACO FUNDAM CLIN PHARM IMMUNOPHARM IMMUNOT INT CLIN PSYCHOPHARM INT J CLIN PHARM TH INVEST NEW DRUG J CHEMOTHERAPY J CLIN PHARM THER J ETHNOPHARMACOL J INT MED RES J LABELLED COMPD RAD	J NAT PROD J VET PHARMACOL THER METHOD FIND EXP CLIN NEUROPSYCHHOPHARMAC OL PHARMACOL RES PHARMACOLOGY PHARMACOPSYCHIATRY PHYTOTHER RES PROG NEURO-PSYCHOPH REGUL TOXICOL PHARM REV PHYSIOL BIOCH P THER DRUG MONIT TOXICON XENOBIOTICA YAKUGAKU ZASSHI
Quarter 4: 7 journals		
ANNU REV PHARMACOLOGOL CAN J PHYSIOL PHARM J ANTIBIOT	NEUROPHARMACOLOGY N-S ARCH PHARMACOL PHARMACOL REV	TRENDS PHARMACOL SCI

The authors

Pascal Bador is a Lecturer at Université Claude Bernard Lyon 1 and member of the URSIDOC research group.

Pascal.Bador@univ-lyon1.fr

ILDS 35.1 Bador et al rev 14.10.06

Chérifa Boukacem-Zeghmouri is a Lecturer at the Université Charles de Gaulle, Lille 3, UFR IDIST.
boukacemc@yahoo.fr

Thierry Lafouge is Professor at Université Claude Bernard Lyon 1 and a member of the URSIDOC research group.
Thierry.Lafouge@univ-lyon1.fr

Hélène Prost is a librarian at INIST-CNRS and specialises in the evaluation of Collections and Document Supply.
prost@inist.fr

Joachim Schöpfel is Head of the E-publishing and Document Supply Department at INIST-CNRS and Lecturer on Scientific Information at the Université Nancy 2.
schopfel@inist.fr