

By Philip M. Davis, Ph.D. Phil Davis Consulting, 432 Mitchell St., Ithaca NY 14850.

pmd8@cornell.edu | 607 272-5461 | <http://phil-davis.org>

November, 25, 2013

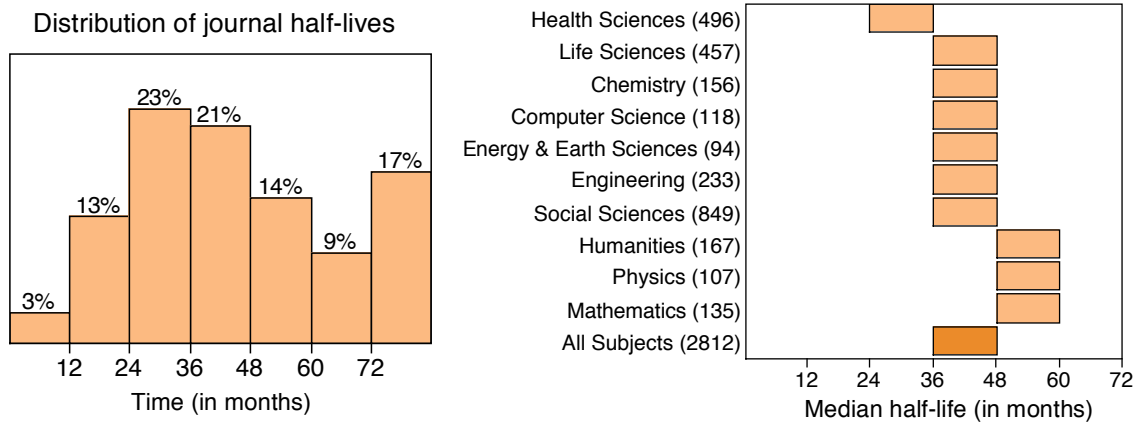
Abstract

An analysis of article downloads from 2,812 academic and professional journals published by 13 presses in the sciences, social sciences, and the humanities reveals extensive usage of articles years after publication.

Measuring usage half-life—the median age of articles downloaded from a publisher’s website—just 3% of journals had half-lives shorter than 12-months.

While journal usage half-lives were typically shorter for journals in the Health Sciences (median half-life: 25-36 months), they were considerably longer for journals in the Humanities, Physics and Mathematics (median half-life: 49-60 months). Nearly 17% (475) of all journals had usage half-lives exceeding six years.

This study illustrates substantial variation in the usage half-lives of journals both within and across subject disciplines.



Data and Methods

The dataset consists of 2,812 journals published by four commercial, two university, and seven learned society/association presses (Table 1).

Due to the number of journals published by the three largest commercial publishers (Elsevier, Springer, and Wiley), 50% of their journals were randomly selected for inclusion in the dataset. As a result, no one publisher contributed more than 25% of journals to the study. Each journal was given equal weight in the analysis.

Table 1. Participating Journal Publishers

Publisher	Journals	%
American Chemical Society	38	1.4%
AIP Publishing	9	0.3%
American Mathematical Society	4	0.1%
American Physical Society	7	0.3%
American Physiological Society	11	0.4%
American Psychological Association	62	2.2%
Elsevier	688	24.5%
Emerald	232	8.3%
IEEE	137	4.9%
Oxford University Press	259	9.2%
Springer Science+Business Media	641	22.8%
University of Chicago Press	37	1.3%
Wiley	687	24.4%
Total	2812	100.0%

Subject classification. Journals were classified by the publisher into one of ten primary subject groups (Table 2). Every subject group in the dataset contained representation from at least four publishers.

Table 2. Participating Journals by Primary Subject Classification

Primary Subject	Journals	%
Chemistry	156	5.6%
Computer Science	118	4.2%
Energy & Earth Sciences	94	3.3%
Engineering	233	8.3%
Health Sciences (incl. Medicine)	496	17.6%
Humanities	167	5.9%
Life Sciences	457	16.3%
Mathematics	135	4.8%
Physics	107	3.8%
Social Sciences (incl. Economics)	849	30.2%
Total	2812	100.0%

Usage half-life is defined as the time taken for a group of articles to reach half of their total number of downloads.¹ Half of a journal's usage is to articles published more recently than the usage half-life.

To estimate the usage half-life for each journal, participating publishers analyzed a sample of all full text article downloads from their websites and calculated the difference (or age) between the sample date and the date of article publications. The median age² of articles from the sample is the usage half-life.

Due to differences between online publishing systems, three publishers were only able to provide half-life calculations in years, rather than in months. Half-lives were standardized across publishers by placing each journal into a 12-month bin (1-12, 13-24, etc.). Those journals with half-lives exceeding 72 months were placed in a single (≥ 73 month) bin.

Exclusions. To reduce bias in half-life calculations, journals that were not published online before 2010 were excluded from the study, as were journals divested to another publisher or discontinued. Full open access journals were also excluded from the dataset as their usage profiles may be different than subscription-access journals.

¹ All participating publishers comply with Project COUNTER specifications for counting article downloads, see: <http://www.projectcounter.org/>

² As usage distributions have very long tails, the median age of articles is a better indication of their central tendency than average age.

Results

Just 3% percent of journals had usage half-lives of 12 months or less, which ranged from as low as 1% for Life Sciences journals to has high as 6% for Engineering journals.

Across all subject disciplines, journal half-lives peaked between two and four years (24 and 48 months). For Chemistry, Computer Science, Energy & Earth Sciences, Engineering, Life Sciences, and Social Sciences, the median journal half-life was three to four years (37 to 48 months). Health Sciences journals revealed shorter half-lives, with a median of two to three years (25 to 36 months), while the Humanities, Physics, and Mathematics all revealed median half-lives of four to five years (49 to 60 months).

Seventeen percent of all journals had usage half-lives that exceeded six years (or 72 months), from as low as 10% of Health Sciences journals to more than one-third (36%) of Mathematics journals (Figure 1). Nearly one quarter of Physics (24%) and Humanities (23%) journals revealed usage half-lives beyond six years.

These findings reveal considerable variation of usage half-lives both within and between subject disciplines.

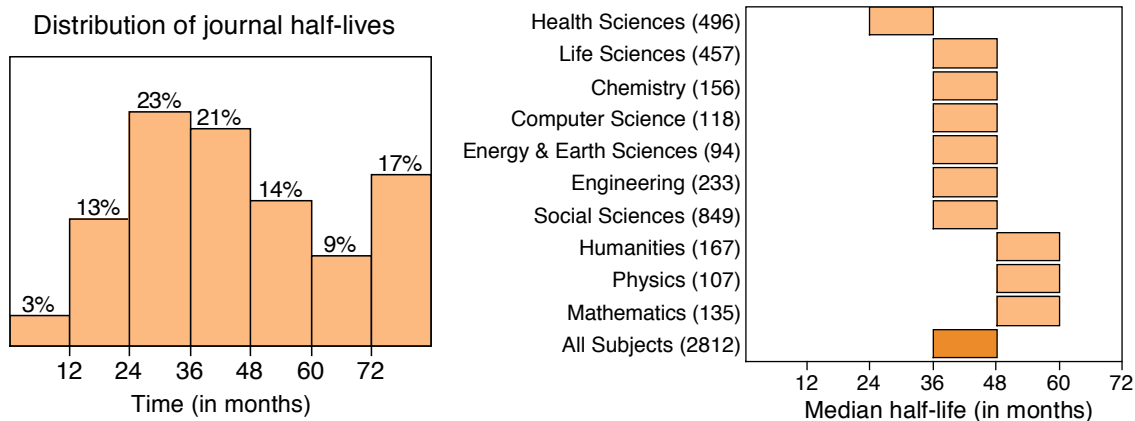


Figure 1. Distribution of journal half-lives

Discussion

This study of 2,812 academic and professional journals published by 13 presses reveals extensive usage of articles from publishers' websites years after publication and considerable variation in usage patterns both within and between subject disciplines. However, these results must be considered within the context of the methods employed in the study.

Estimates derived from sample data. The method for calculating journal usage half-life assumes that a sample of download data is representative of all article downloads from that publisher. Large samples of download data are likely to produce more accurate estimates than smaller samples. Similarly, smaller, low-usage journals may be prone to more variable download patterns than larger, high-usage journals. While individual journal half-life estimates are associated with some degree of error, taken together, they provide a central tendency of usage across various subject disciplines.

Downloads as an indicator of usage. While this study is based on download statistics from publisher websites, other forms of usage may take place when a reader accesses an article from a personal collection, colleague, or public archive.³ It is unknown whether measuring such forms of usage would alter the half-life estimates of journals.

Diversity of publisher access and business models. This study includes 13 publishers operating under a variety of access and business models. While full open access titles were excluded from the dataset, some journals provide public access to their articles after an initial embargo period, while others allow authors to purchase public access to individual articles under a hybrid open access model. Some publishers provide subscribers with access to deep archives of their journals while others sell access to these older articles through separate archival collections. The diversity of access and business models offered by publishers may partially explain the variation of half-life calculations within each discipline.

Generalizability to Funding Agencies. This study is based on estimating the usage half-lives of journals from article download data. It does not distinguish between articles based on funding source, as this information was not consistently documented across journals. Likewise, a single journal may publish articles funded by several separate agencies. Future studies may attempt to track the usage patterns of individual articles through the new FundRef (CrossRef) identifier, although such data will take many years to accumulate.

Acknowledgements

The author was solely responsible for the collection and management of the data, data analysis and interpretation, preparation of the report, and can take responsibility for the integrity of the data and accuracy of the analysis. The study was funded by the Professional & Scholarly Publishing (PSP) division of the Association of American Publishers (AAP).

³ Davis, P.M. 2013. Public accessibility of biomedical articles from PubMed Central reduces journal readership—retrospective cohort analysis. *FASEB Journal* 27(7):2536-2541.
<http://dx.doi.org/10.1096/fj.13-229922>