

Measures of health sciences journal use: a comparison of vendor, link-resolver, and local citation statistics*

Sandra L. De Groote, MLIS, AHIP; Deborah D. Blecic, MLS, AHIP; Kristin E. Martin, MLIS

See end of article for authors' affiliations.

DOI: <http://dx.doi.org/10.3163/1536-5050.101.2.006>

Objective: Libraries require efficient and reliable methods to assess journal use. Vendors provide complete counts of articles retrieved from their platforms. However, if a journal is available on multiple platforms, several sets of statistics must be merged. Link-resolver reports merge data from all platforms into one report but only record partial use because users can access library subscriptions from other paths. Citation data are limited to publication use. Vendor, link-resolver, and local citation data were examined to determine correlation. Because link-resolver statistics are easy to obtain, the study library especially wanted to know if they correlate highly with the other measures.

Methods: Vendor, link-resolver, and local citation statistics for the study institution were gathered for health sciences journals. Spearman rank-order correlation coefficients were calculated.

Results: There was a high positive correlation between all three data sets, with vendor data commonly showing the highest use. However, a small percentage of titles showed anomalous results.

Discussion and Conclusions: Link-resolver data correlate well with vendor and citation data, but due to anomalies, low link-resolver data would best be used to suggest titles for further evaluation using vendor data. Citation data may not be needed as it correlates highly with other measures.

INTRODUCTION

Academic libraries have long struggled with journal subscription price increases that exceed the rate of inflation and, often, the rate of increase for library budgets. Libraries need efficient and reliable means to measure the use of journals to aid in making informed retention and cancellation decisions. Local citation data, in-house journal use, and journal circulation counts were some of the ways to determine print journal usage. However, with the majority of journal subscriptions now online, new methods are available for measuring journal usage. In today's environment, three methods to measure journal use are vendor (often the publisher) platform statistics, link-resolver click-through statistics, and local citation data. None of these can provide a comprehensive picture, because journal use can occur outside the realm of libraries through such means as personal sharing of articles, personal subscriptions, and open access options. But each of these three methods—vendor statistics, link-resolver statistics, and local citation counts—provides measures of use that can be helpful in evaluating journal subscriptions. However, libraries may find it impractical to collect so many types of usage data. The present study measures the correlation between vendor platform statistics, link-resolver click-through statistics, and citation data to determine if one source of data, rather than all three, can be used to reliably indicate high- and low-use journals and assist with making subscription decisions.

Highlights

- The overall Spearman rank order correlation of the vendor and link resolver data was 0.843, allowing libraries to substitute link-resolver data for vendor data in most circumstances, though libraries may wish to look at both for low-use titles.
- Correlations between vendor data and citations, and link-resolver data and citations were slightly lower (0.726 and 0.752, respectively), but there were no cases of highly cited journals that were not also highly used.
- The link-resolver data showed the most spread among usage of titles, but all three measures contained both a large proportion of low-use titles and a small number of high-use titles.

Implications

- Anomalies between the different measures can help libraries identify problems with their vendor accounts or serials knowledgebase links.
- Comparing link-resolver data to vendor data also provides some insight into user behavior and suggests areas for further investigation.

Available metrics

To provide a standardized way for librarians to compare journal use across different publishers and platforms, many vendors and publishers have adopted

* Based on a poster presented at MLA'11, the 2011 Annual Meeting of the Medical Library Association; Minneapolis, MN.

the Counting Online Usage of Networked Electronic Resources (COUNTER) Code of Practice [1]. Journal use can be measured through the number of successful full-text article requests (SFTARs), which indicate how many times articles in a journal are retrieved using a particular journal platform. In contrast to usage statistics for print journals—both circulations and removals from the shelf—that were gathered at the volume level, SFTARs provide an article-level metric.

While SFTARs are a very important measure, using them can be problematic. First, in many cases, due to direct publisher subscriptions as well as subscriptions to aggregated full-text databases, a library may have access to portions of a single journal title on multiple platforms. If a journal is available on multiple platforms, then the statistics from each platform need to be tabulated and merged to get a complete count of all SFTARs. This can present a daunting task when, for example, a library provides access to 40,000 online journals, and some of those titles are available on as many as 10 platforms, as is the case at the study library. Second, sometimes vendor statistics are unavailable or those that are available are incomplete or incorrect due to account problems. Finally, collecting COUNTER statistics can be time consuming. The Standardized Usage Statistics Harvesting Initiative (SUSHI) standard to allow automated harvesting of COUNTER reports is designed to address these issues [2]. Unfortunately, inconsistent decisions on how to implement SUSHI requirements and inaccurate information provided by vendors have made it difficult to use SUSHI to successfully collect COUNTER data from a variety of platforms [3]. (Further challenges with the COUNTER SFTAR data are noted later in the article, as a reflection of the authors' own experience in collecting data for this article.).

An alternative to SFTARs are click-through statistics provided by a link-resolver for all journal titles managed in the knowledgebase, regardless of platform. Link-resolvers (e.g., Ex Libris's SFX; Serials Solutions' 360 Link) use OpenURL standards to provide a link from a bibliographic citation in an index to the location where the full text of the article is available [4]. If the library has access to a journal article on multiple platforms, the link-resolver provides links to all platforms on which the article is available. Because the link-resolver records each click on a given link, each recorded click-through link to a specific journal article or to the journal itself indicates an attempt to access that journal or article. Additionally, libraries using a link-resolver may rely on the same knowledgebase to populate an electronic journal A-Z list or to provide machine readable cataloging (MARC) records for their library catalogs. When a user accesses a journal title through the library catalog or the A-Z list, these click-throughs are counted as well in the link-resolver statistics.

Relying on link-resolver click-through statistics to reflect journal usage has disadvantages. After users click to a journal, they may not access a specific article or, at the other extreme, may access multiple articles, even from other journals on the same platform, and

unlike vendor COUNTER-compliant use statistics, this use will not be counted by the link-resolver. In addition, click-through statistics only record transactions that use the library's tools to access journal content. Users can get to the full-text of articles from other paths than link-resolvers. For example, users at institutions with PubMed's Outside Tool enabled are able to use the link-resolver to access the full text of articles available through the users' institutions or they can click on the publisher links in PubMed. In both cases, the vendor providing the journal will record use of the article, but the link-resolver will only record use in situations when the users employed the Outside Tool feature.

Local citation data, which reflect how many times a particular journal is cited by researchers at a specific institution, provides another measure of use of journals. Citation data carry weight because, with citations, users demonstrate that they value the content of articles by associating them with their own publications. Databases such as Thomson Reuters Web of Science can provide local citation data. Though valuable, local citation data has limitations. It does not demonstrate the overall use of a journal, excluding such uses as consultation for practical purposes (e.g., patient care) or reference by students for assignments. As Rice noted when examining citation analysis for use in collection decisions, "A citation study will tell you what is likely to be used for research, but not what current awareness or review titles supporting instruction are needed" [5]. Nonetheless, citation data can reflect the importance of a journal for specific research purposes, purposes not possible to infer when looking at vendor SFTARs or link-resolver click-through statistics.

Click-through statistics from a link-resolver can provide a tabulated measure of journal use in one easy-to-access and centralized location, making them an attractive alternative to collecting SFTAR data from vendor platforms. However, because vendor SFTAR data are a more complete measure of use than click-through statistics, the authors wanted to test whether a strong positive correlation exists between vendor and link-resolver data. If a strong positive correlation were found, it would suggest that link-resolver data would be an adequate measure of usage for making subscription decisions, obviating the need to undertake the time-consuming collection of SFTARs. If an article is cited by a researcher, the citation reflects engagement by the user with the journal for research purposes. Citation analysis is included in this study to see if the use patterns reflected by the other types of data correlate with citation by faculty members. If citation data reflect different patterns of use compared with vendor COUNTER-compliant SFTARs or link-resolver click-throughs then that would indicate that libraries need to perform citation analysis to understand which journals are important for research on their campuses.

Literature review

At least one previous study examined link-resolver statistics to assist with print journal cancellations.

Gallagher, Bauer, and Dollar examined SFX link-resolver reports in addition to print journal usage at the Yale University's medical library to determine candidates in the print journal collection to cancel [6]. During a 3-month period, SFX statistics demonstrated that of the 3,465 health sciences journals examined, 14.8% (513 titles) were not accessed at all through the SFX link-resolver. In addition, the study found that 10.0% of the online journal titles accounted for 56.8% of all access provided by SFX. In addition to exploring click-through counts, other studies have used link-resolver reports to identify journal titles that had full-text requests but lacked full-text access and used this information to suggest journals to add to the online collection [7].

Previous studies have also examined correlations between various measures of journal use. From 1992 to 1994, Bleic examined the correlation between in-house use, circulation use, and citation use at a health sciences library and found that, for research journals, in-house use, circulation counts, and citation data had high positive correlations, suggesting that if libraries are facing constraints, they could use one method with the knowledge that it significantly correlates with other methods of measuring use [8]. A study by Duy and Vaughan examined the correlation between print journal use data and local citation data with online journal use data provided by the publisher (American Chemical Society, Elsevier, and Wiley) in the subject areas of chemistry and biochemistry. Their study found that there was a significant correlation between print journal use and online journal use. There was also a significant correlation between online journal use and local citation data [9].

The University of Minnesota-Twin Cities Libraries implemented a homegrown click-through journal counter: any time a user clicked on a link on the libraries' web pages or on a link in the online catalog leading to an online journal, a counter recorded the click [10]. This allowed comparison of the local click-through statistics and the vendor statistics. The researchers found a "strong similarity" between the 2 data sets, suggesting that the local click-through statistics were a reliable substitute to the vendor statistics. This study also found that there was an 80/30 rule for journal access in the online environment: 30% of the journals accounted for 80% of the use.

Libraries need efficient and reliable means to measure the use of journals to aid in making informed retention and cancellation decisions. The present study measures the correlation between vendor COUNTER SFTARs, link-resolver click-throughs, and local citation counts for one library's electronic journal collection in the health sciences to determine if one method can be used reliably to assess and identify potential journals for cancellation. Can link-resolver data, even though it is only a partial reflection of use, be used to reliably predict usage determined through other means that are more time consuming to collect (citation and vendor usage data)? Do link-resolver click-through statistics reflect use patterns seen in vendor statistics? Do local citation counts

demonstrate different patterns not reflected in vendor or link-resolver statistics?

METHODOLOGY

The study took place at the University of Illinois at Chicago (UIC), a large urban Research 1 university with six health sciences colleges, a large urban medical center, and three regional medical campuses. Health sciences colleges include medicine, nursing, applied health sciences, dentistry, pharmacy, and public health. Other colleges at the university include engineering, liberal arts and sciences, education, social work, arts and architecture, and urban planning and public affairs. The study university uses Serials Solutions to manage its link-resolver and e-journals A-Z list and to provide MARC records for e-journals.

Despite the multidisciplinary nature of UIC, this study was limited to the health sciences for several reasons. For one, journals accounted for a larger percentage of collections expenditures in health sciences libraries than in other disciplines. In addition, different disciplines differed in their use of the literature. Limiting the study to journals used in the health sciences allowed the results to be focused on a single discipline and thus be more informative, if the result produced practical and applicable findings, such as assisting with subscription decisions. Also, the amount of labor involved in collecting and merging the data necessitated focusing on a subset of interest rather than the entire journal collection of the library. Hierarchical Interface to Library of Congress Classification (HILCC) subject headings assigned by Serials Solutions to titles managed in the knowledgebase were used to help identify appropriate journals, and relevant subject areas such as cellular and molecular biology were included. For general science titles, citation data from only health sciences faculty and researchers were examined to indicate which journals were relevant to health sciences.[†] For this study, journal use data were collected from three sources:

- vendor COUNTER-compliant SFTAR data
- link-resolver click-through statistics
- local citation counts as reported by Web of Science

COUNTER-compliant SFTARs were retrieved from Journal Report 1 (number of successful full article requests) for most of the pertinent vendors and publishers for 2010. A few reports were unavailable or the data incomplete, so the journals covered by these reports were dropped from the study. The 2010 link-resolver data from Serials Solutions were retrieved at the end of January 2011. These data provided combined click-through counts to journal

[†] While this study focuses on health sciences journals, it is likely that researchers from the basic sciences, social work, psychology, biomedical engineering, and other disciplines used the health sciences journals. However, with the Counting Online Usage of Networked Electronic Resources (COUNTER) data and link-resolver data, it is not possible to tell which disciplines are using the journals.

titles from the openURL resolver, e-journal A–Z lists, and links in catalog records to the journal titles. Since UIC uses Outside Tool in PubMed to direct users to full-text content, click-throughs from Outside Tool are sent to the Serials Solutions link-resolver and included in the link-resolver data.

Papers cited in the articles by UIC's researchers were obtained at the end of January 2011 from Thomson Reuters' Web of Science by searching by author affiliation (Univ Illinois and Chicago) and limiting the search to articles published in 2010. The cited references from each retrieved article were entered into a spreadsheet to tabulate each time a journal was cited. Originally, the authors planned to use citation data only from articles published by health sciences faculty and researchers. However, because it was not possible to limit the platform or link-resolver usage data to only uses by individuals in the studied disciplines, local citation data from all disciplines were included. This assured comparability of the data being collected.

Once collected, the vendor and citation data were collated and matched to the title list and click-through data exported from the link-resolver. While the click-through data automatically summed up use for a particular title across all platforms, COUNTER data were provided for each platform on which a portion of a journal run was available. The vendor (COUNTER) data had to be matched first to the appropriate title and then to the appropriate platform to ensure that all data were properly collected, which was complicated because the matching process was not one-to-one. The title list exported from the link-resolver knowledgebase showed that some journals were available from only one source, while others were listed in up to ten different collections.

Some platforms did not supply any COUNTER data (e.g., open access titles published by BioMed Central or government publication). Other platforms were listed multiple times in the knowledgebase but had only a single set of COUNTER data (e.g., one title was available in multiple databases from the same aggregator), and some titles received COUNTER data from a platform not listed as subscribed in the knowledgebase (e.g., a title might have been taken over by another publisher but still had some use data recorded on the former publisher's platform). An example of these complicated collecting issues is the journal *Immunology*, which is available at the study institution through two different EBSCOhost databases—Academic Search Premier and Health Source: Nursing/Academic Edition—and on the Wiley platform, and then *Immunology* additionally had portions freely accessible through PubMed Central. Data collected for this title came from EBSCOhost and Wiley, with some additional data collected through Ovid, for which selected years were available due to post-cancellation access rights on the platform, although that was not indicated in the knowledgebase.

The authors were able to use spreadsheet functions to automate the title matching process using the International Standard Serial Number (ISSN) as the

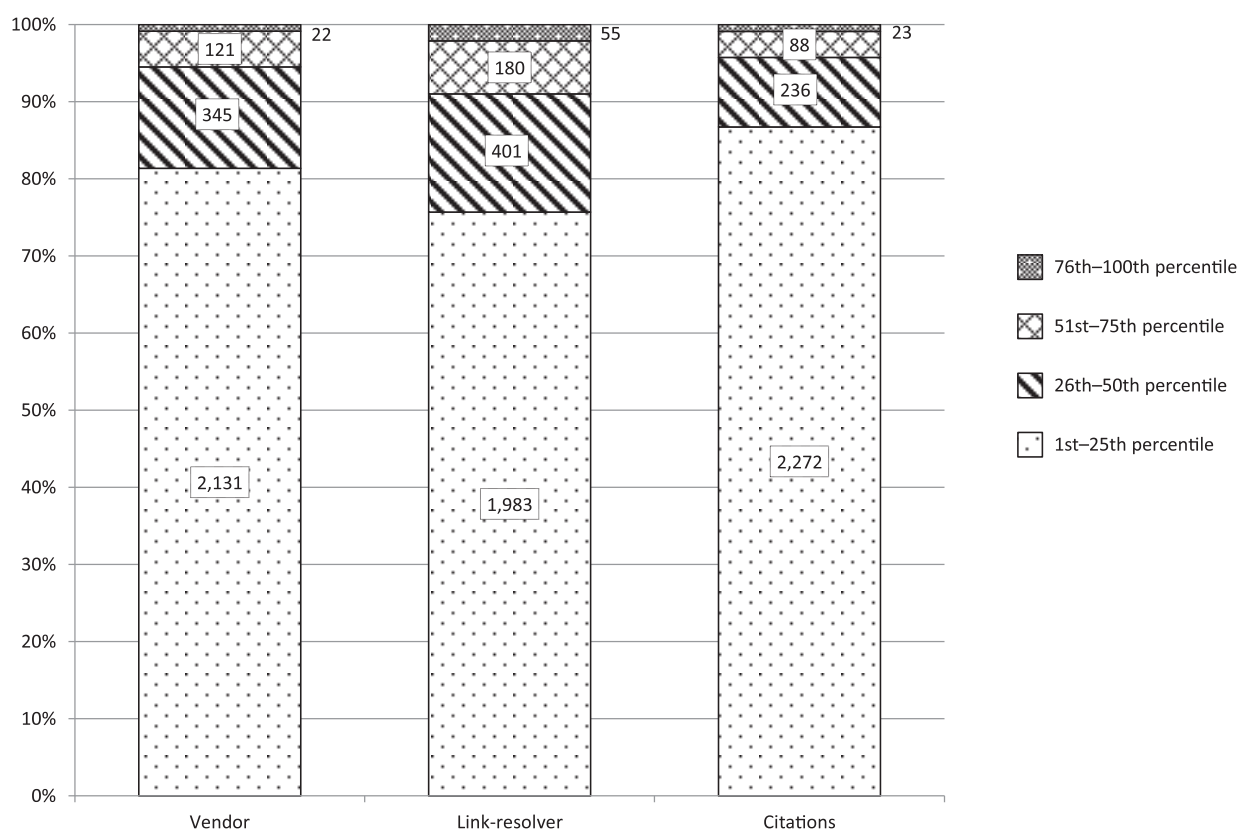
match point, and then results were individually examined to match the COUNTER data to the appropriate platform. Unmatched titles were searched by title and matched up manually. Once all instances of vendor statistics were matched for a given journal title, the SFTARs were summed together, and the sum was used in analysis. Spearman rank-order correlation coefficients were then calculated on the summed vendor SFTARs, link-resolver click-throughs, and local citation counts. The Spearman rank-order correlation coefficient is a measure of how similar the relative rankings of the journal titles are when comparing each ranked list. A result of 1.0 means there is a perfect positive correlation between one ranked list and another: a title that is the fifth highest used on one list is the fifth highest used on the other list. A result of -1.0 means there is a perfect negative correlation so that the highest-used title on one list would be the lowest-used title on the other list, and so on. The further away from 0 the result is, the stronger the relationship between the variables, while a result of 0 would indicate no relationship at all between the lists.

RESULTS

A total of 3,494 current journal titles were identified in the collection as related to the health sciences. Complete vendor data, from all platforms on which the journal was available, was available for 2,619 journals, 75% of the identified titles. Data came from over 30 different platforms. Four hundred sixty-seven titles were removed because they were open access only and thus had no COUNTER statistics available, and an additional 408 titles were removed because complete COUNTER statistics could not be retrieved. Of these 408 titles, the most frequent reason for exclusion was that a title was identified as being part of an aggregator database in the knowledgebase, but no match was found in the statistics collected from the vendor. These titles were generally not core journals for the study library.

When looking at the COUNTER data, only 24% of the titles accounted for 80% of the SFTARs. With the link-resolver data, 29% of the titles accounted for 80% of use. When looking at citations, only 17% of the titles accounted for 80% of use, but 27% of titles received no citations at all in 2010 (compared with 4.5% and 0.8% of titles receiving no SFTARs or click-throughs, respectively). Figure 1 displays a similar concept graphically. If the journals were ranked by usage for each of the 3 measures, a vastly different number of titles would be needed to create 4 equal quartiles of use. For example, for the vendor data, 22 of the highest used journal titles had the same amount of SFTARs as 2,131 of the lowest used journal titles. As the figure demonstrates, the link-resolver data showed the most spread among titles, but for all 3 measures a large proportion of low-use titles accounted for only 25% of use as did a small core of high-use titles.

Figure 1
Number of titles per quartile in a list ranked by usage



The authors then ranked the titles by link-resolver use and next divided the data into 4 almost equal quartiles by number of titles rather than by usage. The number of journals in each quartile did vary slightly because no number was split; in other words, all of the journals with 25 link-resolver click-throughs were put in the same quartile. Vendor SFTARs showed the highest aggregate numbers. In aggregate, link-resolver click-through statistics were about 22%–23% of the number of SFTARs per journal. Click-throughs ranged from 0 to 11,761, and SFTARs from 0 to 71,326. Table 1 shows the values at the 25%, 50% (median), and 75% point in the list of titles ranked by link-resolver use, for each of the 3 measures.

Each set of data (vendor SFTARs, link-resolver click-throughs, and local citations) was compared to each other, and Spearman rank order correlation coefficients were calculated. As shown in Table 2, significant correlations were found between COUNTER SFTARs and link-resolver data, COUNTER SFTARs and local citation counts, and link-resolver click-through data and local citation counts. Figure 2 provides a visual representation of the relationship between the link-resolver data and the vendor COUNTER SFTARs. The twenty-fifth through the seventy-fifth percentiles of the link-resolver data are represented in the scatter chart. The figure demonstrates the correlation between the measures, as well

as the considerable variation in the amount of vendor SFTARs.

The strongest correlation is found between link-resolver and COUNTER data. The strong correlation provides evidence that link-resolver data are a measure of use that can be used reliably in the absence of SFTAR data as an indicator of journal use that reflects the use patterns seen in SFTAR data. Where citation data exist, the correlations of both vendor SFTAR data and link-resolver data with citations also were strong. Many journals that are moderately or heavily used are not cited at all.

Further correlations were executed on specific subsets of the data. Despite efforts to collect all COUNTER vendor data, there were 312 instances where link-resolver click-through statistics were higher than COUNTER SFTARs. While some of these might be due to failure of openURLs to resolve properly and lead users to the correct article, in many cases, the size of the gap between the link-resolver click-throughs and the COUNTER SFTARs suggested that SFTARs were not regularly collected or reported for these titles. Removing anomalies where link-resolver click-throughs exceeded the COUNTER SFTARs strengthened the correlation between link-resolver and COUNTER usage data to 0.857.

Some journals provide free access to the content immediately or after an embargo through open access

Table 1
2010 descriptive statistics for the three datasets by title quartile

List position*	Number of vendor Counting Online Usage of Networked Electronic Resources (COUNTER) successful full-text article requests	Number of link-resolver click-throughs	Number of citations by local authors
25th percentile	68	26	0
50th percentile	283	81	3
75th percentile	757	215	13
Top journal on list	71,326	11,761	1,784

n=2,619 journals.

* Position on a list of titles ranked by number of link-resolver click-throughs.

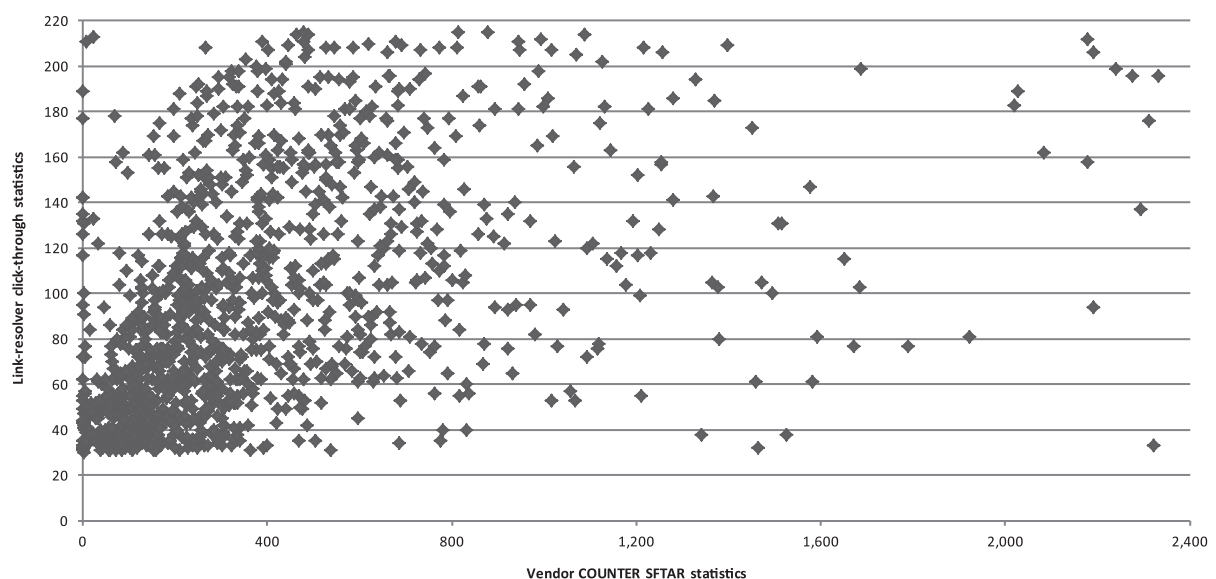
arrangements, which means usage would not be counted through COUNTER unless the user authenticated through the campus proxy server or accessed the article from a campus Internet protocol (IP) range. Removing these journals from the analysis again strengthened the correlation between link-resolver and COUNTER usage data to 0.857, although it slightly weakened the correlation involving citation data to 0.725.

Each of the title quartiles based on an equal number of titles by link-resolver click-throughs was also examined to see if the correlations were affected by where a journal fell on the use spectrum, and those results are listed in Table 2. Although still significant, breaking down the data into quartiles weakened the strength of the correlations. Though this would be expected due to smaller sample size, the correlations weakened especially when the data clustered around a few points of low use in the bottom quartile.

Anomalies in the expected pattern of data were found. Fifty journal titles had citation counts greater than the vendor SFTAR counts, 3 had citation counts

greater than link-resolver click-through counts, and, as noted above, 312 had link-resolver counts greater than vendor SFTAR counts, although only 216 of these titles reported a difference of 5 or greater. Table 3 is a selected list of journal titles illustrating the common use pattern as well as anomalies with link-resolver click-throughs, COUNTER SFTAR statistics, and citation data. The table cautions that despite the strong overall correlations, when the data do not fit the expected use pattern, this lack of fit may indicate a problem elsewhere. For example, there were 1,905 click-throughs for *Diabetes Care*, but only 730 SFTARs. While in some cases link resolution to a specific article may fail or a user may go look at a journal but not choose to retrieve any articles, it is unlikely that would have occurred 1,175 times. It is more likely that the data obtained from vendors are incomplete perhaps due to an account problem with the vendor and further investigation is needed. Similarly, the data for *Cellular Immunology*—32 click-throughs and 10,654 SFTARS—flagged for the authors that the journal was not properly turned on in the library's

Figure 2
Scatter chart of journals from the 25th to 75th percentiles of link-resolver usage



n=1,274 journals.

Note: A few outlying Counting Online Usage of Networked Electronic Resources (COUNTER) successful full-text article requests (SFTARs) counts above 2,400 are eliminated from the chart to make the data fit well in the chart.

Table 2

Link-resolver click-through statistics, vendor successful full-text article requests (SFTARs), and citation data set correlations

Data subsets	Link-resolver use	No. of titles	Link-resolver/vendor		Vendor/citation
			Spearman's correlation	Spearman's correlation	Spearman's correlation
All journals		2,619	0.843	0.752	0.726
1st–25th percentile of ranked titles	≤26	663	0.454	0.392	0.311
25th–75th percentile of ranked titles	≥27–≤214	1,302	0.563	0.408	0.455
75th–100th percentile of ranked titles	≥215	654	0.586	0.511	0.598
25th–100th percentile of ranked titles	≥27	1,956	0.703	0.581	0.622

All correlations were significant at $P < 0.01$.

knowledgebase (i.e., the knowledgebase listed the title as subscribed through an aggregator database, but not on the primary publisher's platform, even though access was available). Users were still able to access content through other methods besides using the link-resolver access tools. With the data in an Excel spreadsheet, it is very easy to run ratios on the data and more closely examine titles that fall at the extremes of proportionality to see if there are problems with the library's vendor accounts or knowledgebase setup.

DISCUSSION

There was a strong positive correlation between the vendor-supplied SFTARs and link-resolver click-throughs, with the vendor SFTARs showing higher usage overall. The results indicate link-resolver click-throughs could, in most cases, reliably suggest high- and low-use journals and can be used when vendor data are missing, incomplete, or unreliable or when staffing does not allow vendor data collection. There was also a positive correlation between both SFTARs and citation counts, and link-resolver click-throughs and citation counts. Many journals that had both SFTARs and click-throughs were not cited, such as *Health Management Technology* and *Cancer Investigation*. Citation counts tell a library which journals are most used for research by faculty, while SFTAR and link-resolver data reflect all types of usage.

Collecting and compiling vendor data from multiple platforms is a time-consuming and error-prone task. Link-resolver data can be very helpful when it is difficult or impossible to obtain complete or reliable vendor data. Link-resolver click-through data can also point out where there is an obvious problem with the COUNTER data, as occurred at the study library for one vendor. Several accounts existed for that vendor's journals, but the vendor supplied the study library with COUNTER data for only one account. High link-resolver click-throughs pointed to a problem with low COUNTER data.

There are limitations, however, in how the link-resolver data can be used to inform subscription decisions. As Table 2 shows, the correlation for the entire dataset is stronger than the correlations for individual quartiles, particularly for the journals in

the lowest quartile. The correlation for the lowest quartile may be weaker in part due to data anomalies, but also likely due to the data clustering around a few points of low use. An increase of just a few uses would be more likely to affect the rank order of a low-use journal than for a journal that already exhibits high levels of use. Table 3 points out some anomalies in data collection and unusual use numbers that could affect the correlation. Thus, it is not possible to assume in all cases that a journal with low click-through statistics should be cancelled. For journals with low click-through counts, librarians are well advised to collect and examine COUNTER SFTARs before deciding to cancel.

Journals in the middle or upper quartile of the link-resolver data are likely in the middle or upper quartile of COUNTER statistics. This correlation relieves librarians of the burden of collecting and examining some COUNTER data and can save significant effort. Collecting, compiling, and matching the COUNTER data took the authors multiple days, even with the sample set limited to current health sciences titles. In particular, the authors found it challenging to match all of the COUNTER data to the appropriate title and database as listed in the knowledgebase. Partially as a result of all of the tedious work expended to collect and collate the COUNTER data, the study university has decided to subscribe to a service that will perform this work on its behalf. However, while a paid service or a SUSHI harvester can greatly reduce the effort, no service or harvester will work for all providers, and the following issues need to be kept in mind regarding COUNTER SFTAR data. First, some providers supply download statistics that are not COUNTER-compliant, while other providers provide no download statistics at all. Second, methods of data collection vary from provider to provider. Frequently an idiosyncratic path must be followed on the platform's administrative website to identify the proper location to download the data. In other cases, the publisher or vendor only supplies the data through email. Third, subscriptions set up under multiple accounts can require a combination of reports. Fourth, with different requirements for different providers, human error can easily occur, with under-collecting and over-collecting possible. For example, data for a

Table 3
Selected journal use statistics, including titles with anomalous results

	Vendor SFTARs	Link-resolver click-throughs	Citation counts
Titles with the common pattern of results: vendor > link-resolver > citations			
<i>Academic Emergency Medicine</i>	992	795	37
<i>Circulation</i>	14,134	2,545	472
<i>Journal of Biological Chemistry</i>	45,526	2,280	1,784
<i>Journal of Family Practice</i>	1,365	808	8
<i>Nature</i>	41,428	4,420	1,015
<i>Proceedings of the National Academy of Sciences (PNAS)</i>	35,994	2,830	1,372
<i>Science</i>	40,155	5,706	997
Titles with zero vendor SFTARs			
<i>American Journal of Clinical Nutrition</i>	0	1,295	140
<i>Biochemistry (Easton)</i>	0	935	364
<i>Infection and Immunity</i>	0	142	122
Titles with zero link-resolver click-throughs			
<i>Internet Journal of Pediatrics & Neonatology</i>	301	0	0
<i>Journal for Specialists in Pediatrics Nursing</i>	799	0	0
Titles with link-resolver click-throughs or citations higher than vendor SFTARs			
<i>Critical Care</i>	19	665	21
<i>Diabetes Care</i>	730	1,905	190
<i>Journal of Studies on Alcohol</i>	496	520	38
Titles with link-resolver click-throughs lower than expected			
<i>Cellular Immunology</i>	10,654	32	5
<i>Developmental and Comparative Immunology</i>	4,069	12	0

single database on a platform can be selected, as opposed to data for the platform as a whole, or, at the other end of the spectrum, statistics can be collected from both the publisher's administrative interface and the publishing platform's administrative interface, resulting in double-counting (e.g., Sage journals, which are hosted by Highwire Press).

Despite the challenges of collecting COUNTER data, some libraries may choose to perform this extra work. Although the correlation coefficient is 0.703 for the top 75% of the journals, about 30% of the variance is still unaccounted for. The noise in the data due to imperfect collection of both COUNTER data and knowledgebase errors that result in inaccurate links to title link-resolver probably lower the correlation strength. Additionally, alternative methods of access to retrieve full text from journals that do not involve link-resolver tools may not be used evenly across all journal titles, creating variation in usage rank order between the COUNTER SFTARs and the link-resolver click-throughs.

In examining the question of whether local citation analysis demonstrates patterns of use not reflected in the COUNTER SFTARs or click-through statistics, the study suggests that libraries do not need to perform citation analysis as an additional measure of use to inform retention and cancellation decisions. For journals that were cited, the study found a strong positive correlation between local citation data and link-resolver click-through data and COUNTER SFTAR data. All journals that were frequently cited were also frequently used, and there were no cases where a highly cited journal had comparatively low SFTARs or link-resolver click-throughs that could not

be explained through a data collection anomaly. However, many journals with moderate or high numbers of click-throughs and SFTARs are not cited, so citation data alone would not be an accurate measure of all use. Citation data as a subset may tell the library which journals are most used for research by faculty, while vendor or publisher statistics and link-resolver data reflect all types of use, including educational and clinical.

In addition to examining the correlations from the study, it is also worthwhile to consider why COUNTER SFTAR statistics recorded much higher levels of use compared to the link-resolver click-through statistics, as this is an area that perhaps warrants further research. Libraries spend substantial dollars on software to facilitate access to online journals, but click-through statistics are comparatively low when weighed against COUNTER SFTARs. In an unpublished survey of health sciences faculty at the study institution, 69% said they used the link-resolver feature to access online journals [11]. At the same time, 46% indicated they would obtain access from the publisher link in a database and 45% indicated they would search Google for access to an article. That such large percentages of respondents use routes other than the link-resolver helps explain the difference between link-resolver and COUNTER data. Interestingly, one complaint made by many of these survey respondents was the difficulty in finding the full text of articles. This would suggest that users are not aware of or do not understand the advantage of utilizing the link-resolver tool and other library routes of access both to identify full-text content across all platforms to which the library subscribes and to

authenticate for access when using a non-campus IP address. Although the study university has customized and branded Google Scholar to lead users through the link-resolver when they attempt to obtain full text, the branding is only applied if the user is identified as having an on-campus IP address or authenticates through the proxy server. It is likely that some users may not gain access to all journal articles available to them through their institutions, particularly when they are off-campus and are using Google or Google Scholar to identify articles.

The study university is not alone in finding that users sometimes fail to use library tools to assist with full-text access. Imler and Eichelberger examined student use of the link-resolver tool at Penn State [12]. Only 54% of student participants located the link-resolver tool right away when their use of the tool was observed. Not all participants equated the text "Get It! Penn State" with locating the full text of an article. The same may be true at UIC, where "Find It @ UIC" is the text provided on the link-resolver icon.

Another reason why the link-resolver statistics may be low as a percentage of the use recorded through vendor SFTARs likely relates to other ways to obtain the full text of an article that do not involve using any of the link-resolver access tools and the limitations of the click-through recording mechanism. Users choosing to go to the home page of a full-text journal or directly to a publisher platform may perform multiple searches and download many full-text articles, but none would be recorded as a click-through because the user never followed a link from the link-resolver to content on or between the platforms. The study university's journal subscription agent has an agreement with publishers to be able to pull into its aggregator databases the full-text content of e-journals to which the study university subscribes. This provides a more seamless experience in obtaining content for the user but also reduces the number of click-throughs uses being recorded. Finally, users may choose to bookmark or go directly to publisher journal sites, to which they linked from tables of content alerts or digital object identifiers, which also bypass link-resolver access tools.

This study also highlighted that some circumstances result in a reduction of SFTARs on the publisher platform. For example, nursing journals in the EBSCOhost database CINAHL Plus with Full-Text were highly used on the EBSCOhost platform, compared to the publisher's platform, indicating patrons' preferences to access the full text seamlessly while searching the index. This study indicates the importance of examining all sources of SFTARs rather than just the publisher platform reports to get a complete measure of use.

Further study of user-preferred database platforms (e.g., Ovid MEDLINE versus PubMed) might also yield information on link-resolver use. Through Ovid MEDLINE, users need to rely on the link-resolver to access the full text, with the exception of links to the full text of articles provided through Ovid. In

PubMed, however, links to the full text on the publisher site are available alongside links to an institution's link-resolver. Users clicking on the publisher's link can go directly to the full text with no additional clicks, but the full text through the publisher is not available in all cases. Sometimes, the institution's link-resolver will provide access to the full text on a platform other than that of the publisher. What impact does database platform have on the use of link-resolvers? The impact of user behavior (e.g., use of bookmarks for journals) might also shed insight on the differences between link-resolver and COUNTER usage data. In addition, it would be useful to perform the same study again over a period of years or among multiple institutions to see if the aggregate ratio of use measured by click-through statistics remains around 22%–23% of use measured by SFTARs.

CONCLUSIONS

Many complexities were found in collecting and analyzing journal use data. For vendor COUNTER statistics, these included identifying, obtaining, and merging statistics from multiple platforms. Services and SUSHI harvesters can offer some assistance, but many complexities will remain until industry standards for the availability and collection of journal use statistics are improved. Link-resolver click-through statistics are more limited in their ability to record use but much simpler to collect and analyze across an entire collection. Because a strong positive correlation was found between click-through statistics and vendor COUNTER SFTARs, librarians can at least in the near term use the link-resolver statistics as one tool to assist in cancellation and retention decisions and journal collection evaluation, especially to identify high-use titles that do not need further evaluation. Citation analysis can identify journals most heavily used in research, but for most titles, it correlates well with other measures of use. Further studies examining user behavior and the access mechanisms used to obtain online journal articles would be helpful both to ensure that users are able to obtain access to the articles they need and to better understand the limitations of different methods that record use.

REFERENCES

1. COUNTER. Counting Online Usage of Networked Electronic Resources [Internet]. 2012 [cited 18 Apr 2012]. <<http://www.projectcounter.org>>.
2. National International Standards Organization. Standardized Usage Statistics Harvesting Initiative (SUSHI) [Internet]. The Organization; 2012 [cited 3 Aug 2012]. <<http://www.niso.org/workrooms/sushi/>>.
3. Pesch O. Perfecting COUNTER and SUSHI to achieve reliable usage analysis. *Serials Lib.* 2011;61(3–4):353–65. DOI: <http://dx.doi.org/10.1080/0361526X.2011.617294>.
4. National Information Standards Organization (US). The OpenURL framework for context-sensitive services [Internet]. NISO Press; 2005 [cited 18 Apr 2012]. <<http://www>>

- .niso.org/kst/reports/standards?step=2&project_key=d5320409c5160be4697dc046613f71b9a773cd9e>.
5. Rice BA. Selection and evaluation of chemistry periodicals. *Sci Tech Lib*. 1983 Fall;4(1):43–59.
 6. Gallagher J, Bauer K, Dollar DM. Evidence-based librarianship: utilizing data from all available sources to make judicious print cancellation decisions. *Lib Coll Acq Tec Serv*. 2005 Jun;29(2):169–79.
 7. Stowers E, Tucker C. Using link-resolver reports for collection management. *Serials Rev*. 2009 Mar;35(1):28–34.
 8. Blecic DD. Measurements of journal use: an analysis of the correlations between three methods. *Bull Med Lib Assoc*. 1999 Jan;87(1):20–5.
 9. Duy J, Vaughan L. Can electronic journal usage data replace citation data as a measure of journal use? an empirical examination. *J Acad Lib*. 2006 Sep;32(5):512–7.
 10. Stemper JA, Jaguszewski JM. Usage statistics for electronic journals: an analysis of local and vendor counts. *Coll Manag*. 2003;28(4):3–22.

11. De Groote SL, Shultz M, Blecic DD. Unpublished survey results of journal and database use of health sciences faculty. Oct 2011.
12. Immler B, Eichelberger M. Do they “get it”? student usage of SFX citation linking software. *Coll Res Lib*. 2011 Sep;72(5):454–63.

AUTHORS' AFFILIATIONS

Sandra L. De Groote, MLIS, AHIP, sgroote@uic.edu, Scholarly Communication Librarian; **Deborah D. Blecic, MLS, AHIP**, dblecic@uic.edu, Bibliographer for the Life and Health Sciences; **Kristin Martin, MLS**, martkr2002@yahoo.com, Electronic Resources Librarian; University Library, University of Illinois at Chicago, Chicago, IL

Received June 2012; accepted September 2012