

Indexes as Tools for Measuring Usage of Print and Electronic Resources

Kathleen Bauer

Libraries are experiencing rapid change as they add electronic resources. These resources are popular with patrons, but their usage is not captured in traditional library statistics. Libraries must find a way to represent the diverse data available for electronic resources in some meaningful way that allows for comparison with statistics for print resources. To do this, the Cushing/Whitney Medical Library has created two indexes, the Electronic and Print Usage Indexes. An index is a simple tool that combines data on items from a group into one number and so represents overall change in the group. The indexes at Cushing/Whitney show that in 1998–1999, patron usage of electronic resources more than doubled, whereas print use declined.



Academic libraries today face a state of change in the composition of their collections and patron usage of those collections. New electronic resources and patrons' desire to use those resources remotely increase every year. In conjunction with this trend, libraries may find traditional measures of usage, such as circulation statistics, declining. The Cushing/Whitney Medical Library at Yale University has seen evidence of this over the past three years. Measurements of print use have declined, and statistics for electronic resource use have risen dramatically. The library now gathers a multitude of numbers that measure print and electronic resource usage but has no uniform reporting method for the data about electronic resource usage and no effective way of comparing numbers for print and electronic use. Given the wide array of data, it is difficult to present statistics in a mean-

ingful way that clearly conveys trends in library usage. What is needed is a method that combines diverse data into a portrait of the state of change in the types of resources patrons are using. This paper explains two indexes created at the Cushing/Whitney Medical Library that can be used easily by librarians to clearly demonstrate usage trends in their libraries.

Usage Measurements: Print and Electronic Resources

Circulation statistics are an accepted way of judging use of a print collection. In addition, some libraries look at shelving data (i.e., the number of volumes shelved) and photocopying activity as indicators of in-house use of the collection.¹ When electronic resources are added to a collection, measuring usage becomes more complex. Patrons may access networked electronic books and journals without

Kathleen Bauer is the Informatics Librarian in the Cushing/Whitney Medical Library at Yale University; e-mail: Kathleen.bauer@yale.edu.

ever having to check them out. An electronic resource may be very popular, but its use will never be reflected in traditional circulation, shelving, or photocopying statistics. These measurements will fail to capture usage of electronic media and will even begin to decline as electronic media replace some print items.² Thus, a library's traditional statistics may indicate that library use is down when the opposite is true.

An important aspect of an index is that it may be used as a tool for comparing changes in very different numbers.

The data available about usage of an electronic resource will vary depending on the resource and how it is accessed. If the product is maintained on the publisher's server, the library may get usage statistics from that publisher. Currently, there is no uniform reporting method, and different publishers typically offer different statistics. For example, one publisher may provide data on each log-on to a site, whereas another

publisher may count search sessions (one log-on may produce multiple searches). At a minimum, the publisher should provide statistics on a monthly basis about how often a particular book or journal is accessed.³ In addition, a library may gather statistics on usage from computer server logs. Given the variety of information available about electronic resources and usage, it is difficult to compare statistics among resources. A typical method is simply to list electronic resources and associated usage statistics, as in table 1.

Table 1 illustrates the drawbacks of listing usage data. Electronic resources are added during the year, so some of the numbers are for part of the year only, and there are no data for the previous year with which to make a comparison. Listing data separately for each item does not give an overall picture of electronic resource usage and provides no frame of reference for comparison with print usage. The data are not collected uniformly: some data may be collected from server logs, and other data may come from publishers. In addition to being gathered in different ways, the data may measure dif-

TABLE 1
Usage of Full-text Electronic Resources, 1998–1999

	1998–1999	1997–1998	Percentage Change
Full-text Books			
Harrison's Online	3216	100	3116%
Medical Letter on Drugs and Therapeutics	560	—	
Merck Manual	939	34	2662%
Scientific American Medicine	546	69	691%
Micromedex	1686	—	
Williams Obstetrics	336	—	
Full-text Journals			
American Journal of Medicine	1099	—	
Annals of Surgery	107	—	
BMJ	3576	—	
Chest	1237	—	
JAMA	5674	—	
Journal of Pediatrics	769	—	
New England Journal of Medicine	5676	—	
RN	555	—	
Science	2202	—	

ferent things, such as individual users, log-ons, or Web accesses. A solution is needed that will reduce the amount of data that are presented and will add a level of abstraction to numbers that should not be compared in their absolute form. An index is a tool that has these qualities.

What Are Indexes, and Why Do People Use Them?

Indexes commonly are used to measure change over time. Famous indexes include the Consumer Price Index (CPI) and the Dow Jones Industrial Average. An index combines different data into one measure that tries to express general change in a group of things. An index in its simplest form works by recording some data about each item in the designated group and totaling those data. In the first, or base, year of the index, that total is set equal to 100. In each subsequent year, the data for each item again are totaled. The index for year t is determined using the following equation:

$$\text{index in year } t = 100 + 100 * (\text{total in year } t - \text{total in base year}) / \text{total in base year}$$

An increase from the base year results in a number greater than 100 and a decrease from the base year results in a number less than 100. An index of 150 indicates an increase of 50 percent since the base year. An index of ninety means a 10 percent decrease.

The CPI is used widely as a measure of inflation in the U.S. economy and is an excellent example of a simple index. It measures the cost of a basket of goods and services, and is calculated monthly by the Bureau of Labor Statistics.⁴ For the CPI, the total cost of the basket of goods is set equal to 100 in the base year. The cost of the same basket of goods is measured again in each succeeding year, and the index is calculated using the formula:

TABLE 2
Cost of Butter and New Automobiles and Corresponding Indexes

	Butter	New Car	Butter Index	New Car Index	CPI
1994	\$1,543	\$18,431	100.0	100.0	100.0
1997	\$2,460	\$20,305	159.4	110.2	108.4

$$\text{CPI in year } t = 100 + 100 * (\text{total cost in year } t - \text{total cost in base year}) / \text{total cost base year}$$

An important aspect of an index is that it may be used as a tool for comparing changes in very different numbers. Changes in small numbers, expressed in absolute terms, can seem very small in comparison to changes in large numbers. Consider one part of the CPI, the per pound price of butter. This is a relatively inexpensive item when compared to other purchases, such as the cost of a new automobile. Converting the cost per item to an index and comparing each to the CPI will help to mask the difference in price and will put the emphasis only on the percentage change. Doing this shows that the inflation in the cost of a new automobile has been much milder than inflation in the cost of butter and that the inflation in the cost of a car is more in line with general inflation in the economy (see table 2).⁵ These comparisons would be much harder to make when looking at absolute numbers.

An index has the additional advantage of combining many data points into one number. An example of this is the Dow Jones Industrial Average, which tracks the value of thirty stocks. Listing the price of each stock would provide more data than most people are willing to look at or are able to comprehend without intensive effort. Combining these thirty stocks into one index has produced a widely accepted and easy-to-understand barometer of activity in the overall stock market.⁶

A final advantage to an index is its ability to represent change over a number of years. Looking at only the percentage

change from year to year may be misleading because the total of each year's percentage change is not equal to the total percentage change from the first year. To see why this is so, think about two different items, widgets and sprockets, both starting in a base year at \$100.00 per gross. Imagine that the widgets increase in cost by two percent, two percent, and six percent, respectively, each year after the base year. The sprockets increase eight percent, one percent, and one percent, respectively. At the end of three years, the widgets cost \$110.28, and the sprockets cost \$110.17. The change in price each year is affected by all the previous increases, so that the total increase for each is not 10 percent and the price at the end is different for each. An index will show how a yearly percentage change affects the price (or whatever is being measured) relative to the year the index began.

As these examples demonstrate, an index may be used to:

- provide a benchmark for making comparisons with individual items;
- mask differences in absolute numbers that can obscure important underlying trends;
- combine several different data points into one easy-to-express and easy-to-comprehend number; and
- measure total change over time.

Given the characteristics of an index, it is an ideal candidate for providing a means of comparing print and electronic usage. The Cushing/Whitney Medical Library chose to create two, the Print and Electronic Usage Indexes, which track usage of print and electronic materials separately.

Methods: Composition of the Indexes

The data chosen for inclusion in the Electronic Usage Index are:

- *Electronic textbooks*: Usage is counted each time an electronic text is accessed from the library's e-textbooks Web page. Data are collected by a script that is run whenever a user selects an e-textbook by clicking on the link to its URL. The script records the access in a

log file and passes the patron on to the URL. These data can only be considered a sample of usage because the script cannot capture the activity of a patron who bookmarks the link to a textbook and in the future bypasses the library's Web site. Although not a complete reflection of use, this sample can be used to gauge the size of change from one year to the next.

- *Ovid full-text sessions*: Usage is counted each time a patron session is initiated in the Ovid electronic journal full-text database, Journals@Ovid. Ovid also provides the library's Medline, Cinahl, and Biosis databases, among others. Access is controlled locally via individual user log-on and password, and this has made it possible to gather precise data on patron usage. Because the medical library purchases access to journals through providers other than Journals@Ovid, these data represent a sample of usage, as is true of the data used for electronic textbooks.

The Electronic Usage Index uses only those statistics that the medical library tracks in-house. The types of data available from publishers are not yet consistent, and even the same publishers sometimes change the type of data they provide. The lack of control the library has over publishers' statistics made them problematic for inclusion in the index. The data used in the index have been included in recent annual reports from the medical library, so they are data that members of the Yale community are familiar with.

The data chosen for inclusion in the Print Usage Index are:

- *Books circulated*: Usage is counted each time a patron checks out a book through a Notis library management system.
- *Photocopies made*: The public photocopiers at the Cushing/Whitney Medical Library record each time a single page is copied. These numbers are tallied at the end of each month. Journals do not circulate, and so photocopying activity is substituted for circulation as a measure of journal usage in the library.

Both indexes are designed to measure

use of a dynamic group of resources. Because so many new types of electronic resources are added each year, large amounts of patron usage would be missed if new items were excluded from the index each year. However, including new resources means that some of the increase in usage must be attributed to the increase in the number of resources offered. The raw usage data are adjusted to account for this growth. The adjustment is computed by dividing the new usage statistics by the corresponding percentage increase in the number of items in that category. The indexes are calculated using the formula:

$$\text{index in year } t = 100 + 100 * (\text{adjusted total in year } t - \text{adjusted total in base year}) / \text{adjusted total in base year}$$

Results and Discussion

These data and the resulting indexes are shown in table 3. A graphical representation is shown in figure 1.

It is evident from the indexes that patron usage of electronic resources increased rapidly in 1998–1999, even after controlling for the effect of the addition of new electronic resources. The Electronic Usage Index of 237.2 indicates that usage in that category more than doubled.

The Print Usage Index of 90.7 reflects a decline in this category of 9.3 percent. The rapid growth in electronic usage gives a clear indication that patrons at the Cushing/Whitney Medical Library are eager to use textbooks and journals in electronic formats. The significant decline in print usage may indicate that patrons are beginning to replace their need for print with electronic information sources.

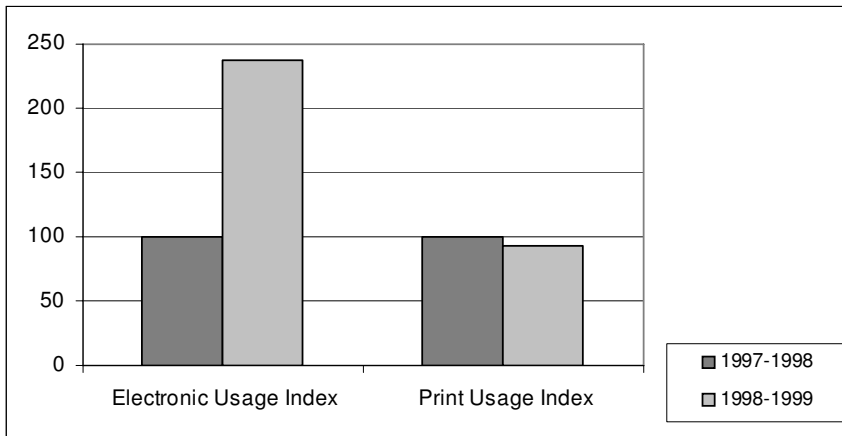
There are caveats to making a comparison between the two indexes. Electronic resources tend to be very current, and so the Electronic Usage Index measures use of current material only. The Print Usage Index looks at all book circulations and all photocopying activity. Perhaps if circulations of only current material were measured in the Print Usage Index, it also would show an increase. That is, it may be only specific segments of the print collection whose use is declining, and this decline is masking an increase in the use of current print materials.

As indicated in the Methods section, the Electronic Usage Index does not reflect use of all e-resources but, rather, is based on a sample. This is also true of other indexes discussed here, such as the CPI. The CPI does not try to measure the inflation in every item purchased in the United States. It looks at a sample of goods, or a “market

TABLE 3
Components and Results for the Electronic Usage Index and the Print Usage Index

Electronic Usage Index	1997–1998	1998–1999	1998–1999 Adjusted	Percentage Change
Electronic textbooks	203.0	12,755.0	4,251.7	1994.4%
Ovid full-text sessions	26,881.0	69,758.0	59,991.9	123.2%
Total	27,084.0	82,513.0	64,243.5	137.2%
Index	100	304.7	237.2	137.2%
Print Usage Index	1997–1998	1998–1999	1998–1999 Adjusted	Percentage Change
Books circulated	32,750.0	29,587.0	29,178.5	-10.9%
Photocopies made	25,709.5	23,958.1	23,822.3	-7.3%
Total	58,459.5	53,545.1	53,000.8	-9.3%
Index	100.0	91.6	90.7	-9.3%

FIGURE 1
Electronic versus Print Usage Indexes



basket," from year to year. By design, the sample used in the Electronic Usage Index is not random (and neither is the sample used in the CPI). Because the sample is not random, there is a danger that it may be skewed in some way.

Two groups of print material should be looked at for usage: material from the last five years and material older than five years.

The exclusion of data from outside sources raises such a concern. The journals the Cushing/Whitney Medical Library has chosen to purchase through Journals@Ovid (the journal group tracked in the Electronic Usage Index) tend to be the most popular and important journals. Perhaps these journals are used in different ways than some of the less popular titles that are not purchased through Ovid. If usage is different for these two groups of journals (Ovid and non-Ovid), the difference could skew the results of the Electronic Usage Index higher or lower.

Areas for Future Study

Of paramount importance in using the Electronic and Print Usage Indexes will be judging their validity in rating relative

usage of the library collection. The indexes will have to be judged against other statistics gathered by the library, such as server statistics, and statistics for individual electronic journals offered outside the Ovid databases. The indexes will be studied to see if the estimates they give of change in usage seem to undercount or exaggerate differences seen in other measurements of usage. One area of concern is that the journals the library purchases in full text from Ovid tend to be popular titles. This could skew the Electronic Usage Index. It would be enlightening to compare usage statistics for electronic journals included and excluded from the Index. The Electronic Usage Index may have to be adjusted to account for its concentration of popular titles.

Another question to be studied is the effect of currency on the use of print and electronic resources. Two groups of print material should be looked at for usage: material from the last five years and material older than five years. Unfortunately, the Cushing/Whitney Library does not have these data available for its print journal collection. A new study of usage of the journal collection by title and date should be started. This will become more important if the library needs to decide

what titles will be carried in print only, what titles should be in electronic format, and what titles should be continued in both formats.

It is clear that the collection of the Cushing/Whitney Medical Library is in a state of change as electronic resources are being rapidly introduced. As electronic resources are added and patrons become fully aware of them, the Electronic Usage Index should continue to show significant increases. At some point, a state of equilibrium should be reached, and the index will be useful for tracking when the rapid growth in electronic resource usage begins to level off. Moreover, the Print Usage Index will be watched over time to see if the decline in print usage accelerates or levels out as some minimum level of print use is established.

Conclusion

Libraries are experiencing changes in the composition of their collections, as electronic resources are added. It is important that libraries report the use of electronic resources because these may be very popular with patrons and their use is generally not captured by traditional library statistics. Because data about electronic resources may come from varied sources, it is difficult to provide the data in an easy-to-understand format. Also, the data a library reports may be increasing to the point where it is difficult to glean meaning from the vast array of numbers. An

index has the advantage of combining many data points into one number and can be used to describe changes in patron usage of print and electronic resources. The Electronic Usage Index created at the Cushing/Whitney Medical Library indicates that patron usage of electronic textbooks and journals is undergoing dramatic growth, whereas the Print Usage Index indicates a decline in the use of traditional printed materials.

The level of accuracy of the indexes remains a question for further study, but two conditions in the current environment argue for guarded acceptance of the dramatic findings produced in this study. One is the anecdotal experience of librarians with patrons at the Cushing/Whitney Medical Library, which is overwhelming preference for digital formats whenever possible. Second, even assuming that the selection of resources for inclusion in the indexes introduced error, the nature of the one-year increase in electronic resource usage is so large (137.2% increase) that even a significant reduction would not erase the obvious advantage of electronic over print. If the Electronic Usage Index increases in the 1999–2000 year in the same dramatic fashion, this will indicate a clear preference of the library's patrons for digital resources. The implication will be that the library must quickly shift priorities from providing patrons with both print and digital resources to making digital resources the focus of library services.

Notes

1. Nancy A. Van House, Beth T. Weil, and Charles R. McClure, *Measuring Academic Library Performance: A Practical Approach* (Chicago: ALA, 1990).
2. Judith Hiott, "Making Online Use Count," *Library Journal* 124, no.16 (Oct. 1,1999): 45–47.
3. International Coalition of Library Consortia, "Guidelines for Statistical Measures of Usage of Web-Based Indexed, Abstracted, and Full Text Resources" (ICOLC, 1998). [rev. Nov. 1998, cited 23 Mar. 2000]. Available online at: <http://www.library.yale.edu/consortia/webstats.html>.
4. Bureau of Labor Statistics, Division of BLS Publishing, *BLS Handbook of Methods* (Washington D.C.: Bureau of Labor Statistics, 1997). [rev. Apr. 1997, cited 23 Mar. 2000]. Available online at: <http://stats.bls.gov/opub/hom/homhome.htm>.
5. ———, *Statistical Abstract of the United States 1998* (Washington D.C.: Bureau of Labor Statistics, 1999).
6. Dow Jones & Company, *What, Exactly, Is the Dow Jones Industrial Average?* (New York: Dow Jones & Company, 2000). [cited 23 Mar. 2000]. Available online at: <http://averages.dowjones.com/faqs.html#1>.