|  |  |
| --- | --- |
| Impact of MEDLINE Selection on Journal Price | August 262011 |
| Project Sponsors: Martha Fishel & Beth Weston | Kristen Greenland |

## Table of Contents

[Abstract 2](#_Toc468194182)

[Introduction 3](#_Toc468194183)

[Methodology 4](#_Toc468194184)

[Results 7](#_Toc468194185)

[Discussion 24](#_Toc468194186)

[Recommendations 25](#_Toc468194187)

## Abstract

**Objective:** The purpose of this project was to determine whether the price of journal titles selected for MEDLINE increases at a higher rate than that of non-selected titles.

**Methods:** Price data for 779 journals considered for inclusion in MEDLINE in 2005-2007 were compiled and analyzed for fiscal years 2001-2011. Titles were divided into those selected and those not selected. Price data was compared for these two groups before and after LSTRC consideration to determine whether a significant increase in price occurred after journals were selected in comparison to the non-selected control group. Journal price data was further analyzed by language, age of publication, non-society vs. society publisher, specific publisher, and subject.

**Results:** Over the ten year period investigated, journal prices increased an average of 10.26% per year. Titles selected for MEDLINE did not show a significant difference in rate of price increase after selection when compared to non-selected titles. MEDLINE titles were consistently more expensive than non-MEDLINE titles, both before and after selection. English titles were consistently more expensive than non-English titles, and titles published by societies were consistently cheaper than those not published by societies. Journals that began publication in 2000 or later increased in price at a higher rate than older journals, both before and after LSTRC consideration. Chemistry journals were the most expensive, followed by Biology and Pharmacy & Pharmacology journals.

**Conclusion:** Journal prices did not increase at a significantly higher rate after selection for MEDLINE, compared with titles that were considered but not selected.

## Introduction

 Health science library collection development and maintenance is becoming more and more difficult as budgets shrink and publications become increasingly expensive. While publishers are tending to move their publications online, many are still available in print. During this transition period, libraries throughout the country are discontinuing print subscriptions in favor of electronic versions. The National Library of Medicine (NLM) collects both the online and print versions to preserve the biomedical literature as part of its core mission, and health science libraries rely on the NLM to serve this function. There are currently over 6,000 titles collected by NLM in both print and electronic versions, and this represents a substantial cost at a time when the budget may be shrinking.

 The NLM subscribes to approximately 10,900 print serial titles, amounting to ~$7.7 million in expenditures during fiscal year 2010. Of this total, ~$5 million was spent on the 4,382 paid titles indexed for MEDLINE. In addition to the cost of subscriptions, NLM spends ~$6 million per year on indexing and data creation for its MEDLINE titles. MEDLINE is NLM’s premier bibliographic database, and each journal included in MEDLINE is indexed using NLM’s Medical Subject Headings (MeSH). Inclusion in MEDLINE increases the visibility of a journal because indexing allows for ready search and retrieval. Publishers are likely to see an increase in the number of subscriptions due to this increase in visibility for their journals, and it is therefore extremely desirable to have a journal included in MEDLINE. Libraries are also more likely to subscribe to MEDLINE journals; MELDINE titles make up 11% of the journals in SERHOLD, the serials database within NLM’s DOCLINE interlibrary loan referral system, but account for 51% of library subscriptions (Maria Collins and Karen Kraly, personal communication). Suggestions on which journals to include in MEDLINE are made by the Literature Selection Technical Review Committee (LSTRC); a panel of external experts that review submitted titles three times a year. Because expenditures for MEDLINE titles far outweigh those for non-MEDLINE titles, it was hypothesized that publishers increase the subscription price of journals once they are selected for inclusion in MEDLINE. In this study, we examined the average cost per year of journal titles before and after selection for MEDLINE to test this hypothesis.

## Methodology

***Compiling the Journal Title List***

The first step in this project was to decide on a set of journals to investigate for differences in pricing. We chose the journals considered by LSTRC in 2005-2007 as our data set because pricing data is available in NLM’s Voyager ILS (version 7.1.0, ExLibris, Des Plaines, IL) from 2000 to present. By choosing all of the journals considered in these years, we could compare those selected for inclusion in MEDLINE to those that were not and compare the time period before selection to the time period after for both groups. We received a list of titles considered in 2005, 2006, and 2007 from Torri Kellough in the MEDLARS Management Section. A total of 414 titles in 2005, 437 titles in 2006, and 418 titles in 2007 were considered by the LSTRC committee, and the aggregated 1,269 titles were our starting data set.

***Filtering the Title List***

 There were several categories of journals that we did not include in our price analysis. We removed titles that were reconsidered and selected for MEDLINE after their designated consideration year but kept those that were reconsidered and were not selected. We also removed journals that do not have a print version because these are purchased through license agreements and individual title price data was not readily available. Titles that were never received by the NLM and are not part of the NLM collection (999 MARC field = NOC) were also removed because we do not have any price data for them. After removing these subsets, we had a total of 998 journals to use in our price comparison analysis.

***Retrieving the Price Data***

 We used Cognos (version 8, IBM, Armonk, NY) to query the Voyager database and extract pricing data for fiscal years 2000-2011. We also retrieved data on the publication dates, publishers, language/country, and MEDLINE status for each title. We set up the Cognos queries to use NLM specific bibliographic IDs (Bib IDs) to retrieve the data and then imported it into an Access database (Microsoft, Redmond, WA).

***Editing the Price Data***

 The price data that we retrieved from Voyager is the invoice price paid by NLM during each fiscal year. Unfortunately, the price paid each year does not always reflect the journal’s price for that year. The most common reason for this is that multiple subscription years are sometimes paid for in one fiscal year. We examined the price data for suspicious patterns (for example: one year there is no price data and the subsequent year the price is twice as high as in the other years) to identify these issues and looked at the details of the invoices in Voyager to determine the price breakdown for the years in question. We removed the price data for the year 2000 because it was the first year price data was entered into Voyager and the data was often incomplete or combined multiple subscription years. The price data was then updated in the Access database.

 There were also a number of journals with missing price data or no price data. We were unable to fill in years where price data was missing because this information is only available on the physical journal issues. These fields were left blank in the spreadsheet and were not factored in to the average price calculations. There were 61 titles with no price data in 2005, 77 in 2006, and 81 in 2007. Titles with no price data were removed from the data set leaving a total of 779 (See Figure 1). Of the 779 titles, 270 were selected for MEDLINE, and 509 were not selected. For a detailed breakdown of the number of titles in each year and category, see Table 1.

**Figure 1: Defining the Data Set**



|  |  |  |
| --- | --- | --- |
| **Table 1: Price Data Compilation** |  |  |
|  | **2005** | **2006** | **2007** |
| **Total** | **414** | **437** | **418** |
| MEDLINE | 149 | 141 | 146 |
| Non-MEDLINE | 265 | 296 | 272 |
| **Filters** |  |  |  |
| Electronic Only | 48 | 49 | 48 |
| Not Our Catalog (NOC) | 12 | 16 | 16 |
| Considered Later | 26 | 31 | 25 |
| No Price Data | 61 | 77 | 81 |
| **Remaining Titles** | **267** | **264** | **248** |
| MEDLINE | 97 | 88 | 85 |
| Non-MEDLINE | 170 | 176 | 163 |

***Normalizing the Price Data***

 Price data was normalized so that data from all three years could be combined into one data set. The average percent increase per year across all three subsets (2005, 2006, and 2007) was calculated at 10.26%. All 2005 prices were increased by this amount, and all 2007 prices were decreased by this amount so that all prices were expressed in 2006 dollars. The years of consideration were then lined up and the data combined into one set. Figure 2 provides a schematic of this process. Price data for this full set of journals was then analyzed using Excel.

**Figure 2: Normalizing the Price Data**



## Results

***Average Journal Price per Year***

 The average journal price was calculated for each fiscal year for all journals considered in 2005, 2006, and 2007 (see Figures 3-5). The average journal price was also calculated for all journals in the full normalized data set (see Figure 6). The average percent increase per year was calculated for the period prior to and after LSTRC consideration and these averages are included in each chart. All percent increase calculations include the consideration year in the “Prior to LSTRC Consideration” time period because publishers determine the price of each journal at the beginning of the year. For example, the percent increase per year prior to LSTRC consideration calculation for 2005 journals used data from FY01 through FY05. Average journal price increased over the ten year period by 186% for 2005 considered journals, 141% for 2006, and by 148% for 2007. Average journal price increased by 135% for the nine year period covered using the entire data set.









***Average Journal Price Comparison***

The average journal price was calculated for each fiscal year for journals considered in 2005, 2006, and 2007 after the titles were divided into selected and non-selected subgroups (see Figures 7-9). The average journal price was also calculated for selected and non-selected titles in the full normalized data set (see Figure 10). The average percent increase per year was calculated for the period prior to and after LSTRC consideration for selected and non-selected titles and these averages are included in each chart. Only the 2006 subset of data showed a large difference between the rate of price increase for selected titles versus non-selected titles after LSTRC consideration.









***Journal Price Comparison by Language***

 Price data was also examined by dividing the data set into English and non-English journals for the full normalized data set. Figure 11 shows the average journal price per year from LSTRC consideration for these two subsets. English and non-English journals were further divided into selected and non-selected titles and the average journal price per year is shown in Figure 12. English journals were much more expensive than non-English journals, but did not increase in price at a higher rate than non-English journals after LSTRC consideration. Non-English journals increased in price at a higher rate after LSTRC consideration, regardless of whether they were selected or not.





***Journal Price Comparison by Age of Publication***

 Price data was further examined by age of publication for the full normalized data set. Journals that began publication before 2000 were separated from those that began publication in 2000 or later. Average journal price per year was then calculated for these subsets and is shown in Figure 13. New journals increased in price at a higher rate than journals that began publication before 2000 both before and after LSTRC consideration. The journals were further subdivided by whether they were selected for indexing, and the average journal price per year for these subgroups is shown in Figures 14 & 15. Selected titles increased at a higher rate than non-selected titles for the new journal subgroup only, and this difference was seen both before and after LSTRC consideration.







***Journal Price Comparison by Society Status***

 Titles in the full normalized data set were also divided based on whether or not they were published by a society. The average journal price per year was calculated for non-society and society publishers and the results are shown in Figure 16. Non-society publications were on average more expensive than society journals but the two groups increased in price at approximately the same rate before and after LSTRC consideration. Titles were further subdivided into selected and non-selected groups and the average journal price per year was calculated. This data is shown in Figure 17, and no difference is seen between the rate of price increase for selected versus non-selected titles after LSTRC consideration.





***Journal Price Comparison by Publisher***

 Price data was also examined by dividing the full normalized data set by specific publisher. Average journal price per year was calculated for publishers that had 20 or more publications in the data set and the results are shown in Figure 18. Elsevier and Wiley-Blackwell were the only publishers that had enough titles to further subdivide into selected versus non-selected titles. Average journal price per year was therefore calculated for Elsevier and Wiley-Blackwell titles in these subgroups and the data is shown in Figures 19 & 20. The data sets were too small to make any real conclusions about price increase trends, however, it is interesting that this small set of Wiley-Blackwell titles, whether selected or not, increase in price dramatically the year after LSTRC consideration.







***Journal Price Comparison by Subject***

 Titles in the full normalized data set were categorized by subject to determine if price differences existed for different areas of research publications. The subject of each title was identified using the Ulrich’s Periodicals Directory (<https://ulrichsweb.serialssolutions.com/>). The most comment main subjects (20 or more titles each) were Biology, Chemistry, Medical Sciences, Pharmacy & Pharmacology, and Psychology and the titles were divided into these groups for analysis (see Figure 21). Chemistry titles were significantly more expensive than titles in all other subject categories. The Biology and Medical Sciences categories were the only ones large enough to further subdivide. Biology journals were divided into selected and non-selected subgroups and the average journal price per year was calculated (see Figure 22). The sample set was too small to make any strong conclusions about price trends. Medical Science journals were also divided into selected and non-selected titles (see Figure 23). Selected titles increased in price at a higher rate than non-selected titles both before and after LSTRC consideration. Because the Medical Sciences category was so large (n=488), we further divided this group into the subject subcategories General, Cardiovascular Diseases, Dentistry, Endocrinology, Oncology, Orthopedics & Traumatology, Psychiatry & Neurology, and Other (see Figure 24). Endocrinology and Oncology journals were the most expensive and increased in price at the highest rates after LSTRC consideration.









## Discussion

 While anecdotal evidence existed for MEDLINE titles increasing dramatically in price after selection, this is the first study to examine whether selection for MEDLINE leads to a significant increase in average journal price per year. We did not see a significant difference between the rate of price increase for MEDLINE selected journals versus non-selected journals after LSTRC consideration. We did find that the titles selected for MEDLINE are significantly more expensive than non-selected titles both before and after LSTRC consideration. This is likely due to the fact that higher quality journals are more expensive to produce. Price is not a factor when the LSTRC committee reviews titles, but if higher quality titles are more expensive to produce, then selecting higher quality journals also means selecting the more expensive journals. This is the likely reason that MEDLINE titles are over three times as expensive on average as non-MEDLINE titles.

 In addition to our study of selected versus non-selected titles for the full set of 2005-2007 journals, we also examined differences in price based on language, length of time published, society status, specific publisher, and subject. We did not find that any of the categories investigated had a difference in rate of price increase after LSTRC consideration, except for non-English journals and this increase occurred regardless of whether the journals were selected or not. We did find that English journals were more expensive than non-English journals, and non-society journals were more expensive than society journals. Subject analysis also revealed several categories of journals that are more expensive. Chemistry journals were a great deal more expensive than any other category, along with Biology and Pharmacy & Pharmacology journals. Within the Medical Sciences subject category, endocrinology and oncology journals were the most expensive.

 There were several issues that arose during this project. It was difficult to compile a list of titles considered by LSTRC in 2005-2007 because the database containing this information is not easy to query. We also found that many of the titles on our list did not have price data in Voyager or the price data was not complete. This was the most difficult issue to deal with in performing price analyses, because there is no other source of price data. We attempted to get price information off the physical copies of some journals in the pilot phase of this study last fall, but many journals do not display the price or display the price in a foreign currency. It was also extremely time consuming to find this data in the stacks and would not have been feasible for our larger data set of almost 1000 titles examined here.

## Recommendations

 One of the major difficulties that arose in the course of this project was retrieval of unique identifiers for journal titles considered by LSTRC. We recommend that the LSTRC database be updated to allow for easy search and retrieval of bibliographic data. Although we did not find that journals selected for MEDLINE increase in price more rapidly than non-selected titles, data compiled in this study can be used to inform budget and collection development decisions. For example, it may be possible to delay acquisition of chemistry titles until they are selected for inclusion in MEDLINE. It may be more difficult to justify acquisition decisions like this for core medical subjects like biology, pharmacy & pharmacology, endocrinology or oncology, which were found in this study to be more expensive than other subjects. However, budget reductions may force difficult collection development decision in the future. The NLM will need to continue collecting all journals indexed for MEDLINE but could use the information in this study to make decisions about non-MEDLINE title acquisition.