5-22-2012

Examining the Impact of the National Institutes of Health Open Access Policy: A Case Study

Nicole R. Webber
University of Northern Colorado, nicole.r.webber@gmail.com

Sandra De Groote

Mary Shultz

Examining the Impact of the National Institutes of Health Public Access Policy: A Case Study

Sandra De Groote, MLIS, AHIP; Mary Shultz, MS, AHIP, Nicole Buccella, MS Candidate
University Library, University of Illinois at Chicago

Introduction & Research Purpose

As of April 7, 2008, all peer-reviewed articles resulting from research funded by the National Institutes of Health (NIH) are required to be submitted to PubMed Central (PMC).

The Law: The NIH Public Access Policy implements Division G, Title II, Section 218 of PL 110-161 (Consolidated Appropriations Act, 2008) which states: SEC. 218. The Director of the National Institutes of Health shall require that all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication.

It is been reported that research made freely available will be accessed and cited more than articles only available through subscriptions. This study seeks to measure the potential impact of the NIH policy to date.

The year 2009 represents the first full year where the NIH Policy has been in place. Articles published in 2009 and funded by the NIH should have a freely accessible version available in PMC by sometime in 2010. What impact has the free accessibility of the PMC articles had on their citation rates compared to articles that appeared in the same journals but not deposited in PMC?

Methodology

- PubMed searches by author affiliation for the institution examined in the study were performed and limited to the year 2009. The retrieved articles were grouped into two categories: articles available in PMC and articles not available in PMC. Verification of presence in PMC was determined by the presence of a PMCID in PubMed.
- A journal's impact factor is based on the number of times a journal is cited and thus, it was important that the two groups of journal articles did not have differing impact factors. A high journal impact factor could be a contributing cause to the number of times an article is cited. To avoid this confounding issue, PMC articles and non-PMC articles from the same journal were matched. If a PMC / non-PMC matching journal pair could not be made, the articles were dropped from the study.
- To avoid further confounding influences, the researchers excluded articles published in open access journals or published in journals where the content became freely accessible after an embargo period.
- The cited references for each publication in each group (PMC articles / non-PMC articles) were downloaded from Web of Science, Google Scholar, and Scopus to determine the total number of citing references for each article.
- The citing articles were examined and the following information was recorded for each article: the total number of citing references found in each database for each article and the total number of citing references for each article (excluding duplicate counts). The total number of citing references were compared to determine what if any differences existed in the number of citing references between the freely available PMC articles VS. the non-open access articles.

Results

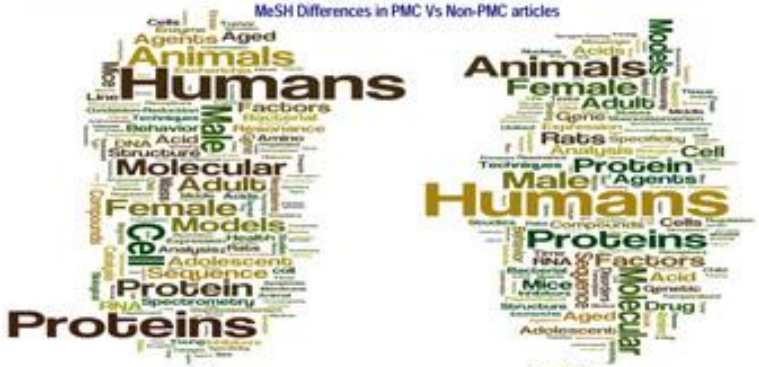
Citing References For Each Article by Database Including Total Unique Citations Retrieved

	WOS Total Cited References	Scopus Total Cited References	Google Scholar Total Cited References	Total Unique Cited References	Average Cited References / Article
PMC articles (n=165)	1969	2142	2436	2779	16.8
Non-PMC articles (n=165)	1544	1727	1947	2225	13.5
Difference in Cited References	425* more in PMC articles	415* more in PMC articles	489* more in PMC articles	476* more in PMC articles	3.3 more per PMC article

*Differences are not statistically significant.

- Typically, Google Scholar found the greatest number of citing reference for each article, while Web of Science presented the fewest number of citing references for both PMC and non-PMC articles.
- The number of citing references for PMC articles was higher than the citing references for non-PMC articles (3.3 more citations per PMC article). However, paired sample t-tests of the two groups in each of the databases did not demonstrate statistically significant differences.
- A total of 330 articles, from 97 journals, remained in the study out the 2,334 articles published by the study institution.
- Of the 330 articles studied:
  - 165 were PMC articles; 165 were non-PMC articles
- Of the 165 articles not in PMC:
  - 155 were funded by some granting agency, 52 of those were funded by NIH; 31 were funded by a non-government agency
- Of the 165 articles in PMC
  - 16 were Publisher PDFs
  - 149 were Author Manuscripts

MeSH Differences in PMC Vs Non-PMC articles



PMC Non-PMC

Discussion/ Conclusions

- The results suggest that the NIH public access policy has led to an increased impact for the research due to availability through PMC. However, differences were not statistically significant.
- The sample size may be too small and not enough time may have passed to allow the articles to be fully cited.

Future Areas of Research

- Obtain a larger sample size; ensure two groups do not differ on other factors in addition to journal impact factors (i.e. topic area, funding)
- Compare articles in PMC due to the NIH Public Access Policy to those not in PMC but that also received NIH funding.
- Study 2009 articles over a longer period of time as most articles will see significant citation impact for many years after the publication date but not as much in the first two years.

(82) Examining the impact of the National Institutes of Health Open Access Policy: A Case Study

Sandra L. De Groote, AHIP, Scholarly Communication Librarian, Daley Library, University of Illinois–Chicago; **Mary Shultz, AHIP,** Regional Head Librarian, Library of the Health Sciences-Urbana, University of Illinois, Chicago, Urbana/Champaign, IL; **Nicole Buccella,** Graduate Assistant, Library of the Health Sciences-Urbana, University of Illinois, Chicago, Urbana/Champaign, IL

Abstract:

Objectives: As of April 7, 2008, all peer-reviewed articles resulting from research funded by the National Institutes of Health (NIH) are required to be submitted to PubMed Central (PMC). It has been reported that research made freely available will be accessed and cited more than articles only available through subscriptions. This study seeks to measure the potential impact of the NIH policy to date.

Methods: Searches by author affiliation for the institution examined in the study were run for the year 2009 in PubMed. The results were further limited to articles funded through NIH, creating two lists of articles: those funded and those not funded by NIH. Verification of presence in PMC was determined by the presence of a PMCID in PubMed. To avoid confounding, the researchers excluded articles published in open access journals. The cited references for each publication in each list were downloaded from Web of Science, Google Scholar, and Scopus to determine the total number of citing references for each article. The total number of citing references were compared to determine what if any differences existed in the number of citing articles between open access articles and non-open access articles.

Results: Typically, Google Scholar found the greatest number of citing reference for each article, while Web of Science presented the fewest number of citing references for both PMC and non-PMC articles. The number of citing references for PMC articles was overall higher than the citing references for non-PMC articles.

Discussion: The results indicate that the NIH public access policy has led to an increased impact of research due PMC availability. However, a methodological flaw was noted where the PMC articles selected for this study came from journals with a slightly higher impact factor than the journals that provided the non-PMC articles. The methodology is being revised to overcome the confounding variable.

Primary Presenter:

Sandra L. De Groote, AHIP

Scholarly Communication Librarian, Daley Library, Chicago, IL