Evidence Based Library and Information Practice

Evidence Summary

PubMed Central: An Essential Resource for Information Professionals and Researchers

A Review of:

Cornell, A., Bushman, B., & Womack, K. (2011). Analysis of journals that did not meet selection criteria for inclusion in the National Library of Medicine collection but have manuscripts in PubMed Central. *Journal of the Medical Library Association, 99,* 168-170. doi: 10.3163/1536-5050.99.2.011

Reviewed by:
Joanne L. Jordan
Research Information Manager
Arthritis Research UK Primary Care Centre, Keele University
Keele, Staffordshire, United Kingdom
Email: j.jordan@keele.ac.uk

Received: 28 Nov. 2012 **Accepted:** 8 Apr. 2013

Abstract

Objective – A review of the journals containing research listed in PubMed Central (PMC), but not selected for inclusion in the National Library of Medicine (NLM) collection. The authors identified reasons why journals had not been included in the collection and if any met the NLM selection criteria and were appropriate for inclusion.

Design – Descriptive study.

Setting – National Library of Medicine, United States.

Subjects – 571 journals that were not included in the NLM collection but had research articles in PMC.

Methods – In October 2009, a report was produced from the NLM library system listing journals tagged as having articles in PMC and not being in the NLM collection. Information was gathered on the journals identified and these were checked against the Collection Development Manual of the NLM and the NLM checklist used for selecting electronic journals. The reason for non-selection of the journal was recorded and the subject category, according to the Library of Congress Classification, was noted.

Recorded reasons why journals were not selected:

- Less than 15% of articles were within scope of NLM collection
- Not enough articles published
- Coverage (lacking original research or not for a scholarly audience)
- Insufficient information to determine reason

For journals where the criteria seemed to be met, the decision on selection to the NLM collection was reviewed.

Main Results – The authors identified 571 journals that had articles in PMC but did not meet the criteria for inclusion in the NLM journal collection. The majority of these journals (73%) were outside the NLM scope and a further 10% had not published a sufficient number of articles to be considered. A further 3% were assessed as not intended for a scholarly audience or lacked original research and another 3% could not be reviewed due to lack of information available. There were 65 journals (11%) that were referred for further review as the selection criteria seemed to be met and 11 of these journals have subsequently been added to the NLM collection. This is in relation to 482 new print and electronic journals in total that were added to the NLM collection in 2009.

However, only 369 of the 571 journals (65%) had one or more articles included in PMC; of these, 238 had one article and 33 had more than four articles in the archive. The reason that some journals had no articles in PMC at the time of this review was due to the time it takes to process new articles and embargos set by the publishers that restrict immediate listing on open access databases such as PMC. A number of these journals may also be new and may not have had a sufficient number of articles or enough information available to be able to include them in the NLM collection. To add context, the authors state that PMC contained over 115,000 NIH-funded articles by the end of November 2010.

The subject areas these non-selected journals were classified under included Engineering (15%); Medicine (14%); Mathematics (10%); Chemistry (10%); and Computer Science (9%). Library Science was assigned to 2% of the journals. The Medicine journals were more likely than those in the other subject areas to be new journals without sufficient articles to be included in the NLM collection.

Conclusion – When the journal title is out of the scope of the NLM collection, an individual article in that journal can still be included in PMC. This provides a solution to the problem of how to collect biomedical research that is not published in biomedical journals. This may be more important in the future as the field becomes more interdisciplinary. This also provides a useful resource for libraries and researchers searching for full-text biomedical articles.

The authors conclude that analyzing the articles from the journals not selected for inclusion in the NLM collection will provide helpful information about the types of biomedical research being published in non-biomedical journals. This will highlight particular areas the NLM should pay attention to in the future.

Commentary

This paper highlights an important issue for health information professionals and researchers; a proportion of health-related research is not published in biomedical journals and therefore is not likely to appear in NLM bibliographic databases such as Medline. The size of this problem will depend on the subject and type of research being conducted and will have less impact on searches for mainstream medical literature than those working closely with, e.g., the social sciences, information technology, engineering, and information science. Those relying on Medline to search for citations may miss articles unless specific subject databases are searched as well, and these can be less well known and more difficult to access and search.

For a systematic review, where it is vital that the search is comprehensive, even a few missing citations could jeopardize the rigour of the review and produce misleading results. Anyone conducting a systematic review should be using a variety of different databases to retrieve citations, but may not have considered adding PMC to the list. Searching PMC as well as Medline may be one way to retrieve a few of the research papers that might otherwise have been missed.

The authors make references to the size of PMC at the time the paper was prepared. However, they have not given much detail about PMC and the non-NIH funded articles that it contains. As of May 2012, PMC contains 2.4 million articles in over 3,000 journals according to their website (http://www.ncbi.nlm.nih.gov/pmc/). The archive contains not only journal articles resulting from NIH-funded research, but also over 1,000 journals that have allowed the full-text articles from all their issues to be deposited in PMC. Funders of research in other countries, such as the National Institutes of Health Research (NIHR) in the U.K., also require researchers to provide open access to articles in a repository. These articles are deposited in UK Pubmed Central, which then becomes part of PMC. The PMC archive is a key resource for information providers and researchers because it provides access to the full text of a diverse selection of health-related research that may otherwise be difficult to obtain.

While it is of concern that health-related research is published in journals not listed in the NLM collection, it is encouraging that this has not gone unrecognized and ways of capturing this information are being explored. The number of additional articles from journals that are not in the NLM collection is small in relation to the size of the archive at present. However, it is important to review this group of articles on an annual basis to ensure that the proportion of articles in journals not selected for the NLM collection is not growing. As the authors say, the world of health research has become more diverse and boundaries between disciplines are not distinct. Organizations such as the NLM will need to keep up with future developments.