

Rex R Robison, PhD, MLS; Douglas J Joubert, MLIS; Alicia A Livinski, MA, MPH; Deborah L Ozga, MLS; Pamela C Sieving, MA, MS – Division of Library Services/ORS

## Objectives and Introduction

This study examined publishing in open access (OA) journals by NIH authors:

1. What are the trends in OA publishing by NIH authors for the period 1997-2007?
2. What proportion of papers published by NIH authors is found in the PMC Open Access Subset?
3. Which PMC search strategy generates the most complete data for tracking open access publishing over time?

The Bethesda Statement on Open Access Publishing (2003; [www.earham.edu/~neters/foa/bethesda.htm](http://www.earham.edu/~neters/foa/bethesda.htm)) requires open access publications to meet two conditions:

- Broad access, use, and distribution rights granted by the author and copyright holder.
- Upon publication, immediate deposit of the work, all supplemental materials, and the permission statement in an online repository.

In April 2008, the National Institutes of Health implemented the NIH Public Access Policy in response to Division G, Title II, Section 218 of PL 110-161 (Consolidated Appropriations Act, 2008). The Policy requires all NIH-funded authors to submit final peer-reviewed journal manuscripts to PubMed Central (PMC). Developed and managed by the National Center for Biotechnology Information, PMC is a free repository of biomedical and life sciences journal articles.

PMC follows the Bethesda Statement on Open Access Publishing in classifying articles as Open Access. Articles that meet these requirements become part of the PMC Open Access Subset, a small portion of all articles in PMC and the source of data for this research.

## Methods

### Search Strategy

PubMed Central was used to find OA publications authored by NIH intramural staff. PMC has several advantages for this type of search including:

- Indexing of affiliations for all authors;
- An OA designation for journals meeting the Bethesda definition of OA;
- Format options for exporting results to reference management software;
- Familiar, easy to use interface.

On October 1, 2008, the following search strategy was run in PMC:

(national institutes of health[ad] OR NIH[ad] OR 20892[ad]) OR ("center for scientific review"[ad] OR "gerontology research center"[ad] OR "national institute of allergy and infectious diseases"[ad] OR National Institute of Environmental Health Sciences[ad] OR national institute on drug abuse[ad] OR national institute on aging[ad] OR national cancer institute[ad] OR national eye institute[ad] OR "national heart lung and blood institute"[ad] OR national human genome research institute[ad] OR "national institute on alcohol abuse and alcoholism"[ad] OR "national institute of arthritis and musculoskeletal and skin diseases"[ad] OR "national institute of biomedical imaging and bioengineering"[ad] OR "national institute of child health and human development"[ad] OR "National Institute of General Medical Sciences"[ad] OR "National Institute of Mental Health"[ad] OR "National Institute of Neurological Disorders and Stroke"[ad] OR "National Institute of Nursing Research"[ad] OR "National Library of Medicine"[ad] OR "Fogarty International Center"[ad] OR "National Center for Complementary and Alternative Medicine"[ad] OR "National Center for Research Resources"[ad] OR "National Institute on Deafness and Other Communication Disorders"[ad] OR "National Institute of Dental and Craniofacial Research"[ad] OR "National Institute of Diabetes and Digestive and Kidney Diseases"[ad]) AND (montana[ad] OR mt[ad] OR north carolina[ad] OR nc[ad] OR maryland[ad] OR md[ad] OR bethesda[ad] OR rockville[ad] OR Baltimore[ad] OR phoenix[ad] OR poolersville[ad]) OR (NCI[ad] OR NHLBI[ad] OR NIMH[ad] OR NIDA[ad] OR NINDS[ad] OR NIGMS[ad] OR NIAID[ad] OR NIAHNS[ad] OR NIBIB[ad] OR NICHD[ad] OR NIDDK[ad] OR NIEHS[ad] OR NIDCR[ad] OR NIA[ad] OR NIDCR[ad] OR NINDS[ad] OR NIMH[ad] OR NIAAA[ad] OR NLM[ad] OR CSR[ad] OR NCRAR[ad] OR NCMHD[ad]) Limits: Open Access Articles

Results were downloaded into EndNote® X1 and exported into Microsoft Excel® for analysis.

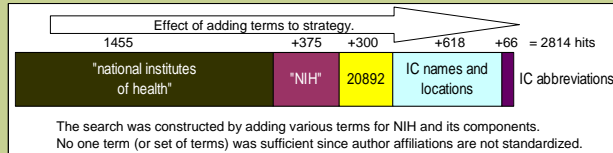
Scopus™ was used to determine the total publication output of NIH authors. It was selected based on its Affiliation Search Tool and 15,000 indexed journals.

- On October 1, 2008, Scopus was searched with a similar search strategy, i.e., variations of "NIH" and names of the institutes and centers (ICs). Variations in institution names ("unmatched citations") that had fewer than 100 hits were ignored. Permutations on this search, e.g. individual terms and year limits, were performed to generate our data.

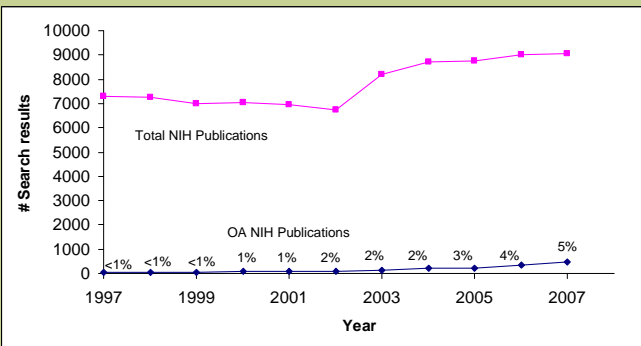
## Results

- We identified 2814 papers in the PMC Open Access Subset by NIH authors; 1786 were published between 1997 and 2007.
- Our strategy retrieved a total of 12,250 NIH-authored papers in the entire PMC archive
- Scopus identified 85,959 NIH-authored papers; this number includes articles available under traditional paid access and various free programs including OA
- Articles by NIH authors included in the PMC Open Access Subset were published in 151 OA journals.
- A complex search strategy was developed to achieve both highly sensitive and highly specific results.

## Development of Search Strategy



## Publications by NIH Authors 1997-2007

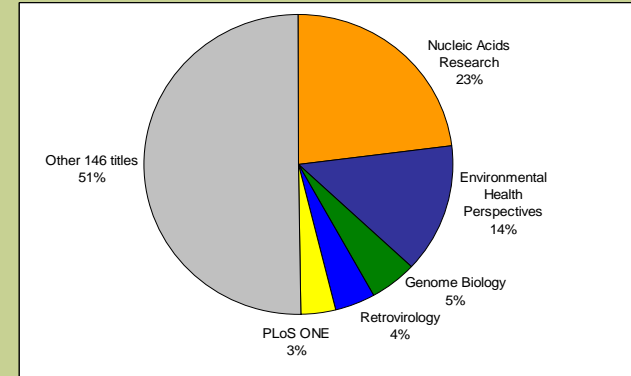


## OA Publishing: Top 5 NIH Institutes/Centers 2005-2007

NLM	20%
NIEHS	8%
NHGRI	5%
NIAID	5%
NCI	4%

Data are for 2005-2007. Analysis disregards ICs with fewer than 100 total publications for this time period.

## OA Journals In Which NIH Authors Most Frequently Published 1997-2007



## Impact Factors\* of OA Journals In Which NIH Authors Most Frequently Published

1. Nucleic Acids Research = 6.954
2. Genome Biology = 6.589
3. Environmental Health Perspectives = 5.636
4. Retrovirology = 4.040
5. PLoS One = no impact factor

\*2007 Science Edition of the Journal Citation Reports® was used to determine impact factors.

## Conclusions

- NIH intramural researchers published in OA journals more frequently at the end of the study period than at the beginning;
- OA articles by NIH researchers are 2.1% of their total publications (1786/85,959) for 1997-2007;
- These OA articles represent 16.6% (1786/12,250) of NIH-authored publications in the general PMC archive;
- A complex search strategy was necessary to achieve both highly sensitive and highly specific retrieval;
- Lack of consistency in authors' identification of their institutional affiliation is the primary reason that these searches are difficult; indexing policies are also inconsistent over time.