Examination and Evaluation of Publications Resulting from Research Projects Listed in HSRProj

Project Lead

Erin Foster, NLM Associate Fellow

Project Sponsors

Patricia Gallagher, Librarian
Lisa Lang, Director

National Information Center on Health Services Research (NICHSR)

11 February 2015

Table of Contents

Abstract

OBJECTIVE:

For this project, the Associate examined a subset of projects from the Health Services Research Projects in Progress (HSRProj) database and explored what publications resulted from these projects. The main objectives are to develop methods of retrieving publications and consider how to automate the retrieval of publications from the HSRProj database.

METHODS:

To complete the project objectives, the Associate performed database and website searches for publications resulting from HSRProj projects and developed search protocols and search strategies to retrieve publications.

RESULTS:

This project resulted in the identification of search strategies to retrieve publications and the development of protocols for searching databases and websites. The Associate also generated a list of databases and webpages, maintained by the Supporting Agencies that house or point to publications resulting from funded projects.

DISCUSSION:

The Associate gained familiarity in searching for grant funded publications and exposure to the various databases and websites that house these publications. The project presented challenges in the search and recovery of HSRProj project publications, particularly in searching for grey literature. The Associate also identified organizations that can partner with the National Information Center on Health Services Research and Health Care Technology (NICHSR) in future efforts to recover and retain project publications.

RECOMMENDATIONS:

The Associate recommends several actions to be taken in light of the results. These recommendations include: integration of identified search strategies into HSRProj, continued exploration of Supporting Agencies included in HSRProj, development of partnerships with organizations and foundations for future work on this project, and consideration of methods to archive publications once included in HSRProj.

Background

The National Institutes of Health (NIH) Revitalization Act of 1993 created the National Information Center on Health Services Research and Health Care Technology (NICHSR) to focus on information related to health services research and health care technology.[1] The main goals of NICHSR are to make the results of health services research available, improve access to these results, and support the information infrastructures that house and disseminate health services research. NICHSR maintains several databases related to health services research, including: Health Services Research Projects in Progress (HSRProj), Health Services and Sciences Research Resources (HSRR), and Health Services/Technology Assessment Text (HSTAT).[2] The products and services provided by NICHSR serve health professionals, health administrators, health policy makers, and affiliated information professionals.

This project focuses on the HSRProj database, which contains citations of health services research projects in progress. AcademyHealth and the Cecil G. Sheps Center, at the University of North Carolina, developed the HSRProj database in 1990 with funding from the Pew Charitable Trust. HSRProj became available through NICHSR starting in 1995 and the database now contains 28,448 project records. HSRProj is freely available and provides access to projects funded through ongoing grants and contracts awarded by federal and state governments, private organizations, and foundations. The purpose of the database is to provide information about research in progress before published results are available.[3]

Projects in HSRProj are included by request of principal investigators conducting health services research or organizations funding health services research. AcademyHealth manages submissions to HSRProj through email or use of an online form. The selection criteria for inclusion in HSRProj are that the project focuses on health services research and is ongoing or within 6 months of completion by the time it appears in HSRProj. Completed projects are available in HSRProj for five years following projects' conclusion; after that, projects are archived. Of the records in HSRProj, about 10,000 are ongoing or completed, the rest are archived. All HSRProj projects - whether ongoing, completed, or archived - are discoverable when searching the database.

HSRProj does not currently include information about, or links to, publications resulting from projects. The inclusion of publication information in HSRProj project records could enhance the value of the database for researchers and policy makers by providing a holistic view of completed projects and their outcomes. Additionally, there is particular interest in identifying "grey literature" produced from these projects. Grey literature is defined as "that which is produced on all levels of government, academics, business, and industry in print and electronic formats, but which is not controlled by

commercial publishers."[4] While traditionally published literature, such as journal articles, are informative and useful in describing project outcomes, grey literature provides an additional lens into projects. As discussed in the Summary Report of the Health Services Research and Health Policy Grey Literature Project, conducted by AcademyHealth, grey literature provides information about project contexts and implications that are not necessarily included in traditionally published articles.[5] Additionally, grey literature is generally written for a lay audience and, as such, provides a level of accessibility not typically offered by journal articles or other types of traditionally published literature.

Project Objectives

This project explored including publication information, for both traditionally published and grey literature, in HSRProj. The Associate examined a subset of projects from the HSRProj database and investigated what publications resulted from these projects. Specific objectives of the project include database and website searching for HSRProj project publications, development of search protocols or search strategies to recover these publications, and consideration of ways in which to automate the process of retrieving publications.

Methods

The subset of projects examined included 1400 projects completed in 2009. The year 2009 was chosen to allow time for the generation of publications from these projects. Initially, the Associate was provided with a spreadsheet of HSRProj projects completed in 2009 that included Project IDs[1] and the start/end dates of the projects as listed in the HSRProj database. In addition to this information, the Associate requested another spreadsheet with the Grant IDs[2] and Supporting Agencies[3] associated with these projects.

Database and website searches for publications

The project proposal listed PubMed, PubMed Central (PMC), and Google as initial places to search for publications resulting from HSRProj projects. In the case of projects funded by the National Institutes of Health (NIH), the NIH RePORTER system was used as a method of searching for publications resulting from funded projects. This system links to PubMed and PMC, which allowed the Associate to view associated publications in these databases. Additionally, through use of the NIH RePORTER system, the Associate discovered grants from other federal agencies, listed in the HSRProj subset, with links to publications included in PubMed and/or PMC.

---

[1] Unique identification numbers assigned to projects in the HSRProj database
[2] Grant numbers assigned by Supporting Agencies
[3] Organizations or institutions that fund projects included in HSRProj

The majority of project publications identified in PubMed and PMC are scholarly articles so, in order to capture the grey literature outcomes (e.g. annual reports, progress reports), the Associate focused on exploring the websites of Supporting Agencies, particularly the websites of non-government Supporting Agencies. Through these searches, the Associate identified databases and webpages that include grey literature as well as information about traditionally published articles.

Development of search protocols and search strategies

To begin, the Associate focused on those projects with a Grant ID listed in HSRProj. Since grant numbers act as unique identifiers, the Associate used this information to create targeted search strategies to identify publications. Initial searches began with NIH funded projects and focused on PubMed and PMC. The Associate searched PubMed, specifically the grant number field, to identify publications in PubMed and PMC. Additionally, the NIH RePORTER system was searched using the grant number to identify publications from NIH funded projects in PubMed and PMC. Publications from other federal agencies were identified using the grant number in NIH RePORTER and retrieved using the grant number field tag in PubMed. During the course of searching, the Associate used documentation provided through the NLM Technical Bulletin and PubMed Help to determine the specific portions of the grant number needed to retrieve publications in PubMed and PMC.

For those HSRProj projects where the grant number was not useful or not available to use in identifying publications, the Associate focused on retrieving publications using the Project Title and Principal Investigator (PI) name information included in the HSRProj records. The Associate began searches in Supporting Agencies' publications databases using a combination of Project Title and PI name to find relevant publications. Unlike the searches completed using grant numbers, searches generated using Project Titles and/or PI names are not as targeted and tend to require manual inspection of the search results. For those Supporting Agencies that maintain webpages of publication citations, rather than databases, searching consisted of manual scans of the listed citations.

Results

The suggested search strategies and search protocols developed in this project are listed below. Additionally, the Associate generated a list of databases and webpages maintained by Supporting Agencies that house or point to publications resulting from funded projects.

Search strategies

The strategies listed below were developed through searches of PubMed and PMC and referral to the NLM Technical Bulletin and PubMed Help.[6],[7] The NIH Institutes and Centers (ICs) listed are those

ICs present in the HSRProj list of projects completed in 2009. These strategies recover publications from PubMed and PMC through use of the Grant ID listed in HSRProj. These strategies are structured as 8 characters (2 letter institute code followed by a 6 digit number with no spaces or hyphens) followed by the grant number field tag. The institution codes can be found in PubMed Help's Appendices.[7] The six digit number is taken from the grant number assigned to a project. For example, if a project is assigned the grant number U10-CA023516-01, the search strategy would be CA023516[gr].

National Institutes of Health Institutes and Centers	Suggested search strategy
Fogarty International Center (FIC)	TWXXXXXX[gr]
National Cancer Institute (NCI)	CAXXXXXX[gr]
National Center for Complementary and Alternative Medicine (NCCAM)	ATXXXXXX[gr]
National Center for Research Resources (NCRR)	RRXXXXXX[gr]
National Eye Institute (NEI)	EYXXXXXX[gr]
National Heart, Lung, and Blood Institute (NHLBI)	HLXXXXXX[gr]
National Human Genome Research Institute (NHGRI)	HGXXXXXX[gr]
National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)	ARXXXXXX[gr]
National Institute of Child Health and Human Development (NICHD)	HDXXXXXX[gr]
National Institute of Dental and Craniofacial Research (NIDCR)	DEXXXXXX[gr]
National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)	DKXXXXXX[gr]
National Institute of Environmental Health Sciences (NIEHS)	ESXXXXXX[gr]
National Institute of General Medical Sciences (NIGMS)	GMXXXXXX[gr]
National Institute of Mental Health (NIMH)	MHXXXXXX[gr]
National Institute of Nursing Research (NINR)	NRXXXXXX[gr]
National Institute on Aging (NIA)	AGXXXXXX[gr]
National Institute on Alcohol Abuse and Alcoholism (NIAAA)	AAXXXXXX[gr]
National Institute on Drug Abuse (NIDA)	DAXXXXXX[gr]
National Institute on Minority Health and Health Disparities (NIMHD)	MDXXXXXX[gr]
National Library of Medicine	LMXXXXXX[gr]

The strategies listed below also retrieve publications from PubMed and PMC for the Agency for Healthcare Research and Quality (AHRQ), the Centers for Disease Control and Prevention (CDC), and the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Mental Health Services (CMHS) using the Grant IDs listed in HSRProj.

Health and Human Services (HHS) Agencies	Suggested search strategy
AHRQ	HSXXXXXX[gr]
CDC – National Center for Chronic Disease Prevention and Health Promotion	DPXXXXXX[gr]
CDC – Public Health Practice Program Office	PHXXXXXX[gr]
CDC – National Institute for Occupational Health	OHXXXXXX[gr]

and Safety	
CDC – National Center for Injury Prevention and Control	CEXXXXXX[gr]
SAMHSA, CMHS	SMXXXXXX[gr]

Additionally, publications from Health Resources and Services Administration (HRSA), Office of Rural Health Policy can be retrieved through use of the Grant ID; however no grant numbers are listed for these projects in HSRProj. Publications from the Department of Veterans Affairs (VA) are also included in PubMed and PMC; however the grant numbers included in HSRProj do not align with grant documentation provided by PubMed Help. Further exploration of VA projects is needed to determine the best method of retrieving publications resulting from VA funded projects.

Search protocols

For those HSRProj projects where grant numbers were not useful or not available, the Associate used Project Title and PI name information to search for project publications. For those Agencies that have publications databases available, a combination of Project Title and PI name captured relevant publications when available, however the search engines provided by these databases limit the degree to which searches can be customized. As a result, this method of searching tends to be labor intensive and requires manual inspection of retrieved results.

Consequently, the Associate explored Google as a method of retrieving publications, both traditionally published and grey literature, from Supporting Agencies' websites. Using Google to limit searches to Supporting Agencies' websites (i.e., site: Supporting Agency URL), as a method of publications retrieval, can be more effective than going through Supporting Agencies' databases. Additional notes about search protocols for select Supporting Agencies are included in Appendix II: Notes on search protocols.

Supporting Agencies' databases or webpages

This is a listing of databases or webpages maintained by Supporting Agencies that house or point to publications resulting from funded projects. A list of Supporting Agencies that do not have databases or webpages for this purpose is included as Appendix I: Supporting Agencies continued.

Supporting Agencies	Database / Webpage
Agency for Healthcare Research and Quality (AHRQ)	PubMed & PubMed Central (PMC)
American Heart Association (AHA)	PMC & AHA Journals
Atlantic Philanthropies	What We're Learning
Blue Cross Blue Shield of Michigan Foundation	Research
California Breast Cancer Research Program (CBCRP)	CBCRP Research Portfolio

Canadian Institutes of Health Research (CIHR)	PMC Canada
Centers for Disease Control and Prevention (CDC)	PubMed & PMC
Commonwealth Fund	Publications
Department of Defense	Defense Technical Information Center (DTIC)
Department of Health, United Kingdom	PubMed
Medical Research Council, United Kingdom	Gateway to Research & Europe PubMed Central
National Institutes of Health (NIH)	PubMed & PMC
National Patient Safety Foundation	Grants Summary Reports
New York State Health Foundation	Resources and Reports
Nova Scotia Health Research Foundation	Research Results
Pennsylvania Department of Health	CURE
Robert Wood Johnson Foundation (RWJF)	RWJF Research & Publications
Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Mental Health Services (CMHS)	PubMed & PMC
United Hospital Fund	United Hospital Fund

Discussion

During the course of this project, the Associate gained experience searching for publications resulting from funded projects, discovered databases and webpages that house or provide information about these publications, and learned about the variety of Supporting Agencies that fund health services research. The primary challenge came in the searching for HSRProj project publications – the Associate experienced the most success using the grant number to develop search strategies. However, not all grant numbers could be incorporated into a search strategy. For these projects, and for those without grant numbers listed in HSRProj, the Associate used other information (i.e. project title, PI name) to discover publications. Unfortunately, these searches did not result in defined search strategies, but did provide insight into potential use of Google for retrieval of publications (see Search protocols). Discovering grey literature was particularly difficult with the most successful identification coming from searches of Supporting Agencies' databases.

When exploring the websites of Supporting Agencies, specifically non-governmental organizations, the Associate noticed that a higher number of organizations have databases devoted to grants than to publications. The grant databases list information about grant-funded projects, but do not include project publications, or links out to project publications. Some organizations have separate databases that contain publications resulting from these grant-funded projects, but no organizations link these systems. This observation led the Associate and project sponsors to consider the efforts of the health services research community to demonstrate the value of grant-funded projects by providing public access to publications.

As a result of this discussion, project sponsor Lisa Lang encouraged the Associate to explore the website of Grantmakers in Health and look for discussions or publications that address the idea of returning value by providing access to publications. The Associate did not find information on this topic on the Grantmakers in Health website, but the discussion encouraged NICHSR to reach out to Grantmakers in Health in hopes of collaborating around this issue. Additionally, the Associate identified the Robert Wood Johnson Foundation (RWJF) as another potential partner in efforts to provide public access to health services research outcomes.

Recommendations

Integrate search strategies into HSRProj

In response to one of the project objectives, the Associate recommends NICHSR incorporate the search strategies, identified in the Search strategies section, into the HSRProj database. Including these search strategies in HSRProj, through development of a search script to parse the grant number provided, will automate the process of retrieving publications. The process of developing and integrating this script into HSRProj can be achieved through discussion with the Office of Computer and Communications Systems (OCCS) at NLM as well as the National Center of Biotechnology Information (NCBI). In addition, the Associate suggests that an additional section be added to HSRProj project records that lists a link out to publications retrieved using the appropriate search strategies.

Continue to explore Supporting Agencies in HSRProj

Federal agencies not fully explored in this project include: Centers for Medicare & Medicaid Services, Department of Veterans Affairs (VA), National Science Foundation (NSF), United States Agency for International Development (USAID), United States Department of Agriculture (USDA), and United States Department of Education, National Institute on Disability and Rehabilitation Research (NIDRR). Those Supporting Agencies listed in Appendices

Appendix I: Supporting Agencies continued also warrant further examination. Additionally, this project was limited to those HSRProj projects completed in 2009; other Supporting Agencies included in HSRProj should be explored in order to determine how they provide, or do not provide, access to publications resulting from funded projects.

Reach out to organizations and foundations

As mentioned in the Discussion, the Associate encourages NICHSR to reach out to Grantmakers in Health and the RWJF to discuss efforts to provide public access to publications resulting from health services research projects. Specific to RWJF, developing a partnership can also encourage

synchronization between HSRProj and RWJF's "Research and Publications" database. These initial partnerships can grow to include additional organizations and foundations as deemed appropriate and beneficial. The development of these relationships aligns with NICHSR's mission to support information infrastructures that house and disseminate health services research.

Consider archiving methods for publications in HSRProj

In order to provide consistent access to health services research project publications in HSRProj, NICHSR should consider methods of archiving publications. Grey literature is especially vulnerable and can disappear more easily than journal articles published as a result of HSRProj projects. As such, the Associate recommends that NICHSR consider methods of retaining grey literature publications so that HSRProj can continue to provide access to these types of publications. Conducting an initial literature search can help determine how other institutions and organizations handle this issue. Additionally, should partnerships develop between NICHSR and other institutions and organizations around issues of public access to health services research outcomes, the issue of preservation of grey literature can also be raised. Establishing a place to house these publications will also need to be addressed – perhaps exploring options at NLM such as the Digital Collections repository. Copyright issues will also need to be taken into account, most likely with individual organizations and foundations (perhaps PIs as well).

References

[1] National Library of Medicine. (2002). *Facts Sheets - National Information Center on Health Services Research and Health Care Technology (NICHSR)*. Retrieved from http://www.nlm.nih.gov/pubs/factsheets/nichsr_fs.html

[2] National Library of Medicine. (1999). *Finding and Retrieving HSR: Tools and Databases*. Retrieved from http://www.nlm.nih.gov/nichsr/db.html

[3] National Library of Medicine. (2004). *HSRProj Frequently Asked Questions (FAQs)*. Retrieved from http://wwwcf.nlm.nih.gov/hsr_project/faq_proj.cfm

[4] GL'99 Conference Program. Fourth International Conference on Grey Literature: New Frontiers in Grey Literature.GreyNet, Grey Literature Network Service. Washington D.C. USA, 4-5 October 1999.

[5] AcademyHealth. (2006). *Health Services Research and Health Policy Grey Literature Projet: Summary Report*. Retrieved from http://www.nlm.nih.gov/nichsr/greylitreport_06.html

[6] Knecht, L. (2006). *NIH Grant Numbers in PubMed Citations*. Retrieved from http://www.nlm.nih.gov/pubs/techbull/mj06/mj06_grant_numbers.html

[7] PubMed Help. (2004). *Grant Number Information Found in the GR Field in MEDLINE/PubMed*. Retrieved from http://www.nlm.nih.gov/bsd/grant_acronym.html

Acknowledgments

Appendices

Appendix I: Supporting Agencies continued

The Associate was not able to find databases that house, or point to, publications from funded projects for these Supporting Agencies. Further research and exploration of these organizations is needed.

Supporting Agencies
Academic Medicine and Managed Care Forum, Aetna
American Cancer Society
American Diabetes Association
American Society of Health-System Pharmacists Research and Education Foundation
Blue Shield of California Foundation
Brain and Behavior Research Foundation
InHealth
John A. Hartford Foundation
Michael Smith Foundation for Health Research
North Carolina Department of Health and Human Services, Division of Medical Assistance
Oncology Nursing Society Foundation
Paso del Norte Health Foundation
Sheba Medical Center
W.K. Kellogg Foundation
William T. Grant Foundation

Appendix II: Notes on search protocols

These search protocols are based on searches of the databases and webpages listed in the Supporting Agencies' databases or webpages section.

Agency for Healthcare Research and Quality (AHRQ)

See Search strategies

American Heart Association (AHA)

AHA public access policy identifies PMC as the repository for AHA funded research. In PubMed, AHA publications can be found using: (PI name) AND (american heart association). Requires manual searching to find relevant articles after using this strategy. AHA also makes journal articles published in HA journals available after six months.

Atlantic Philanthropies

Database of evaluation summaries, reports, case studies, and speeches from 2004 to the present. Searched for publications using Program Titles and PI names as included in HSRProj, but was unable to locate publications.

Blue Cross Blue Shield of Michigan Foundation

Webpage of publication citations from 1987-2013. Located publications from HSRProj projects by scanning for PI name on webpage. Publications found in PubMed following identification of article.

California Breast Cancer Research Foundation (CBCRP)

Research database of publications from 1995-2013. Located publication using PI name. Publication found in PubMed following identification of article.

Canadian Institutes of Health Research (CIHR)

Publications housed in PMC Canada according to the CIHR public access policy. PMC Canada is currently unavailable due to cyber-attack.

Centers for Disease Control and Prevention (CDC)

See Search strategies

Commonwealth Fund

Database of reports, briefs, case studies, journal article summaries, and other grey literature publications. Located publications through use of Project Title and PI name. Website also indicated that majority of articles are available through PubMed as well.

Department of Defense

In response to the White House Office of Science and Technology Policy (OSTP) memo from February 2013, the DOD released a 2014 memo that stated final peer-reviewed scholarly publications (and data) must be included in DTIC. Searched by PI name to find publications.

Department of Health, United Kingdom

Using documentation in PubMed Help discovered that (traditionally published) publications from projects can be included in PubMed. Publications found using Project Title, PI name, and "Department of Health/United Kingdom" in PubMed search. Initially looked at the Publications database on gov.uk, but with further exploration concluded it was limited to publications from Department of Health, UK rather than from projects they fund.

Medical Research Council, United Kingdom

Two databases to search for MRC UK materials. In the Gateway to Research database, located publication using Project Title as in HSRProj. Searched Europe PMC using Grant Lookup; used PI name to find publication. Linked to PubMed citation in Europe PMC record.

National Institutes of Health (NIH)

See Search strategies

National Patient Safety Foundation

A webpage that links to a "Research Program Summary of Progress" pdf which includes citations for published articles.

New York State Health Foundation

A database of resources and reports. Located publication through searching Program Name (as part of Project Title) in "Grant Outcomes Reports" section.

Nova Scotia Health Research Foundation

A webpage of select research publications from funded research. HSRProj projects not listed, but may be useful for other projects included in HSRProj.

Pennsylvania Department of Health

A database of Commonwealth Universal Research Enhancement Program (CURE) research grants including progress reports, final reports, and performance reviews. Located publications using a combination of Project Title and PI name.

Robert Wood Johnson Foundation

A database of research and publications. Located publications through several different methods: search of HSRProj project title and/or search by PI name (Last name First initial). Both methods require manual determination of relevant publications based on search results.

Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Mental Health Services (CMHS)

See Search strategies

United Hospital Fund

A database of publications and presentations. Located publications using Project Title AND PI name.