MEDLINE Search Strategies vs. Relevant Retrieval: How Closely Do They Match for a Research Evaluation Topic?

Presented at the 2007 American Evaluation Association Conference
November 10, 2007

Catherine Pepper, MLIS, MPH², John Araujo, PhD¹, Mona Choi, PhD³, Jerald O’Hara¹, Robin M. Wagner, PhD, MS¹, Guijing Wang, PhD¹, Trevor Woollery, PhD¹

¹Office of Public Health Research, Office of the Chief Science Officer
²Coordinating Center for Infectious Diseases
³National Center for Public Health Informatics
Centers for Disease Control and Prevention
Atlanta, GA
Premise

- No standardized research portfolio evaluation method reported in literature
- Research evaluation is informed and advanced by review of relevant scientific knowledge
- Retrieval of this knowledge depends on:
  - Accessibility of knowledge
  - Quality and extent of indexing of literature databases
  - Searching skills
How Do You Retrieve “Relevant” Literature?

Step 1. Define research question

Step 2. Identify knowledge bases

Step 3. Execute search queries in databases

Step 4. Filter with inclusion/exclusion criteria

Search query

Concepts

- Public health
- Economics
- Policy
- Biomedical sciences
- Research
- Business management
- Construct search strategy
- Search skill
- Concept – Term translation
- Date
- Setting
- Publication type
- Purpose of study
Our Research Question

- What are “best practice” methods, models, and metrics used by organizations for conducting portfolio research evaluation?
  - Discipline: Economics, policy and management, evaluation, health services research, etc.
  - Method: Formal management systems, bibliometric analysis, User assessments, etc.
  - Model: Payback, logic
  - Metric: Input, process, output, outcome, or impact

- Method: Literature Review

- Goal: To apply this knowledge towards framing a methodology for evaluating health impact of a public health research portfolio
Step 1. Search Queries

- Main concepts:
  - What are methods, models, and metrics used by health organizations in various disciplines for conducting portfolio research evaluation?
  - How is public health impact or payback from research measured or evaluated?
**Step 2. What are the appropriate sources for these queries?**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public health</td>
<td>MEDLINE, Cochrane</td>
</tr>
<tr>
<td>Economics</td>
<td>MEDLINE, EconLit</td>
</tr>
<tr>
<td>Policy</td>
<td>MEDLINE</td>
</tr>
<tr>
<td>Biomedical sciences</td>
<td>MEDLINE, Web of Science, Cochrane</td>
</tr>
<tr>
<td>Business, Management</td>
<td>ABI/INFORM</td>
</tr>
</tbody>
</table>
## Step 3. Search Strategies

### Keyword vs. Index Term

<table>
<thead>
<tr>
<th><strong>Keyword</strong></th>
<th><strong>Index Term</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Must appear somewhere in text</td>
<td>Describes content of article</td>
</tr>
<tr>
<td>Often appended by authors</td>
<td>Databases: Subject headings</td>
</tr>
<tr>
<td>At mercy of authors’ wording</td>
<td>Attached by trained human indexers</td>
</tr>
<tr>
<td>Google</td>
<td>Synonym / word variant control</td>
</tr>
<tr>
<td>Helpful when no index term exists for concept</td>
<td></td>
</tr>
</tbody>
</table>
### Step 3. Search Strategies (cont.)

Varies with different databases

<table>
<thead>
<tr>
<th>Search Term</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payback</td>
<td>ABI/Inform</td>
<td>MEDLINE</td>
</tr>
<tr>
<td>Health Policy</td>
<td>MEDLINE</td>
<td>EconLit</td>
</tr>
<tr>
<td>Portfolio Research</td>
<td>[none]</td>
<td>[none]</td>
</tr>
</tbody>
</table>
Searching for Portfolio Evaluation Literature

- "Public Health" Journals: 700
- Indexing: Public health concepts “squeezed” into existing terminology
  - Same terms mean different things in different domains
  - Need creative search strategies using both
    - Indexed vocabulary
    - Non-indexed vocabulary

Guest Editorial: Veterans Health Administration's pain research portfolio and publications.
PMID: 17551868 [PubMed - in process]

Creating Tomorrow's Leaders in Cancer Prevention: A Novel Interdisciplinary Career Development Program in Cancer Genetics Research.
PMID: 17542713 [PubMed - in process]

A hybrid neurogenetic approach for stock forecasting.
IEEE Trans Neural Netw. 2007 May;18(3):551-64.
PMID: 17526350 [PubMed - in process]

Eur J Cancer. 2007 May 21; [Epub ahead of print]
PMID: 17521904 [PubMed - as supplied by publisher]

Health burden and funding at the Centers for Disease Control and Prevention.
PMID: 16476645 [PubMed - indexed for MEDLINE]
Health burden and funding at the Centers for Disease Control and Prevention.

Curry CW, De AK, Ikeda RM, Thacker SB.

Centers for Disease Control and Prevention, Atlanta, Georgia, USA. ccurry@cdc.gov

BACKGROUND: The relationship between domestic funding for selected conditions to the Centers for Disease Control and Prevention (CDC) and the burden of disease and disability in the United States was assessed systematically. METHODS: Using mortality, years of potential life lost (YPLLs), disability-adjusted life years (DALYs), hospital days, hospital discharges, and direct medical costs of conditions, 34 high-burden conditions addressed by CDC programs were identified, and information was collected about the funds spent on each by CDC during fiscal year (FY) 2003. The 34 conditions were grouped into 15 categorical areas, and the relationship between budget and burden was analyzed using correlation and regression methods for each of the categorical areas and for each measure of burden. RESULTS: Of CDC’s total FY 2003 budget of $6.9 billion, 62% ($4.3 billion) of funding was allocated to one of the 34 conditions studied. A positive relationship between budget and burden was identified for all measures of burden, although the correlations varied for the different conditions. CONCLUSIONS: Although examination of the relationship of CDC’s budget to burden measures provides insight into the agency’s portfolio of investments, this exercise also highlights a number of limitations with this approach and the currently available burden measures. Assessment of key public health functions such as emergency preparedness and the collection of vital statistics require development of metrics different from the burden measures used in this analysis. Investment in the development of such metrics warrants consideration.

PMID: 16476645 [PubMed - indexed for MEDLINE]
The burden of disease and disability in the United States has been a focus of public health policy and the Centers for Disease Control and Prevention. The burden of disease and disability includes both medical costs and lost productivity. The burden of disease and disability is estimated by the number of years of potential life lost (YPLLs), hospital days, hospital discharges, and direct medical costs of caring for persons with a given condition. The burden of disease and disability is measured by calculating the relationship between budget and burden for each of the 34 conditions studied. A positive relationship between budget and burden was identified for all measures of burden, although the correlations varied for the different conditions. The relationship between budget and burden was analyzed using correlation and regression methods for each of the categorical areas and for each measure of burden. RESULTS: Of CDC's total FY 2003 budget of $6.9 billion, 62% ($4.3 billion) of funding was allocated to one of the 34 conditions studied. A positive relationship between budget and burden was identified for all measures of burden, although the correlations varied for the different conditions. CONCLUSIONS: Although examination of the relationship of CDC's budget to burden provides insight into the agency's portfolio of investments, this exercise also highlights a number of limitations with this approach and the currently available burden measures. Assessment of key public health functions such as emergency preparedness and the collection of vital statistics require development of metrics different from the burden measures used in this analysis. Investment in the development of such metrics warrants consideration.

PMID: 16476645 [PubMed - indexed for MEDLINE]
Health burden and funding at the Centers for Disease Control and Prevention.

Curry CW, De AK, Ikeda RM, Thacker SB

Centers for Disease Control and Prevention, Atlanta, Georgia, USA. ccurry@cdc.gov

BACKGROUND: The relationship between domestic funding for selected conditions to the Centers for Disease Control and Prevention (CDC) and the burden of disease and disability in the United States was assessed systematically. METHODS: Using mortality, years of potential life lost (YPLLs), disability-adjusted life years (DALYs), hospital days, hospital discharges, and direct medical costs of conditions, 34 high-burden conditions addressed by CDC programs were identified, and information was collected about the funds spent on each by CDC during fiscal year (FY) 2003. The 34 conditions were grouped into 15 categorical areas, and the relationship between budget and burden was analyzed using correlation and regression methods for each of the categorical areas and for each measure of burden. RESULTS: Of CDC's total FY 2003 budget of $6.9 billion, 62% ($4.3 billion) of funding was allocated to one of the 34 conditions studied. A positive relationship between budget and burden was identified for all measures of burden, although the correlations varied for the different conditions. CONCLUSIONS: Although examination of the relationship of CDC's budget to burden measures provides insight into the agency's portfolio of investments, this exercise also highlights a number of limitations with this approach and the currently available burden measures. Assessment of key public health functions such as emergency preparedness and the collection of vital statistics require development of metrics different from the burden measures used in this analysis. Investment in the development of such metrics warrants consideration.

MeSH Terms: “Medical Subject Headings”

- Budgets/statistics & numerical data*
- Budgets/trends
- Centers for Disease Control and Prevention (U.S.)*
- Cost of Illness*
- Health Status Indicators*
- Hospitalization/statistics & numerical data
- Humans
- Morbidity/trends
- Mortality/trends
- Planning Techniques
- Primary Prevention/economics*
- Program Evaluation
- Quality-Adjusted Life Years
- Research Support/economics*
- United States/epidemiology

8-18 MeSH terms are assigned to each article

Supplement MeSH searching with textwords

PMID: 16476645 [PubMed - indexed for MEDLINE]
Current priorities in health research funding and lack of impact on the number of child deaths per year.

Leroy JL, Habicht JP, Pelto G, Bertozzi SM

Mexican National Institute of Public Health, Cuernavaca, Mexico. jleroy@correo.insp.mx

We determined the proportion of research on childhood mortality directed toward better medical technology (i.e., by improving old technology or creating new technology) compared with research on technology delivery and utilization. We also estimated mortality reductions from a research-funding strategy focusing primarily on developing technology compared with one that also focused on delivery and utilization. Ninety-seven percent of grants were for developing new technologies, which could reduce child mortality by 22%. This reduction is one third of what could be achieved if existing technologies were fully utilized. There is a serious discrepancy between current research and the research needed to save children's lives. In addition to increased research on the efficacy of treatment, there is an even greater need for increased research on delivery and use of technology.

Publication Types:
- Research Support, Non-U.S. Gov't

MeSH Terms: Medical Subject Headings
- Biomedical Research Economics
- Biomedical Technology
- Child
- Child Health Services/Utilization
- Child Mortality/Trends*
- Child, Preschool
- Data Collection
- Delivery of Health Care
- Developing Countries
- Financing, Organized Classification
- Financing, Organized Statistics & Numerical Data*
- Foundations*
- Health Priorities
- Health Services Needs and Demand
- Health Services Research Economics*
- Humans
- National Institutes of Health (U.S.)*
- Research Support/Classification
- Research Support/Statistics & Numerical Data*
## Search Strategy: MEDLINE

<table>
<thead>
<tr>
<th>#</th>
<th>Search History</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(portfolio adj4 research).tw.</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>research translation.tw.</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>(research adj4 translat$).tw.</td>
<td>1693</td>
</tr>
<tr>
<td>4</td>
<td>&quot;outcome and process assessment (health care)&quot;/ or &quot;outcome assessment (health care)&quot;/ or &quot;process assessment (health care)&quot;/</td>
<td>33527</td>
</tr>
<tr>
<td>5</td>
<td>3 and 4</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>Health Policy/ or exp Policy Making/</td>
<td>25176</td>
</tr>
<tr>
<td>7</td>
<td>portfolio.tw.</td>
<td>758</td>
</tr>
<tr>
<td>8</td>
<td>6 and 7</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>exp &quot;Costs and Cost Analysis&quot;/</td>
<td>71301</td>
</tr>
<tr>
<td>10</td>
<td>9 and 7</td>
<td>48</td>
</tr>
<tr>
<td>11</td>
<td>10 not 8</td>
<td>43</td>
</tr>
<tr>
<td>12</td>
<td>exp research/ or research support/</td>
<td>341029</td>
</tr>
<tr>
<td>13</td>
<td>12 and 7</td>
<td>124</td>
</tr>
<tr>
<td>14</td>
<td>13 not (8 or 10)</td>
<td>114</td>
</tr>
<tr>
<td>15</td>
<td>(health adj2 impact).tw.</td>
<td>3032</td>
</tr>
<tr>
<td>16</td>
<td>15 and 12</td>
<td>315</td>
</tr>
<tr>
<td>17</td>
<td>exp Public Health/</td>
<td>2003298</td>
</tr>
<tr>
<td>18</td>
<td>16 and 17</td>
<td>270</td>
</tr>
<tr>
<td>19</td>
<td>methods.fs.</td>
<td>957386</td>
</tr>
<tr>
<td>20</td>
<td>18 and 19</td>
<td>67</td>
</tr>
<tr>
<td>21</td>
<td>exp evaluation studies/</td>
<td>322730</td>
</tr>
<tr>
<td>22</td>
<td>16 and 21</td>
<td>131</td>
</tr>
<tr>
<td>23</td>
<td>ec.fs.</td>
<td>133718</td>
</tr>
<tr>
<td>24</td>
<td>**&quot;Outcome Assessment (Health Care)&quot;/</td>
<td>9929</td>
</tr>
<tr>
<td>25</td>
<td>15 and 23 and 24</td>
<td>5</td>
</tr>
</tbody>
</table>
"In this article we examine the merits of a ‘portfolio approach’ to generating research funds in . . . “
Search Strategy: ABI/INFORM

Keyword
Keyword
Index Term

Marked Documents

Marked Documents lets you save documents you find useful. To add documents to your marked items:

- With Results displayed--Select the check box to the left of a document title.
- With a document displayed--Select the "mark document" check box on the document display page.

Once you have documents marked, you can create your bibliography, email marked documents, export citations, or create a web page with links to your marked items.

* The maximum number of documents you can save is 50.

Recent Searches

1. (translation w/4 research) AND (health)
   Database: ABI/INFORM Research
   Look for terms in: Citation and abstract
   Publication type: All publication types
   5 results as

2. (portfolio w/4 research) AND (health)
   Database: ABI/INFORM Research
   Look for terms in: Citation and abstract
   Publication type: All publication types
   15 results as

3. LSU(PAYBACK METHOD)
   Database: ABI/INFORM Research
   Look for terms in: Citation and abstract
   Publication type: All publication types
   107 results as
Step 4. Inclusion/Exclusion Criteria

- Inclusion criteria
  - Peer-reviewed
  - Journal articles
  - 1995 – present

- Exclusion criteria
  - Not program outcome assessment
# Step 4. Team Review

<table>
<thead>
<tr>
<th>Database</th>
<th>Total # References Retrieved from Search</th>
<th>Screened by Searcher</th>
<th># Relevant Articles Reviewed to Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE</td>
<td>244</td>
<td>184</td>
<td>12</td>
</tr>
<tr>
<td>ABI/Inform</td>
<td>27</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Web of Science</td>
<td>15</td>
<td>8</td>
<td>TBD</td>
</tr>
<tr>
<td>Cochrane</td>
<td>37</td>
<td>25</td>
<td>TBD</td>
</tr>
<tr>
<td>EconLit</td>
<td>134</td>
<td>46</td>
<td>TBD</td>
</tr>
<tr>
<td>Internet</td>
<td>---</td>
<td>69</td>
<td>35</td>
</tr>
<tr>
<td>TOTALS</td>
<td>457</td>
<td>346</td>
<td>48</td>
</tr>
</tbody>
</table>
Conclusion

- Must search the gamut of knowledge databases to identify all of the relevant literature
- **Do not** rely solely on Google
- Search strategies must be tailored to the database
  - Same indexing terms can have different meanings in different databases, e.g.,
    - In the ERIC database, “portfolio assessment” means “student evaluation”
- The use of screening criteria is helpful for finding “relevant” literature
- For complex searches, enlist a reference librarian
The Bigger Picture

- Develop process with specialized expertise
- Characterize accurately a public health portfolio for later retrieval and reporting
  - OPHR Strategic Objectives
- Pave way to achieve Excellence in Science Objective B1:
  - Sustain and expand intramural, extramural, and intra-agency research that meets the highest standards of quality to provide the foundation for evidence-based public health decisions and practices (Popovic)
- Support goals-driven research – must be tagged appropriately
  - Identify the knowledge gaps and the evidence
  - Evidence must be retrievable
References and Resources

- PubMed Searching Tutorials
  - Searching the Public Health Literature
    www.sph.umich.edu/mi-info
  - PubMed Tutorial—Northwestern
    www.galter.northwestern.edu/tutorials/pubmed
Contact Information

Cathy Pepper, MLIS, MPH
Public Health Informatics Fellow
Informatics Office, Office of the Director,
Coordinating Center for Infectious Diseases
Centers for Disease Control and Prevention

Phone: (404) 639-6018
Address: 1600 Clifton Road NE
Mail Stop E-08
Atlanta, GA 30333
E-mail: cpepper@cdc.gov