Mapping the Multidisciplinary Field of Public Health Services and Systems Research

Jenine K. Harris, PhD, Kate E. Beatty, MPH, Jesse D. Lecy, PhD, Julianne M. Cyr, BA, Robert M. Shapiro II, MALS

Context: Public health services and systems research (PHSSR) is the field of study charged with evaluating the public health system. PHSSR currently lacks a clear identity integrating the many theories, approaches, and disciplines contributing to the field.

Evidence acquisition: Experts in PHSSR were consulted to identify 11 key published PHSSR studies. With these articles as a starting point, a newly developed citation data collection system was used to collect a sample of 2986 documents connected to the key articles through citation linkages. Data were collected in October 2009.

Evidence synthesis: Citation network methods and latent position cluster modeling were used to examine the network of documents. A subset of 108 documents comprising the backbone of the network was identified through main-path analysis. Four unique clusters were identified within the main path. The core cluster consisted of older articles focused on local health department activities, partnerships, and effectiveness. The three non-core clusters focused on public health law, behavioral interventions, and national performance standards. Although all non-core clusters cited the core, there was little crosstalk among the non-core clusters, a pattern consistent with multidisciplinary fields.

Conclusions: PHSSR appears to be a multidisciplinary field with research happening in silos across different research areas. Developing transdisciplinary research connections across PHSSR is necessary to meet national PHSSR goals. (Am J Prev Med 2011;41(1):105–111) © 2011 American Journal of Preventive Medicine

Context

Public health services and systems research (PHSSR) is concerned with the “organization, financing, and delivery of public health services within communities and the impact of these services on public health.”1 Efforts to understand and strengthen the public health system date back more than 90 years to the first report of the American Public Health Association’s (APHA’s) Committee on Municipal Health Department Practice in 1921.2,3 Early PHSSR following this report was characterized by a lack of consistent focus.3 Although PHSSR goals have become more formalized and focused in the past 20 years, it remains difficult to draw clear boundaries that would serve to identify PHSSR. The field has been referred to by several different acronyms4,5 and no specific PHSSR keywords exist among the Medical Subject Headings (MeSH) used to index MEDLINE/PubMed and other health science databases.6 This lack of identity may stem from the nature of the field itself. PHSSR currently combines a range of theoretic and methodologic perspectives, including health services research, epidemiology, biostatistics, economics, sociology, psychology, political science, information science, and operations research.3 Given the importance of the public health system, there is a need to better understand the field that is charged with examining and improving it.7,8 This study aims to elucidate PHSSR through a new citation network approach.9,10

Evidence Acquisition

By following the citation connections among articles within a field, citation network approaches aid in understanding research development11,12 and relationships among research areas.1,2,13

Data Collection

Citation Network Analyzer (CNA), a citation data collection tool developed by one of the authors in 2007, was designed as an efficient
and inclusive sampling framework using snowball sampling techniques. Specifically, seed articles in a research area are identified and a network is generated by following citations forward in time from the seeds using a specified sampling rate and a defined number of levels or generations (Lecy JD, Mergel IA, Schmit HP. Moving beyond the boundaries of network research in public administration. Under review [unpublished manuscript]). What results from this process is a network of documents that represents research trends in a field of study (Appendix A, available online at www.ajpmonline.org) (Lecy JD, Mergel IA, Schmit HP. Moving beyond the boundaries of network research in public administration. Under review [unpublished manuscript]).

The CNA uses Google Scholar and samples based on PageRank, which is calculated internally by Google Scholar’s search engine (Lecy JD, Mergel IA, Schmit HP. Moving beyond the boundaries of network research in public administration. Under review [unpublished manuscript]). PageRank gives higher weights to publications cited by important papers and to publications cited in shorter reference sections, because shorter reference sections are likely to be selective. Maslov and Redner demonstrated that PageRank appropriately identifies important articles in a research area (Lecy JD, Mergel IA, Schmit HP. Moving beyond the boundaries of network research in public administration. Under review [unpublished manuscript]). Only documents with high PageRank centrality are retained by CNA. This strategy identifies compact samples that retain important linkages, key publications not used as seed articles, and other articles that are cited for many reasons including to express support or opposition to an argument or approach. The resulting network is highly inclusive and does not necessarily represent a homogeneous voice.

The Center for Public Health Systems and Services Research (CPHSSR) at the University of Kentucky was consulted to identify PHSSR experts. The CPHSSR used their experience developing a comprehensive PHSSR database with the National Library of Medicine, along with information from the database itself, to provide a list of seven experts. The experts were asked via e-mail to identify three to five of the earliest top articles in PHSSR. Four of the seven experts responded to produce a list of 19 unique seed article candidates. Articles published in the previous 5 years and articles cited fewer than ten times in Google Scholar were removed, producing a list of 11 seed articles (Table 1). A sample of 20% for three levels was collected. Data were collected in October 2009; data management and analysis took place between December 2009 and June 2010.

### Table 1. Seed articles

<table>
<thead>
<tr>
<th>Title</th>
<th>Study</th>
<th>Source</th>
<th>Citation count*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local health department effectiveness in addressing the core functions of public health</td>
<td>Turnock (1994)²²</td>
<td>Public Health Reports</td>
<td>73</td>
</tr>
<tr>
<td>Not-so-strange bedfellows: models of interaction between managed care plans and public health agencies</td>
<td>Halverson (1997)²⁴</td>
<td>Milbank Quarterly</td>
<td>55</td>
</tr>
<tr>
<td>From measuring to improving public health practice</td>
<td>Turnock (1997)³</td>
<td>Annual Review of Public Health</td>
<td>41</td>
</tr>
<tr>
<td>Behind the curve? What we know and need to learn from public health systems research</td>
<td>Mays (2003)¹</td>
<td>Journal of Public Health Management and Practice</td>
<td>17</td>
</tr>
</tbody>
</table>

*Using Google Scholar
Analysis
Three types of analysis were conducted to examine the network: (1) descriptive; (2) main-path analysis; and (3) latent position cluster modeling.

Descriptive analysis: Identifying key journals, articles, and authors. Descriptive network measures were used to identify prominent journals, articles, and authors in the network. Citation networks are inherently directed; links represent the flow of information from article A to article B. If article B cites article A, information flow is shown as A→B. Descriptive network measures included (1) in-degree: how many sources an article cites in the network; and (2) out-degree: how many times an article was cited in the network.

Main-path analysis: examining the backbone of the research area. Main-path analysis is used to identify articles that act as the backbone for a body of research by calculating traversal weights for all nodes and links in the network. These weights indicate the proportion of all paths through the network between each source article (article that cites no others) and a sink article (article that is cited by no others) that contain a specific node or link.Traversal weights indicate how important a particular node is to holding the network together. A traversal weight cutoff of at least 1 SD above the mean traversal weight was selected in order to identify the main path.

Latent position cluster modeling: finding subfields. Recent developments in statistical network modeling and software allow researchers to test hypotheses about network structure. Exponential random graph modeling (ERGM) is one type of statistical network modeling that allows prediction of the probability of the observed overall network structure, or the likelihood of any specific link in an observed network. ERGM provides results with interpretations similar to logistic regression. One variation of ERGM allows the development of latent position cluster models, which estimate the probability of each node belonging to a particular cluster of nodes.

Latent position cluster modeling was used to identify research clusters within the main path. Two measures were examined to select the number of clusters that was the best representation of the observed data: (1) clustering probabilities and (2) Bayesian Information Criterion (BIC). Specifically, probabilities associated with belonging to clusters were calculated for each node in models with two to six clusters and BIC was calculated for each cluster model. Once the optimal number of clusters was identified, articles in each cluster were examined to identify themes.

Evidence Synthesis
The network included 2986 unique documents published between 1993 and 2009. Most documents were journal publications (n = 2808, 94%); however, there were also 88 books (3%) and 90 (3%) miscellaneous document types (e.g., reports, dissertations). Because most documents were journal articles, the term article is used to represent any document in the network in the paragraphs that follow.

More than 2000 unique lead authors (n = 2283) contributed an average of 1.3 (SD = 1.3) articles to the network (range = 1–39). Most authors (n = 1908) had one article in the network. The ten authors contributing the highest number of articles to the network were Russ E. Glasgow (n = 39); Lawrence O. Gostin (n = 25); Nancy Krieger (n = 19); Ross C. Brownson (n = 10); James G. Hodge (n = 9); Glen P. Mays (n = 9); Steven H. Woolf (n = 9); Richard L. Spoth (n = 8); and Bernard J. Turnock (n = 8).

There were 761 unique journals contributing an average of 3.9 (SD = 10.5) articles to the network (range = 1–150). The ten journals contributing the most articles to the network were American Journal of Public Health (n = 150); American Journal of Preventive Medicine (n = 115); British Medical Journal (n = 100); Journal of Public Health Management and Practice (n = 74); Journal of the American Medical Association (n = 55); Journal of General Internal Medicine (n = 49); Diabetes Care (n = 43); Pediatrics (n = 42); New England Journal of Medicine (n = 40); and Health Affairs (n = 39).

Highly cited articles. According to Google Scholar, articles in the network were cited between 0 and 4161 times overall. Within the network, articles were cited between 0 and 387 times by other articles in the network (this number is necessarily lower given that the network is a sample of articles). Table 2 shows the top ten cited articles in the network. Although the top ten articles were not dominated by any single author or journal, two appeared in Milbank Quarterly and two were first-authored by Glasgow. Overall, each article in the network cited an average of 1.3 (SD = 0.9) other articles in the network (range = 0–13).

The main path. In order to determine which articles constituted the backbone of the field, the main path was identified using the mean traversal weight plus 1 SD (0.035). The main path included 108 articles with 61 different lead authors published in 33 journals and four books between 1994 and 2009. The main path included seven of the seeds along with articles and books from all three search levels. On average, main-path articles were cited 12.5 times within the network (SD = 16.8), which was significantly more often than the 0.9 (SD = 8.1) times for non-main-path articles (t = –7.2, p < 0.001). The Journal of Public Health Management and Practice contributed 22 articles (20%) to the main path, whereas the American Journal of Preventive Medicine and the American Journal of Public Health each contributed 12 articles (11%) despite having higher frequencies in the overall network. Glasgow, Gostin, and Mays contributed the most articles to the main path with 12, 10, and 8 articles, or 11%, 9%, and 7%, respectively. Figure 1a shows the main path with nodes sized by their traversal weight.
Node colors show clustering based on latent position cluster modeling (see below). A list of the articles in the main path grouped by cluster is included as in Appendix B (available online at www.ajpmonline.org).

**Latent position cluster modeling.** Latent position cluster modeling was used to identify subfields. Graphically, the best fit appeared to be the four-cluster model (Appendix C, available online at www.ajpmonline.org). Probabilities from each model were compared to the cluster BIC for the model and it was found that the four-cluster model had the best combination of high probability of cluster membership and low BIC (Appendix D, available online at www.ajpmonline.org).

Article characteristics and titles were examined by cluster to identify patterns related to how authors, journals, and topics make up the four clusters. Cluster 4 appears to be a “core” set of PHSSR articles cited by authors in Clusters 1–3. The core articles focus primarily on local health departments, including the assessment of effectiveness, the role of managed care in local public health, and the role of community partnerships in public health. Seven (23%) of the 31 core cluster articles were published in the *Journal of Public Health Management and Practice*; five (16%) in *Public Health Reports*; and three each (10%) in the *Journal of the American Medical Association* and the *American Journal of Public Health*. Turnock published five (16%) of the articles in this cluster, whereas Handler published four (13%). Articles in the core were generally older than articles in other clusters and did not appear to represent an active area of PHSSR, but instead represented a common ancestry or bridge connecting the three more current clusters of research.

Cluster 1 consisted of articles examining the National Public Health Performance Standards Program (NPHPSP). Of the 34 articles in Cluster 1, there were 14 (41%) published in the *Journal of Public Health Management and Practice*; six (18%) in the *American Journal of Preventive Medicine*; and five (15%) in the *American Journal of Public Health*. Only three authors contributed more than one article to this cluster: Mays (n=6); Leischow (n=2); and Green (n=2).

Cluster 2 (n=18) was made up almost exclusively of publications from public health law with ten of these being articles published by Gostin, a well-known public health law scholar. Six (33.3%) of the articles in Cluster 2 were published in the *Journal of the American Medical Association*; three (17%) in the *Journal of Law, Medicine, and Ethics*; and two (11%) in the *American Journal of Public Health*. Other than Gostin, no author was represented multiple times in this cluster.

Cluster 3 (n=25) included numerous articles on different strategies for translating and utilizing...
behavioral interventions. Five (20%) of the articles in this cluster were published in *Annals of Behavioral Medicine*, whereas several other journals contributed two articles. There were 12 articles (48%) by Glasgow in Cluster 3, whereas Greenhalgh and Green each authored two Cluster-3 articles.

There was a significant difference ($F(3, 104) = 42.6, p < 0.001$) between clusters in the average cluster-membership probability, indicating that some clusters had a more clear membership than others. Figure 1a shows the constituents in each of the four main-path clusters based on the highest cluster probability for each node. Nodes in the core cluster (Cluster 4) had the lowest average probability (59%) of correct assignment, whereas nodes in the behavioral intervention cluster (Cluster 3) had the highest (91%). The NPHPSP cluster (Cluster 1) was the largest ($n = 34$), whereas the public health law cluster (Cluster 2) was the smallest ($n = 18$). When the clusters were collapsed into single nodes, Cluster 4 (black) emerged as a core cluster in the network, being cited by (but not citing) articles in each of the clusters.

Figure 1. Citation network main path
Note: (a) Nodes sized by traversal weight and node colors based on the four latent position clusters identified through exponential random graph modeling; (b) collapsed to show linkages between clusters

July 2011
the other three clusters. With the exception of one citation link between the NPHPSP cluster and the public health law cluster, there were no other cross-cluster citations aside from citations of articles in the core. The collapsed network (Figure 1b) shows the total number of citation linkages between clusters.

**Discussion**

Although there have been calls for the development of a consistent identity and direction for PHSSR, this study is the first to examine the current state of PHSSR empirically. Using a newly developed citation network strategy, the identity struggle PHSSR faces was demonstrated. The large number of lead authors and journals, along with the diversity of journals that figure prominently (e.g., *Health Affairs, Diabetes Care, Journal of Public Health Management and Practice*), shows the range of theories, topics, and methods contributing to PHSSR.

Based on citation patterns, PHSSR appears to be a multidisciplinary field with research happening in parallel across different disciplines. The lack of cross-cluster citation among Clusters 1–3, despite sharing the common ancestry of Cluster 4, may point to an important opportunity for PHSSR to integrate areas such as NPHPSP standards (Cluster 1) and behavioral research (Cluster 3). In 2006, Lenaway and colleagues identified national goals for PHSSR, including “Explore the relationship between social determinants of health and system performance” and “Explore the relationship between performance and health outcomes.” Social determinants of health and health outcomes research currently fall largely in PHSSR Cluster 1, whereas system performance appears in Cluster 3. Meeting these goals, and others, would necessitate bringing together PHSSR areas that currently are unconnected.

There are a few limitations with this study. First, CNA moves forward in time to collect the sample of articles. Although articles published in the most recent 5 years were removed, three of the seeds were less than 10 years old. In addition, minor data-quality issues were encountered with the use of Google Scholar, including occasional misspellings of identifying information and multiple editions of some books. Any mistakes found were corrected. Finally, CNA is a new data collection tool and there is still much to learn in using it.

**Conclusion**

Developing new connections among subfields and integrating the many theories and methods used in PHSSR may aid in meeting national goals and developing a clearer identity for the field. One strategy that may facilitate these new connections is a new strategic focus on the development of transdisciplinary PHSSR teams. However, developing a transdisciplinary field of PHSSR would first require shifts in thinking and strategy among those funding PHSSR, conducting PHSSR, and training new researchers in PHSSR.

This research was supported by a mini-grant from Assuring the Future of Public Health Systems & Services Research, a program of the University of Kentucky Center for Public Health Systems and Services Research, funded by the Robert Wood Johnson Foundation.

No financial disclosures were reported by the authors of this paper.

**References**

39. Iglehart JK. The American health care system. Managed care. (0028-4793 (Print)).

Appendix

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.amepre.2011.03.015.