

Evolution of Coauthorship in Public Health Services and Systems Research

Michael E. Bales, MPH, PhD, Stephen B. Johnson, PhD, Jonathan W. Keeling, MS, Kathleen M. Carley, PhD, Frank Kunkel, MS, Jacqueline A. Merrill, RN, MPH, DNSc

Context: Public health services and systems research (PHSSR) focuses on the structure, organization, and legal basis of domestic public health activities and their effect on population health. An accurate description of the field is needed to empower funding agencies and other stakeholders to coordinate PHSSR activities and to foster the development of the field. The purpose of the study is to characterize the emerging community of researchers engaged in PHSSR. This study (1) describes dynamics of this growing community and (2) identifies research themes, subgroups within the field, and collaboration among groups.

Evidence acquisition: Coauthorship network visualization of selected research publications in the MEDLINE bibliographic database between 1988 and May 2010.

Evidence synthesis: PHSSR has emerged gradually with noticeable growth after 1994 and after 2004. The network of PHSSR research has a core-periphery structure. The core includes highly collaborative researchers focusing on topics pertaining directly to PHSSR, such as workforce, quality improvement and performance, law, and information infrastructure. The periphery consists of groups publishing either on general health services research topics or on epidemiologic and clinical topics.

Conclusions: Although a nucleus group of productive and engaged individuals participate in PHSSR, most also publish broadly on health services research and population health. This trend suggests that this emerging field cannot yet support a singular focus on PHSSR. Lack of funding sources and defined career paths likely contribute to this pattern. An overview of collaboration in PHSSR is an important step in advancing a coordinated research agenda and attracting sustainable funding streams for this field.

(Am J Prev Med 2011;41(1):112–117) © 2011 American Journal of Preventive Medicine

Context

The public health system is the structure, organization, and legal basis of domestic public health activities.¹ As evidenced by the Prevention and Public Health Fund established in recent health reform legislation,² robust public health services and systems are essential for population health. Yet remarkably little is known about how modern public health systems can best achieve vital population health outcomes.^{3–7}

In response, a new field of inquiry, public health services and systems research (PHSSR) is emerging to supply the needed evidence. PHSSR is focused on the factors that contribute to system performance, including organization, financing, and delivery of public health services and their impact.^{8–10} Researchers in the field formulate, translate, and apply research evidence to guide system-wide improvement.¹¹ Although similar inquiry is found in the literature from the early 1900s, rapid growth in the field, especially since 2004, suggests an emerging community of practice.

Communities of practice are critical to the growth and maturity of any field.^{12,13} An overview of the scholarly activity in a community of practice allows stakeholders to understand the composition of a community, as well as factors influencing its development. Researchers may find such knowledge useful for seeking out experts and potential collaborators, or for identifying research topics. Policymakers may benefit when planning or fostering

From NMBI Systems (Bales), the Department of Biomedical Informatics (Johnson, Keeling, Merrill), Columbia University, New York, New York; and the School of Computer Science, Carnegie Mellon University (Carley, Kunkel), Pittsburgh, Pennsylvania

Address correspondence to: Jacqueline A. Merrill, RN, MPH, DNSc, Department of Biomedical Informatics, College of Physicians and Surgeons, Columbia University, Vanderbilt Clinic, 5th Floor, 622 West 168th Street, New York NY 10032. E-mail: jacqueline.merrill@dbmi.columbia.edu.

0749-3797/\$17.00

doi: 10.1016/j.amepre.2011.03.018

collaboration among groups. Funding agencies may gain insight into the dynamics of a field, thereby positioning themselves to identify research agendas for a given domain.

The objectives of this study are (1) to describe dynamics in the emergence of the PHSSR community; (2) to identify research themes, subgroups within the field, and collaboration among groups; and (3) through characterizing the field, to empower stakeholders to foster robust development.

Network Analysis and Visualization

A network analysis of coauthorship based on citation data was used to provide insights into the growth of the PHSSR domain. The networks consist of nodes and links. Nodes represent authors, publications, and journals. Links connect nodes, representing relationships such as authorship (links from authors to publications) or coauthorship (links from authors to authors).

The science of mapping knowledge domains was first described in 2004.¹⁴ A variety of network typologies can be created from citation data. Among the most common are coauthorship and co-citation networks.¹⁵ A layout algorithm is applied to position the nodes spatially. Mathematical equations can be applied to measure the structure of the network or to compare networks. Network analysis and visualization have been used to describe the structure of a variety of fields.¹⁶⁻²²

The current work is distinguished from prior work in one key respect: acquisition of data. The networks are similarly derived from queries of a bibliographic database. However, the goal for the current study was to retrieve all articles written by specific authors, regardless of topic. Traditional keyword-based approaches are not suitable for this, requiring us to generate a technique to identify articles written by selected individuals.

ReCiter Program for Author Name Disambiguation

A refinement was made of the ReCiter algorithm that addresses the problem of nonstandardized author names in MEDLINE (SBJ et al., unpublished observations, 2010). ReCiter is unique among tools for disambiguating author names in MEDLINE (e.g., in Torvik²³) in that its search results are always up-to-date. ReCiter downloads records in response to a query, taking a last name and first name or first initial as input, optionally with middle initial, article titles, journal names, and MeSH keywords. To maximize recall, the system conducts a general search of MEDLINE using only author name and retrieves all matching articles. The resulting list is partitioned into groups, each corresponding to a different author identity. To maximize precision, the algorithm selects the author

group that best matches the terms in the input, producing a list of PubMed IDs authored by the target individual.

Sciologer Platform for Social Network Analysis and Visualization

The software used in this study is Sciologer, a multipurpose platform for exploratory network analysis and visualization.^{24,25} Sciologer can represent multiple types of nodes using different icons. When applied to bibliographic data, Sciologer generates network diagrams of authors, publications, institutions, journals, keywords, common terms, and grants.

Evidence Acquisition

Identification was made of publications produced by a subset of respondents to a survey of 2067 individuals involved in PHSSR (JAM et al., unpublished observations, 2011). These individuals were identified through participation in PHSSR meetings, conferences, or other activities supported by the Robert Wood Johnson Foundation (RWJF), the primary institutional funder for PHSSR at this time. The response rate was 41%. Ninety of 742 respondents did not consider themselves to be members of a PHSSR community.

From the 652 remaining respondents, names of the most productive and engaged were selected based on survey responses matching at least three of four criteria during the past 3 years: (1) authored a PHSSR publication; (2) received funding for PHSSR; (3) presented PHSSR at meetings; and (4) shared resources (i.e., data, staff, or personnel) with key groups (e.g., the Association of State and Territorial Health Officials, the National Association of County and City Health Officials, the National Public Health Performance Standards Program, the Public Health Accreditation Board, the Public Health Foundation, or the University of Kentucky Center for PHSSR). These criteria were met by 133 respondents, representing the nucleus of the PHSSR community. On June 18, 2010, a ReCiter was used to generate a list of MEDLINE articles authored or coauthored by these 133 people. The output was used in Sciologer to visualize structure and development in PHSSR as exemplified by the work of these individuals.

Coauthorship Network

The first network captured scholarly output from before 1988 through 2010, using 1950 as a cutoff. Sciologer was used to explore the titles of all articles in each visually discernable cluster. If at least two articles were thematically related, terms associated with the theme were assigned to the cluster. If no two articles were thematically related, or if the cluster consisted of only one article with

many coauthors, a theme was not assigned. The themes were validated with expert opinion and each was assigned a descriptive label.

Development of the Community over Time

A second network captured development over time intervals corresponding to key events in the emergence of the field: (1) publication of the influential *Future of Public Health* in 1988; (2) publication of the *Essential Services of Public Health* in 1994; (3) the terrorist attacks in 2001; and (4) start of a 10-year initiative to support PHSSR by the RWJF in 2004.^{26,27} The interval beyond 2004 was divided resulting from a sharp increase in scholarly output. Thus, the network shows coauthorship over six time periods: before 1988, 1989–1994, 1995–2001, 2002–2004, 2005–2007, and 2008–2010.

Analysis of Journals Reflecting Research Domains

A third network captured groupings of PHSSR authors based on journal preference. To this network, links were added between authors and journals in which they have published. Groups of thematically related journals were identified and the network was labeled with those themes.

Evidence Synthesis

Coauthorship Network of a Nucleus of Productive and Engaged Survey Respondents

Based on the output of ReCiter, 118 of the 133 most productive and engaged survey respondents (88.7%) had authored or coauthored at least one article indexed in MEDLINE through June 17, 2010. Fifteen people had not authored or coauthored an article in MEDLINE. These individuals are likely involved in nonscholarly activities, for example as administrators or program directors. The number of articles totaled 2344, with the number published per author ranging from 1 to 231 ($M=19.9$, $SD=34.7$). Table 1 describes the outcome of this analysis.

Coauthorship Network Through 2010

Figure 1 shows the coauthorship network of publications by the most productive and engaged survey respondents from before 1988 to 2010. The image on the left includes 2344 publications, representing the work of 118 unique authors and their coauthors. Topic labels were added manually. Positions of nodes are determined by a force-directed placement algorithm that positions linked nodes closer together in space.²⁸ Node coloring is based on a three-dimensional color space in which nearby nodes are assigned similar shades. While much of the network consists of dense clusters dominated by the work of individuals within research groups or departments, the network

Table 1. Analysis of publications for the most productive and engaged survey respondents

Total survey respondents matching three of four criteria: (1) published, (2) funded, (3) presented at meetings, (4) shared resources	133
Author-name MEDLINE queries that retrieved at least one result ^a	124
Queries for which ReCiter selected at least one article	115
Respondents for whom articles were added after manual review of ReCiter output	3
Respondents represented in the network, excluding coauthors	118
Average papers per respondent	20
Median papers per respondent	8

^aSome queries returned results for other authors with similar names.

core is made up of individuals from multiple institutions. The inset view of the network in Figure 1 shows an oval that signifies the network core, an area with the most intense collaboration on PHSSR topics including public health preparedness, performance, and law.

The network's core is discernable but not visually dense—the densest clusters are on the periphery, sometimes around authors who have published dozens of single-author papers, such as a series of commentaries. In other cases dense clusters occur around many authors at a single institution, often a research lab or department. By contrast, scholarly collaboration in the core appears to be among authors drawn together by PHSSR, rather than institutional affiliation. Authors are drawn together into the core when the force-directed placement algorithm is applied because they are linked to PHSSR articles. Links to coauthored articles on other topics simultaneously pull authors away from the center. The result is a coarse-grained mesh of PHSSR at the core of the network, resembling a stretched fishing net.

Key Subgroups in the Coauthorship Network

As shown in Figure 1a and 1b, the network includes a number of visually identifiable subgroups. Some are formed around the work of one highly prolific author, rather than a specific concept. The periphery in Figure 1a includes research on a variety of public health–related concepts. These pertain mainly to population health. The core includes specific PHSSR concepts: *Preparedness*, *Performance*, *Workforce*, and *Law*.

Supplemental Networks

Figure A-1 (Appendix A, available online at www.ajpmonline.org) shows the growth of the community based on coauthorship at six time intervals showing how

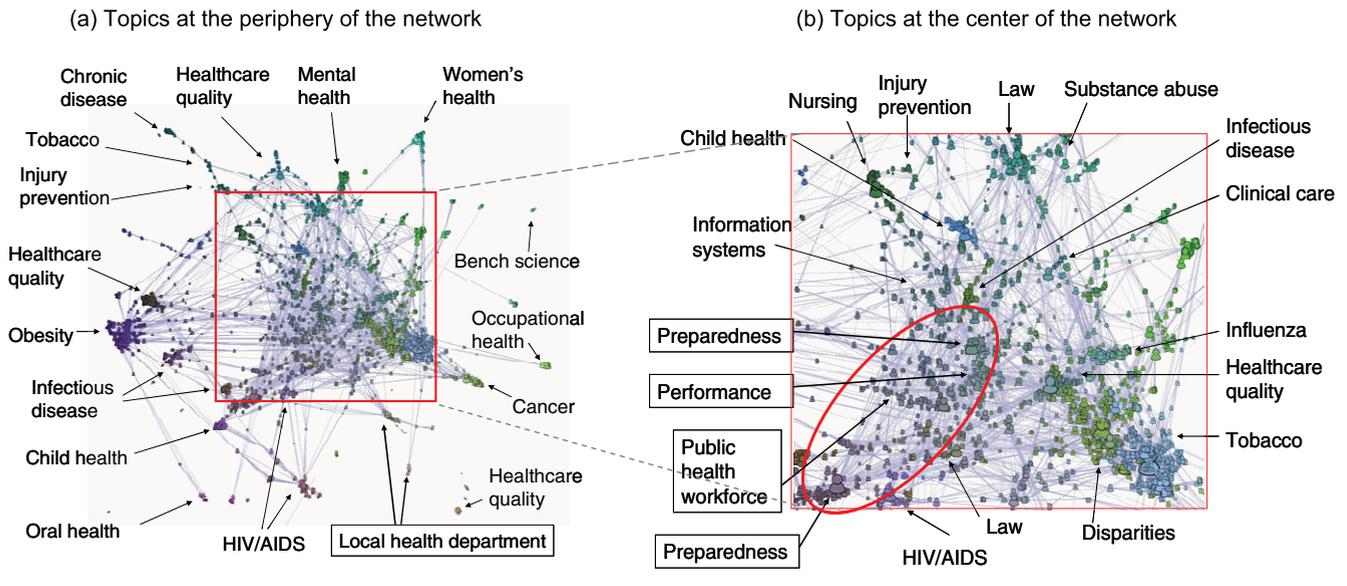


Figure 1. Coauthorship network of 2344 research articles on PHSSR, from before 1988–2010, representing the work of 118 unique authors and their coauthors

Note: Nodes are authors or publications. Authors are linked to their coauthors and to papers they have written. Node size reflects number of links. Node shading is based on a three-dimensional color space in which nearby nodes are assigned similar colors. Topics pertaining most directly to PHSSR appear in boxes.

PHSSR, public health services and systems research

the scholarly output of the PHSSR community has developed.

During this time, the 118 authors published articles in a total of 490 unique journals. Table 2 (available online at www.ajpmonline.org) gives the titles of 20 journals in which these authors published most frequently. Figure A-2 (Appendix A, available online at www.ajpmonline.org) is a coauthorship network of 118 authors showing only journals.

Discussion

In this analysis, PHSSR is described as an emerging, yet discernable community of practice. The results have several limitations that bear discussion. First, the networks provide only a partial overview of all PHSSR research from before 1988 to 2010. There may be researchers who do not appear because they did not participate in the survey, or because their publications are not indexed in MEDLINE (e.g., technical reports, grey literature, or work indexed in other bibliographic databases). Second, although network labels were validated, preliminary labeling was done by only one rater. As a result, some labels might differ from those that would emerge from a consensus-based process involving multiple raters. That said, this research is a first step, and the ratings of one expert provide a sufficient overview at this stage.

The network for all years (before 1988 through 2010) is composed of a core–periphery structure in which the topics related most directly to PHSSR are drawn toward

the core. At the network’s visual center are recent publications in which “public health services and systems research” occurs in the title. Immediately outside the core, loosely distributed groups publish on a variety of topics, including health services and population health. The resulting structure describes a core of PHSSR with trajectories toward peripheral clusters of authors and publications on decreasingly related topics.

The community is centered on a nucleus of individuals from multiple institutions engaged in research related directly to performance and infrastructure. However, there is little evidence from the current analysis of a collaborative network sustained over time. This suggests that a focused trajectory of productivity is yet to be achieved. Individuals involved in PHSSR publish broadly on health services research and population health. This suggests that this emerging field cannot yet support a singular focus on PHSSR. This may change if PHSSR matures into a well-formed research discipline with dedicated funding and infrastructure such as academic fellowships or endowed professorships.

The growth in the coauthorship network (Figure A-1, available online at www.ajpmonline.org) must be interpreted in the context of the data and methods. First, because scholarly output across science is growing exponentially, growth in any community is expected. To understand whether PHSSR has grown more rapidly than science as a whole would require further research. Second, because the data used to produce the networks were

Table 2. The journals in which a nucleus of 118 authors published most frequently from before 1988 through 2010

Frequency	Journal
197	<i>Journal of Public Health Management and Practice</i>
127	<i>Wisconsin Medical Journal</i>
78	<i>American Journal of Public Health</i>
61	<i>The AIDS Reader</i>
61	<i>Health Affairs (Millwood)</i>
59	<i>JAMA: The Journal of the American Medical Association</i>
49	<i>Public Health Reports</i>
41	<i>American Journal of Preventive Medicine</i>
39	<i>Health Services Research</i>
35	<i>Nursing Economic\$</i>
29	<i>Journal of General Internal Medicine</i>
28	<i>Journal of Law, Medicine & Ethics: A Journal of the American Society of Law, Medicine & Ethics</i>
23	<i>Medical Care</i>
22	<i>The New England Journal of Medicine</i>
22	<i>Pediatrics</i>
18	<i>Public Health Nursing</i>
16	<i>Annual Review of Public Health</i>
16	<i>The Journal of the Kentucky Medical Association</i>
16	<i>Maternal and Child Health Journal</i>
15	<i>American Journal of Managed Care</i>

Note: Journals less associated with public health such as *Wisconsin Medical Journal*, the *AIDS Reader*, *Nursing Economics*, and *Pediatrics* are evidence that these authors are active in fields other than public health services and systems research.

derived from publications of individuals who responded to a survey in January 2010, the networks are centered on the current time. This is another reason one might expect the networks to have more nodes and links in more recent time periods. The networks do not include publications of individuals who may have been very active several years ago. Individuals who did not participate in the survey are included in the network if they coauthored with any of the 133 most productive and engaged respondents.

Conclusion and Implications

Descriptive, exploratory research is a first and important step in providing actionable knowledge to stakeholders

wishing to foster growth and mature collaboration in PHSSR. A number of approaches have been considered to understand collaborative structure and growth. In the future, measurement of network structure may be combined with visual inspection to measure and validate levels of collaboration in networks. A method that quantifies patterns in collaboration over time might allow predictive models of future collaboration in given community of practice.

Funding for this research is provided by the Robert Wood Johnson Foundation.

MEB received consulting fees for conducting this study. JWK is a National Library of Medicine pre-doctoral trainee supported by NLM grant T15-LM007079. KMC received consulting fees for her participation in this study. JAM was the principal investigator of this study, which is part of a larger project to study the field of PHSSR with funding support from the Robert Wood Johnson Foundation.

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

References

- Lister SA. An overview of the U.S. Public Health System in the context of emergency preparedness. Washington DC: Congressional Research Service, 2005.
- Patient Protection and Affordable Care Act. Prevention and Public Health Fund, U.S.; 2010.
- Beaulieu J, Schuttfeld FD, Kelly AV. Recommendations from testing of the National Public Health Performance Standards Instruments. *J Public Health Manag Pract* 2003;9(3):188-97.
- Handler A, Issel M, Turnock B. A conceptual framework to measure performance of the public health system. *Am J Public Health* 2001; 91(8):1235-9.
- Lenaway D, Halverson P, Sotnikov S, Tilson H, Corso L, Millington W. Public health systems research: setting a national agenda. *Am J Public Health* 2006;96(3):410-3.
- Mays GP, Smith SA, Ingram RC, et al. Public health delivery systems: evidence, uncertainty, and emerging research needs. *Am J Prev Med* 2009;36(3):256-65.
- Salinsky E, Gursky S. The case for transforming governmental public health. *Health Aff* 2006;25(4):1017-28.
- AcademyHealth. Public Health Systems Research Interest Group. www.academyhealth.org/content.cfm?ItemNumber=1128&navItemNumber=2035.
- Lenihan P. The public health system: an idea whose time has come. *J Public Health Manag Pract* 2005;11(2):165-7.
- Scuttfeld FD, Marks JS, Perez DJ, Mays GP. Public health services and systems research. *Am J Prev Med* 2007;33(2):169-71.
- Scuttfeld FD, Bhandari MW, Lawhorn NA, Lamberth CD, Ingram RC. Public health performance. *Am J Prev Med* 2009;36(3):266-72.
- Merrill JA, Hripcsak G. Using social network analysis within a department of biomedical informatics to induce a discussion of academic communities of practice. *J Am Med Inform Assoc* 2008;15(6):780-2.
- Wenger EC, Snyder WM. Communities of practice: the organizational frontier. *Harv Bus Rev* 2000;78(1):139-45.
- Shiffrin RM, Borner K. Mapping knowledge domains. *Proc Natl Acad Sci U S A* 2004;101:5183-5.

15. Godin B. On the origins of bibliometrics. *Scientometrics* 2006; 68(1):109–33.
16. Bales ME, Johnson SB, Weng C. Social network analysis of interdisciplinarity in obesity research. Washington DC: American Medical Informatics Association Annual Symposium, 2008.
17. Bassecouard E, Lelu A, Zitt M. Mapping nanosciences by citation flows: a preliminary analysis. *Scientometrics* 2007;70(3):859–80.
18. Boyack KW, Klavans R, Borner K. Mapping the backbone of science. *Scientometrics* 2005;64(3):351–74.
19. Eggers S, Huang Z, Chen H, et al. Mapping medical informatics research. In: Chen H, Fuller SS, Friedman C, Hersh W, eds. *Medical informatics: knowledge management and data mining in biomedicine*. New York: Springer, 2005.
20. Klavans R, Boyack KW. Toward a consensus map of science. *J Am Soc Inform Sci Technol* 2009;60(3):455–76.
21. Reid EF, Chen H. Mapping the contemporary terrorism research domain. *Int J Hum Comput Stud* 2007;65(1):42–56.
22. Watts DJ. The “new” science of networks. *Annu Rev Sociol* 2004; 30:243–70.
23. Torvik VI, Smalheiser NR. Author name disambiguation in MEDLINE. *ACM Trans Knowl Discov Data* 2009;3(3):11.
24. Bales ME. *Sciologer: visualizing and exploring scientific communities*. New York NY: Columbia University, 2009.
25. Bales ME, Kaufman DR, Johnson SB. Evaluation of a prototype search and visualization system for exploring scientific communities. San Francisco CA: American Medical Informatics Association Annual Symposium, 2009.
26. IOM. *The future of public health*. Washington DC: National Academy of Science Press, 1988.
27. Public Health Functions Steering Committee. *Public health in America: vision, mission, and essential services*. Washington DC: U.S. Public Health Service, 1994.
28. Beyer D. *CCVisu: automatic visual software decomposition*. Leipzig, Germany: International Conference on Software Engineering, 2008.

Appendix

Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.amepre.2011.03.018](https://doi.org/10.1016/j.amepre.2011.03.018).

Did you know?

The *AJPM* online archive includes issues back to 1998.

Visit www.ajpmonline.org today!