Published in final edited form as:

Am J Prev Med. 2011 July ; 41(1): 112–117. doi:10.1016/j.amepre.2011.03.018.

Evolution of Co-authorship in Public Health Systems and Services Research

Michael E. Bales, MPH, PhD,
NMBI Systems, New York, NY

Stephen B. Johnson, PhD,
Department of Biomedical Informatics, Columbia University, New York, NY

Jonathan W. Keeling, MS,
Department of Biomedical Informatics, Columbia University, New York, NY

Kathleen M. Carley, PhD,
School of Computer Science, Carnegie Mellon University, Pittsburgh PA

Frank Kunkel, MS, and
School of Computer Science, Carnegie Mellon University, Pittsburgh PA

Jacqueline A. Merrill, RN, MPH, DNSc
Associate Research Scientist, Department of Biomedical Informatics, College of Physicians and Surgeons, Columbia University, Vanderbilt Clinic, 5th Floor, 622 West 168th St., New York, NY 10032, Phone: 212-305-3194, Fax: 212-342-0663

Jacqueline A. Merrill: jacqueline.merrill@dbmi.columbia.edu

Abstract

Background—Public health systems and services research (PHSSR) examines the organization, financing, and delivery of public health services and the impact of these activities on population health. An accurate description of this PHSSR is needed to empower funding agencies and other stakeholders, to coordinate PHSSR activities, and to foster the development of the field.

Purpose—To characterize the emerging community of researchers engaged in PHSSR. This study 1) describes dynamics of this growing community; and 2) identifies distinct topics being researched, communities of practice within PHSSR, and collaboration among groups.

Methods—Co-authorship network visualization of selected research publications in the Medline bibliographic database.

Results—PHSSR has emerged gradually since 1988, with noticeable growth after 1994 and after 2004. The network of PHSSR research has a core-periphery structure. The core of this network includes highly collaborative researchers focusing on topics pertaining directly to PHSSR, such as the public health workforce, quality improvement and performance, law, and information infrastructure. The periphery consists of groups publishing either more generally on various health services research topics, or on epidemiologic, clinical, or health sciences topics.

Conclusions—While a nucleus group of productive and engaged individuals participate in PHSSR, most authors are also involved in general health services research, issues of population health, or health science topics unrelated to PHSSR. An overview of collaboration in PHSSR is an important step in advancing a coordinated research agenda and attracting sustainable funding streams for this field.

Background

The public health system is the structure, organization, and legal basis of domestic public health activities.¹ The 21st century public health system is confronted with the emergence of new pathogens, terrorism, natural disasters, and unprecedented demographic, political, and socioeconomic shifts.^{2, 3} Changes in federal immigration laws and population aging require assessment and policy development to assure rapidly changing public health needs are met.⁴ Attrition and budget limitations have reduced the local public health work force by over 15% during the last few years, just as experts are calling for an increase in the preventive and monitoring services that these workers provide.^{5–7} These circumstances challenge health officials, planners, and policymakers at all levels of government to make the most effective, efficient, and equitable use of scarce public resources. Yet remarkably little is known about how modern public health systems can best achieve vital population health outcomes.^{8–12}

In response to the dearth of empirical evidence a new field of inquiry, **Public Health Systems and Services Research (PHSSR)**, is emerging focused on the factors that contribute to system performance, including organization, financing, and delivery of public health services and the impact of these activities on population health.^{13–15} PHSSR seeks to formulate, translate and apply research evidence to guide system-wide improvement.¹⁶ Although such inquiry is found in the literature from the early 1900's, rapid growth in the field, especially since 2004, suggests an emerging community of practice that includes researchers, students, public health practitioners, federal agency staff, funders, and policymakers.¹⁷

Communities of practice are among the most important structures of any academic field.^{18, 19} An overview of the scholarly activity of a community of practice allows stakeholders to understand the current composition of a community, as well as the factors influencing its growth and development. Researchers can use the results to understand how their own work fits within the greater context of research in the field. They may find such knowledge useful when seeking out experts and potential collaborators, preparing grants, or identifying research topics. Professional groups and policymakers may benefit when planning development programs or when fostering synergy and collaboration among groups. Funding agencies may gain insight into the dynamics of a field, thereby positioning themselves to identify research agendas that meet emerging needs for knowledge in a given domain.

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

Network analysis and visualization

We use network analysis of co-authorship based on citation data to provide insights into growth of the PHSSR domain. The network consists of nodes and links. Nodes represent authors, publications, and journals. Links connect the nodes, representing relationships such as authorship (links from authors to publications) or co-authorship (links from authors to authors). In this research network analysis is used to produce visual maps of the body of scholarly output for specific individuals publishing in the PHSSR domain.

The science of mapping knowledge domains was first described as a cohesive area of inquiry in 2004.²⁰ There are a variety of types of networks that can be created from citation data. Among the most common are co-authorship and co-citation networks.²¹ To represent

the structure of a domain, a network analyst selects sources of bibliographic data, retrieves a list of articles, and decides what elements will constitute the nodes of the network (e.g. authors and publications), and what constitutes a link between nodes (e.g. author to author or author to publication). The nodes and links are represented in an adjacency list or in a matrix. A layout algorithm is applied to place these data elements in either a two-dimensional or three-dimensional space. This produces a visual representation of the network. A variety of mathematical equations can be applied to measure the structure of the network or to make comparisons between two or more networks. Among these are measures of density (proportion of links that are present to the total number of potential links), centrality (average number of links per node), and measures of prominence of subgroups.²² Network analysis and visualization have been used to elicit information about the structure of a variety of fields, including the overall structure of science.^{23–29}

The current work is distinguished from prior work in one key respect: acquisition of data. As with other work on mapping knowledge domains, our networks are derived from data obtained through queries of a bibliographic database, but to characterize the scholarly work of individuals involved in PHSSR, we wished to retrieve all articles written by a specific group of authors, regardless of topic. These lists, in the form of article identifiers (PubMed IDs or PMIDs), were used to formulate queries. A significant challenge with this approach arises from the lack of standardization of author names in the Medline database. For example, a query for “Smith J” will return articles by Jane A. Smith as well as John B. Smith, among others.

ReCiter program for author name disambiguation

To address the problem of ambiguous names, an algorithm named ReCiter was applied that is designed to identify the unique set of articles authored by a given person.³⁰ ReCiter is unique among tools for disambiguating author names in Medline in that its search results are always up-to-date.³¹ ReCiter downloads records from Medline 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 the last name and first initial only 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

Sciologer is a multi-purpose platform for exploratory network analysis and visualization and was used to produce networks.^{32, 33} Sciologer is designed for use with multiple types of nodes (multipartite networks), each represented using different icons. When applied to bibliographic data, Sciologer generates network diagrams of authors, published articles, institutions, journals, keywords, common terms, grants, and other elements. The system is extensible to other domains, for example, nodes might be various classes of genes and proteins. Sciologer employs a unique three-dimensional approach which assigns similar colors to nearby nodes to help users discern groups visually.

Automap software to extract semantic concepts from publications

AutoMap is software for extraction of network data from natural language texts, developed at Carnegie Mellon. The main functions of AutoMap are to extract, analyze, and compare natural language texts in terms of concepts and themes.^{34, 35} When applied to the titles and abstracts from a set of publications AutoMap can extract content analytic data (words and frequencies) to identify groups of topics and the resulting themes that structure a body of scholarly work.

Methods

In this research network analysis and visualization of co-authorship is used to characterize the structure and development of the PHSSR community. An internet-based survey was conducted of 2067 individuals identified through their participation in PHSSR meetings, conferences, or other activities supported by the Robert Wood Johnson Foundation (RWJF)^a. The response rate was 41%. Ninety of 742 respondents did not consider themselves to be public health systems and services researchers. From the 652 remaining respondents we selected the names of the most productive and engaged based on survey responses that matched at least three of four criteria: 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^b all during the past three years. These criteria were met by 133 survey respondents, representing the nucleus of the PHSSR community.

ReCiter produced a list of all Medline articles authored by these 133 persons. Last name, first name, middle initial, institutional affiliation, ZIP code, and words commonly used in PHSSR articles and journals were input to the ReCiter algorithm. The ReCiter output was used in Sciologer and AutoMap to generate the networks described below.

Co-authorship network

The first network provides an overview of scholarly output from before 1988 through 2010 using 1950 as a cutoff date. Sciologer produced a co-authorship network of publications identified by the ReCiter algorithm. A public health expert (M.B.) visually examined the co-authorship network and extracted short phrases descriptive of topics being researched in clusters of authors and publications. The expert used the Sciologer interface to explore the titles of all articles in each visually discernable cluster. If at least two articles were found to be thematically related, terms associated with the theme were assigned to the cluster. If no two articles were thematically related, or if the cluster was comprised of only one article written by many co-authors, the cluster was not assigned a theme. The themes assigned to each cluster were validated by a second public health expert (J.M.) who substituted concepts generated through the semantic analysis of the titles and abstracts.

Development of the community over time

A second network provides an overview of development over time intervals corresponding to key events in the emergence of the field: 1) publication of an influential report *The Future of Public Health* in 1988; 2) publication of the Essential Services of Public Health in 1994; 3) the World Trade Center and anthrax attacks in 2001; and 4) a ten-year commitment to support PHSSR by the Robert Wood Johnson Foundation in 2004.^{36, 37} Due to a sharp increase in scholarly output the years beyond 2004 were divided into two intervals. Thus, the network is partitioned to show co-authorship over six time periods: before 1988; 1989–1994; 1995–2001; 2002–2004; 2005–2007; and 2008–2010.

Author case studies

The third network provides a snapshot of work being conducted by selected individuals. These were produced by using a single author's list of article PMIDs as input to Sciologer. To produce comparable, information-rich examples, three authors were selected randomly

^aRWJF is the primary institutional funder for PHSSR at this time.

^bThe 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.

from among the nucleus of 133 respondents having authored between 50 and 90 publications. These networks were labeled in the same manner as the co-authorship network.

Analysis of journals reflecting research domains

A fourth network shows groupings of PHSSR authors based on journal preference. We modified the Sciologer linking schema to add links between authors and journals in which they have published. Groups of thematically related journals were identified and the network was labeled to highlight these themes.

Semantic analysis of concepts

A fifth network showing authors in relation to research concepts used Automap software to parse the titles and abstracts of publications retrieved by the ReCiter algorithm. Approximately 28,000 unique terms were extracted after common typo correction. This list was dramatically reduced in a stepwise fashion, first by deletion of stop words (e.g. *and*, *but*, *that*), and terms occurring with ubiquitous frequency (i.e. frequency > 3000). Non-informative adjectives and adverbs (e.g. *robust*, *frequently*) and non-specific nouns (e.g. *study*, *time*, *problem*) were also removed. Bi-grams (e.g. public health) and n-grams (e.g. Robert Wood Johnson) were converted to single concepts. The term set was further reduced using expert opinion (JM, KC) to remove plurals and to generalize or harmonize semantically equivalent terms into categories, such as *anatomical terms*, *clinical processes*, or *anti-microbial agents*.³⁸ The resulting 125 concept categories guided labeling of the co-authorship network visualizations.

These concepts were mapped back to the nucleus of 133 survey respondents. This produced a network in which the nodes represent concepts and authors. Links represent occurrence of the concept in the titles and/or abstracts of that author. The Girvan-Newman grouping algorithm was applied to progressively remove the lowest number of ties between concepts and authors to identify subgroups containing more links than would be expected at random.³⁹ The algorithm identifies subgroups and gives an overall score between 0 and 1, where 30 signifies groups with significantly more links than expected. The concept-to-author groupings produced by Girvan-Newman were reviewed iteratively by the expert team (JM, MB, JK) until consensus on a thematic label for each group emerged.

Results

Co-authorship network of a nucleus of productive and engaged survey respondents

Based on the output of ReCiter, the nucleus of 133 most productive and engaged survey respondents had authored or co-authored a total of 2,344 articles indexed in Medline through June 17, 2010. Of these, 118 (88.7%) published at least one article. Among these individuals, the number of articles published per author ranged from 1 to 231 ($\bar{x}$ = 19.9, SD = 34.7). Table 1 describes the outcome of this analysis.

Co-authorship network before 1988 through 2010

Figure 1 shows the co-authorship network of publications by the most productive and engaged survey respondents from before 1988 to 2010. On the left the image includes 2,344 publications, representing the work of 118 unique authors and their co-authors. Topic labels on the periphery of the network were added manually using terms derived from semantic analysis of title and abstracts. Nodes are either authors or publications. Authors are linked to their co-authors and to the papers they have written. The size of each node is determined by its node degree (number of links to other nodes). Positions of nodes are determined by a force-directed placement algorithm which positions linked nodes closer together in space.⁴⁰ Node coloring is based on a three-dimensional color space in which nearby nodes are

assigned similar colors. While much of the network consists of dense clusters of research dominated by the work of individuals within research groups or departments, the network core is comprised of individuals from multiple institutions. The inset view of the network in Figure 1 shows an oval that signifies the network core, comprised of an area with the most intense collaboration on topics directly related to PHSSR.

Key subgroups in the co-authorship network

As shown in Figure 1a and 1b, the PHSSR community 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 of the network in Figure 1a includes research on a variety of public health related concepts. With the exception of *Healthcare Quality* and *Local Health Dept*, these pertain mainly to population health, rather than to general health services or PHSSR. The core of the network includes specific PHSSR concepts: *Preparedness*, *Performance*, *Public Health Workforce*, and *Law*.

Time-based network visualization shows growth of network over time

Figure 2 shows the growth of the community based on co-authorship at six time intervals showing how the scholarly output of the PHSSR community has developed. In each network, the nodes are either authors or publications. Authors are linked to co-authors and to the papers they have written. The two earliest intervals feature individual groups of authors/co-authors working independently, with few links between groups. Recent intervals show the emergence of the community of practice, with an increasing number of links between groups.

Case studies

The case studies show three individual authors. The leftmost author has published exclusively on managed care and related topics. The author in the middle has published on public health topics including obesity and tobacco, but also health services research and research on public health law and ethics. The rightmost author has published a number of articles on managed care, oral health and depression. The publication histories of these randomly selected authors are illustrative of the trajectories of many of the 118 authors studied.

Journals publishing PHSSR

The 118 authors published articles in a total of 490 unique journals. Table 2 lists the titles of 20 journals in which these authors published most frequently from before 1988 through 2010. The journals represent a wide scope of health topics. Journals in which these authors publish most frequently (*Journal of Public Health Management and Practice* and *American Journal of Public Health*) are particularly associated with public health services and systems. Other journals generally associated with public health are *Public Health Reports*, *American Journal of Preventive Medicine*, *Public Health Nursing*, and *Annual Review of Public Health*. Journals less associated with public health such as *Wisconsin Medical Journal*, the *AIDS Reader*, *Nursing Economics*, and *Pediatrics* are evidence that this group of authors are active in research fields other than PHSSR.

Figure 4 is a co-authorship network of 118 authors from before 1988 to June 2010, showing only journals. Thematically related journals are often spatially proximate to one another because links between authors and journals draw related journals closer together in space. The image shows how individuals in specific communities of practice publish in common journals. Among the identifiable communities are dental care, diabetes, environmental health, infectious disease, medicine, nursing, and women's health.

Six themes derived from semantic analysis

Our analysis of the author to concept network using the Girvan-Newman algorithm produces six groups with a score of 0.32. The groups were labeled by theme. Theme 1: *At-risk Populations* included 12 concepts such as mental health, substance abuse, social equity, and HIV/AIDS. Theme 2: *Disease Prevention (primary, secondary or tertiary prevention)^c* included 17 concepts related to clinical processes and pathophysiology. Theme 3: *Healthcare Quality* included 25 concepts such as hospital, insurance, and cost. Theme 4: *Workforce and Infrastructure* included 13 concepts such as health professional, outcome, and finance. Theme 5: *Performance* included 30 concepts such as organization, jurisdiction, and policy. Theme 6: *Population Health* included 28 concepts such as disease, tobacco, epidemiology, and disparities. Figure 5 shows a matrix representation of this network indicating the density for each theme. The highest densities are in Themes 2, 4, and 5 where about one third of possible links are present between authors and concepts. About one quarter of possible links are present for Themes 1, 3, and 6.

Discussion

This research is a foundational step in characterizing the development of the PHSSR community.

The PHSSR co-authorship network is comprised of a core-periphery structure

The network for all years (before 1988–2010) is comprised of a core-periphery structure in which the topics related most directly to PHSSR are drawn towards the center. The core of the network is comprised of publications that pertain directly to PHSSR, including workforce, performance, and law. At the network's visual center are recent publications in which "public health systems and services research" occurs in the title. Immediately outside the core, loosely-distributed groups publish on a variety of topics, including health services and population health. Topics less related to PHSSR appear more distant from the core. The resulting structure might be described as a core of PHSSR at the center, with trajectories radiating toward peripheral clusters of authors and publications on decreasingly related topics.

Publications at the core of the co-authorship network pertain most directly to PHSSR

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, e.g., a series of commentaries. In other cases dense clusters occur around many authors at a single institution, suggesting a research lab or department. By contrast, scholarly collaboration in the core appears to be among authors drawn together by PHSSR, rather than by a common institutional affiliation. Authors are drawn together into the core because they are linked to PHSSR articles (a "force" draws nodes together when the force-directed placement algorithm is applied). Many co-authored articles on other topics that appear elsewhere in the network. These links simultaneously draw 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.

^cPrimary prevention aims to prevent disease (e.g. immunizations). Secondary prevention is focused on early detection (e.g. screening) and treatment. Tertiary prevention aims to reduce the impact of established disease and restore functioning, typically through clinical care.

Collaboration over time

When examining the networks in each of the three last time intervals it was difficult to identify sustained collaboration among two or more people. Although three-year time intervals may be too short to show sustained collaboration, lack of funding opportunities, lack of resources, or potential lack of awareness about the PHSSR work of others may limit meaningful collaboration between groups.⁴¹

The network of journals illustrates the diverse provenance of the PHSSR community of practice

The network of journals is well-suited to showing the backgrounds of authors in the PHSSR community. The clustering of journals by topic makes it clear that PHSSR practitioners may be affiliated with a number of other distinct communities of practice, including nursing, psychiatry, public health law, and infectious disease epidemiology.

Case studies illustrate variation in the career paths of specific authors

While the case studies offer a portrait of the scholarly output of just three individuals, they illustrate substantial diversity among the authors in the network. They show how individuals who consider themselves PHSSR researchers are also active in a variety of more established areas of inquiry.

Semantic analysis helps define areas of research

The six areas of research defined by the semantic analysis characterize scholarly work at the nucleus of the community. These include at-risk populations and clinical health services, themes not encompassed in the definition of PHSSR. PHSSR is a comparatively new field and these areas of inquiry likely represent research established earlier in the authors' careers. Limited funding for PHSSR encourages inquiry on topics for which there is dedicated funding, such as clinical services, risk behaviors, and chronic diseases. Despite this, higher densities (shown in Figure 5) in the Themes 4 (workforce and infrastructure) and 5 (performance) suggest areas of sustained PHSSR activity.

The density of the networks from 1989–1994 and 1995–2001 (figure 2) increased particularly following the publication of the Essential Services in 1994. After the World Trade Center and anthrax attacks in 2001 a decrease in density suggests a shift in these authors' activities away from research. Growth resumes in 2005–2007 following the RWJF funding initiative. This pattern suggests that productivity in PHSSR is sensitive to external events affecting public health systems, and is stimulated by development of standard frameworks, such as the Essential Services, and by dedicated funding.

Limitations

The results of this research are subject to some limitations. First, the networks provide an incomplete overview of all PHSSR research from before 1988 to 2010. There may be researchers that do not appear in the networks, either because they did not participate in the survey from which the data were derived, or because their publications are not indexed in Medline (e.g., technical reports, grey literature, or work indexed in other bibliographic databases). Second, although the network labels were validated by an expert (J.M.), preliminary labeling was done by only one rater (M.B.) As a result, some labels likely differ from those that might emerge from a consensus-based process involving multiple raters. That said, this research is meant to be a first step in outlining the composition and development of PHSSR, and we feel the ratings of one expert provide a sufficient overview at this stage.

The growth in the co-authorship network must be interpreted in the context of the data and methods. First, because scholarly output across science is growing exponentially, rapid growth in any community is expected. To understand whether PHSSR has grown more rapidly than science as a whole would require further research. Second, since the data used to produce the networks are 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^d.

Future research

Descriptive, exploratory research is a first step in characterizing collaboration in a given community. We have considered a number of approaches to understand collaboration. Measures of network structure may be combined with visual inspection to measure and validate levels of collaboration in networks. A method to identify patterns in collaboration over time might allow predictive models of future collaboration in given community of practice.

Conclusion and Implications

Robust public health systems and services are essential for population health. Efforts to study and improve public health systems are of vital importance as illustrated by the Prevention and Public Health Fund funding established in recent health reform legislation.⁴² In this analysis PHSSR is described as an emerging, yet discernable community of practice. The community is centered on a nucleus of individuals from multiple institutions engaged in research related directly to performance, quality improvement, and infrastructure. However there is little evidence from our analysis of a collaborative network sustained over time, suggesting that a focused trajectory of productivity is yet to be achieved. Members of the PHSSR community publish broadly on health services research and population health suggesting that this emerging field cannot yet support a singular focus on PHSSR. A descriptive characterization of the community is an important step in providing actionable knowledge to stakeholders wishing to foster growth in PHSSR.

Acknowledgments

This research was funded by the Robert Wood Johnson Foundation.

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

References

1. Congressional Research Service. An Overview of the US Public Health System in the Context of Emergency Preparedness. Library of Congress; 2005.
2. Baker EL, Porter J. Practicing management and leadership: creating the information network for public health officials. *J Public Health Manag Pract*. 2005; 11(5):469–73. [PubMed: 16103826]
3. Centers for Disease Control and Prevention. Report on Terrorism Preparedness and Emergency Response (TPER)-Funded Activities. Atlanta: Centers for Disease Control and Prevention; 2009. Public Health Preparedness: Strengthening CDC's Emergency Response.
4. Johnson JH, Sabol BJ, Baker EL. The crucible of public health practice: Major trends shaping the design of the management academy for public health. *Journal of Public Health Management and Practice*. 2006; 12(5):419–425. [PubMed: 16912602]

^dThe 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 co-authored with any of the 133 most productive and engaged respondents.

5. National Association of County and City Health Officials. NACCHO Survey of Local Health Departments' Budget Cuts and Workforce Reductions. 2009.
6. Robert Wood Johnson Foundation and Trust for America's Health. Shortchanging America's Health: A State-By-State Look at How Public Health Dollars Are Spent. 2009 Oct. [cited 2010 March]; Available from: <http://www.rwjf.org/files/research/20091007phspendingpolicyhighlightrevised.pdf>
7. Seelye, K. The New York Times. 2010 Mar 1. Public Health Departments Shrinking, Survey Finds.
8. 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. [PubMed: 11499110]
9. Salinsky E, Gursky S. The case for transforming governmental public health. *Health Affairs*. 2006; 25(4):1017–1028. [PubMed: 16835182]
10. Lenaway D, Halverson P, Sotnikov S, Tilson H, Corso L, Millington W. Public health systems research: Setting a national agenda. *American Journal of Public Health*. 2006; 96(3):410–413. [PubMed: 16449601]
11. Beaulieu J, Schuttfeld FD, Kelly AV. Recommendations from testing of the National Public Health Performance Standards Instruments. *Journal of Public Health Management and Practice*. 2003; 9(3):188–97. [PubMed: 12747315]
12. Mays GP, Smith SA, Ingram RC, Laura J, Racster M, Lamberth CD, Lovely ES. Public Health Delivery Systems: Evidence, Uncertainty, and Emerging Research Needs. *American Journal of Preventive Medicine*. 2009; 36(3):256–65. [PubMed: 19215851]
13. Lenihan P. The public health system: An idea whose time has come. *Journal of Public Health Management and Practice*. 2005; 11(2):165–167. [PubMed: 15711447]
14. Scuttfeld FD, Marks JS, Perez DJ, Mays GP. Public health services and systems research. *American Journal of Preventive Medicine*. 2007; 33(2):169–171. [PubMed: 17673106]
15. AcademyHealth. Public Health Systems Research Interest Group. 2010 Jul 6. 2010 Available from: <http://www.academyhealth.org/content.cfm?ItemNumber=1128&navItemNumber=2035>.
16. Scuttfeld FD, Bhandari MW, Lawhorn NA, Lamberth CD, Ingram RC. Public Health Performance. *American Journal of Preventive Medicine*. 2009; 36(3):266–72. [PubMed: 19215852]
17. Merrill J, Btoush R, Gupta M, Gebbie K. A History of Public Health Workforce Enumeration. *Journal of Public Health Management and Practice*. 2003; 9(6):459–70. [PubMed: 14606184]
18. Wenger, EC.; Snyder, WM. *Harvard Business Review*. 2000. Communities of practice: The organizational frontier.
19. Merrill JA, Hripcsak G. Using social network analysis within a department of biomedical informatics to induce a discussion of academic communities of practice. *Journal of the American Medical Informatics Association*. 2008; 15(6):780–782. [PubMed: 18756000]
20. Shiffrin RM, Borner K. Mapping knowledge domains. *Proceedings of the National Academy of Sciences*. 2004; 101:5183–5185.
21. Godin B. On the origins of bibliometrics. *Scientometrics*. 2006; 68(1):109–133.
22. Wasserman, S.; Faust, K. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press; 1994.
23. Eggers, S.; Huang, Z.; Chen, H.; Yan, L.; Larson, C.; Rashid, A., et al. Mapping medical informatics research. In: Chen, H.; Fuller, SS.; Friedman, C.; Hersh, W., editors. *Medical Informatics: Knowledge Management and Data Mining in Biomedicine*. Springer US; 2005.
24. Bales, ME.; Johnson, SB.; Weng, C. Social network analysis of interdisciplinarity in obesity research. *American Medical Informatics Association Annual Symposium*; 2008; 2008.
25. Reid EF, Chen H. Mapping the contemporary terrorism research domain. *International Journal of Human-Computer Studies*. 2007; 65(1):42–56.
26. Bassecoulard E, Lelu A, Zitt M. Mapping nanosciences by citation flows: A preliminary analysis. *Scientometrics*. 2007; 70(3):859–880.
27. Boyack KW, Klavans R, Borner K. Mapping the backbone of science. *Scientometrics*. 2005; 64(3): 351–374.

28. Klavans R, Boyack KW. Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*. 2009; 60(3):455–476.
29. Watts DJ. The “new” science of networks. *Annu Rev Sociol*. 2004; 30:243–270.
30. Johnson SB, Bales ME, Dine D. ReCiter: A tool for identifying a target author’s unique list of Medline publications. *J Am Med Inform Assn*. 2010 (In preparation).
31. Torvik VI, Smalheiser NR. Author name disambiguation in Medline. *ACM Transactions on Knowledge Discovery from Data*. 2009; 3(3):11. [PubMed: 20072710]
32. Bales, ME. *Sciologer: Visualizing and Exploring Scientific Communities*. New York, NY: Columbia University; 2009.
33. Bales, ME.; Kaufman, DR.; Johnson, SB. Evaluation of a prototype search and visualization system for exploring scientific communities. *American Medical Informatics Association Annual Symposium*; 2009; San Francisco, CA. 2009.
34. Carley, KM.; Columbus, D.; Bigrigg, M.; Kunkel, F. *AutoMap User’s Guide*. Pittsburgh: Carnegie Mellon University, School of Computer Science, Institute for Software Research; 2010. CMU-ISR-10–121Report No.: CMU-ISR-10–121
35. Diesner J, Carley KM. Conditional Random Fields for Entity Extraction and Ontological Text Coding. *Journal of Computational and Mathematical Organization Theory*. 2008; 14(3):248–262.
36. Institute of Medicine. *The Future of Public Health*. Washington, D.C: National Academy of Science Press; 1988.
37. Public Health Functions Steering Committee. *Public health in America: vision, mission, and essential services*. Washington, D. C: US Public Health Service; 1994.
38. Lewis, ET.; Diesner, J.; Carley, KM. *Using Automated Text Analysis to Study Self-Presentation Strategies*. Pittsburgh, PA: Carnegie Mellon University; 2001 Jul 12.
39. Girvan M, Newman MEJ. Community structure in social and biological networks. *Proceeding of the National Academy of Science*. 2002; 99(12):7821–26.
40. Beyer, D. *CCVisu: Automatic visual software decomposition*. International Conference on Software Engineering; Leipzig, Germany. 2008.
41. Merrill JA, Wilson RV, Keeling JW, Chen TV. *Public Health Services and Systems Research: Mapping an Evolving Community of Practice*. *AJPM*. in press.
42. Patient Protection and Affordable Care Act. *Prevention and Public Health Fund*; US: 2010.

a. Topics at the periphery of the network

b. Topics at the core of the network

Figure 1.

Co-authorship network of PHSSR research, from before 1988–2010, with topics labeled at the periphery and core. The image includes 2,344 research articles on PHSSR, representing the work of 118 unique authors and their co-authors. In the network, nodes are either authors or publications. Authors are linked to their co-authors and to the papers they have written. The size of each node is determined by its node degree (number of links to other nodes.) Positions of nodes are determined by a force-directed placement algorithm which positions linked nodes closer together in space. Node shading is based on a three-dimensional color space in which nearby nodes are assigned similar colors. Annotation of research areas is done manually based on expert review.

Figure 2.

Networks representing the publications of 118 most engaged and productive survey respondents over a six selected time intervals illustrating the growth of scholarly output. In each network, the nodes are either authors or publications. Authors are linked to co-authors and to papers they have written. Colors change across time periods. A group in a dark shade in one time period may not correspond to a dark-shaded group in the next time slice.

Figure 3.

Three networks representing the career paths of specific authors illustrating the variety of work being conducted by individuals in the community. Networks were produced by using an individual author's list of publication identifiers as input to Sciologer. Each network was labeled in the same manner as the co-authorship network.

Figure 4.

Co-authorship network of PHSSR research, from before 1988 to June 2010, showing only journals. Journal labels are grouped to illustrate communities of practice in the network.

Figure 5.

A matrix representing a network with 125 concepts on the x-axis and of 118 authors on the y-axis from the period before 1988 to 2010. The concepts are divided into six themes: 1) At-risk Populations, 2) Disease Prevention (primary, secondary, tertiary), 3) Healthcare Quality, 4) Workforce and Infrastructure, 5) Performance, and 6) Population Health. Rows are sorted by cumulative author link weight, greatest at bottom, and columns are sorted within each theme by cumulative concept link weight, greatest on the left. Cells are shaded according to strength of the link between author and concept, darkest being strongest. Density is the proportion of links between authors and concepts that are present compared to all that are possible. The highest densities are in Themes 2, 4, and 5 where about one third of possible links are present between authors and concepts. About one quarter of possible links are present for Themes 1, 3, and 6.

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 1 result	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 co-authors	118
Average papers per respondent	20
Median papers per respondent	8

Table 2

The journals in which a nucleus of 118 published authors in the Public Health Systems and Services Research community public most frequently in the period before 1988 through 2010.

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