Development of a Public Health Nursing Data Infrastructure

Karen A. Monsen, Ph.D., R.N.,1 Betty Bekemeier, Ph.D., M.P.H., R.N.,2 Robin P. Newhouse, Ph.D., R.N.,3 and F. Douglas Scutchfield, M.D.4

1School of Nursing, University Of Minnesota, Minneapolis, Minnesota; 2School of Nursing, University of Washington, Seattle, Washington; 3School of Nursing, University of Maryland, Baltimore, Maryland; and 4Peter P Bosomworth Professor of Health Services Research and Policy, College of Public Health and Medicine, Lexington, Kentucky

Correspondence to:
Karen A. Monsen, School of Nursing, University Of Minnesota, 5-140 Weaver-Densford Hall, 308 Harvard Street SE, Minneapolis, MN 55455. E-mail: mons0122@umn.edu

ABSTRACT An invited group of national public health nursing (PHN) scholars, practitioners, policy-makers, and other stakeholders met in October 2010 identifying a critical need for a national PHN data infrastructure to support PHN research. This article summarizes the strengths, limitations, and gaps specific to PHN data and proposes a research agenda for development of a PHN data infrastructure. Future implications are suggested, such as issues related to the development of the proposed PHN data infrastructure and future research possibilities enabled by the infrastructure. Such a data infrastructure has potential to improve accountability and measurement, to demonstrate the value of PHN services, and to improve population health.

Key words: data infrastructure, nursing workforce, Omaha System, public health nursing standards, public health systems, quantitative research.
enable public health leaders and policymakers to advance and to maximize PHN practice, and to meet national population health goals.

The meeting participants agreed that the PHN data infrastructure gap is consistent with the gap identified by the IOM for the broader public health system (Committee on Public Health Strategies to Improve Health, Institute of Medicine, 2011). Currently, there is a great deal of variation in data collection and reporting of PHN data across public health departments and other organizations, registries, and surveys. Adequate data sources are not consistently available at either the state or national level to assess programs and services, levels of service (individual, family, community, and system), population health data, penetration of service delivery within target populations, the PHN workforce, PHN-specific interventions, or PHN-sensitive outcomes (Scutchfield, Knight, Kelly, Bhandari & Vasilescu, 2004; Mays et al., 2006; Bhandari, Scutchfield, Charnigo, Riddell & Mays, 2010; Committee on Public Health Strategies to Improve Health, Institute of Medicine, 2011). Due to these deficits, it is difficult to fully identify and describe the PHN workforce or to attribute the contribution of PHN interventions to population health outcomes.

Using a modified Delphi process, the meeting participants generated 13 suggested research questions that were specific to guiding development of a proposed PHN data infrastructure (Table 1). (Kenney, Hasson & McKenna, 2011). The questions were validated and revised 1 month later at a special session held at the 2010 Annual Meeting of the American Public Health Association. Based on the 13 questions, four key research topics were identified as priorities: (1) determining the data elements in a minimum dataset that capture practice and outcomes, (2) outcome indicators that are PHN-sensitive at multiple levels, (3) issues of validity and reliability of the essential PHN data elements, and (4) use of existing databases to improve quality and safety of PHN practice (Issel et al., in press). This article summarizes the strengths, limitations, and gaps specific to PHN data; proposes a research agenda for the development of a PHN data infrastructure; and provides recommendations for future research that may be enabled by the proposed PHN data infrastructure.

### Table 1. Research Questions to Guide the Development of a PHN Data Infrastructure

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the minimum data elements related to public health nursing practice?</td>
</tr>
<tr>
<td>Which data elements ought to be in a minimum dataset?</td>
</tr>
<tr>
<td>What are the core variables essential to measure public health nursing practice outcomes?</td>
</tr>
<tr>
<td>What are public health nursing sensitive outcome indicators at: 1) individual, 2) family, and 3) community levels?</td>
</tr>
<tr>
<td>What are sensitive indicators of public health nursing-community partnerships (e.g., # of stakeholders, cohesiveness, appropriate representation)</td>
</tr>
<tr>
<td>What is the minimum common unit on which data ought to be gathered?</td>
</tr>
<tr>
<td>Which metrics (related to public health nursing processes and outcomes) are reliable and valid and at what level(s)?</td>
</tr>
<tr>
<td>Which public health nursing metrics are similar across settings and can be used across settings?</td>
</tr>
<tr>
<td>What predicts adoption of public health nursing sensitive/relevant metrics by an organization?</td>
</tr>
<tr>
<td>How can we develop a database of databases with a corresponding decision tree for selecting key variables related to public health nursing practice and outcomes?</td>
</tr>
<tr>
<td>How can current databases (e.g., Omaha System) be used to improve public health nursing safety and quality of practice?</td>
</tr>
<tr>
<td>What gaps exist in outcomes data currently being collected by/about public health nursing?</td>
</tr>
<tr>
<td>What are limitations (gaps) of metrics needed for decision making?</td>
</tr>
</tbody>
</table>

*Questions deemed priority by the participants in the October 2010 agenda setting conference.

### State of the Science of Public Health Nursing Data

In scientific and health care disciplines, vast quantities of data are becoming available due to advances in technology (Hey, Tansley & Tolle, 2009). Multiple sources of PHN data are available to support research examining the PHN workforce, practice, and outcomes (Magee, Lee, Giuliano & Munro, 2006; Zeni & Kogan, 2007). Each data source contributes an important perspective to the broader field of public health systems and services research (PHSSR) and to PHN research specifically. Similarly, each data source has limitations that must be addressed. Table 2 describes major PHN data sources presented by PHN data experts at the invited conference, including variables, and limitations for selected resources files, national and state surveys, registries, and clinical data.

The Health Resources and Services Administration (HRSA) Bureau of Health Professions Area…
<table>
<thead>
<tr>
<th>Data source</th>
<th>Description</th>
<th>Variables</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compiled resources</td>
<td>County-specific health resource information system designed to be used by planners, policymakers, researchers, and other professionals interested in the nation’s health care delivery system. May be purchased from Quality Resource Systems, Inc. (HRSA, 2011)</td>
<td>Health professions, health facilities, hospitals, vital statistics, population and economics, utilization, expenditures, and health professions training with geographic codes and descriptors which enable it to be linked to other files</td>
<td>Race/ethnicity varies depending on the data set. Reporting periods are based on the availability of each data element</td>
</tr>
<tr>
<td>Area resource file (ARF)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Registries</td>
<td>A collaborative effort between the CDC Immunization Safety Office and eight managed care organizations to monitor immunization safety, rare and serious events following immunization. (Centers for Disease Control &amp; Prevention, 2011c)</td>
<td>Population health outcomes, such as influenza outbreaks</td>
<td>Limited to defined geographic areas and participating organizations</td>
</tr>
<tr>
<td>Vaccine safety datalink</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special child health services registry (New Jersey)</td>
<td>A confidential record of infants and children who have birth defects and special health care needs (mandated conditions) or who are at risk for developing such needs (State of New Jersey Department of Health &amp; Human Services, 2011)</td>
<td>Birth defects, chronic health conditions</td>
<td>Specific to New Jersey, difficult to verify meaning of variables, difficult to link by time and location to variables from other databases</td>
</tr>
<tr>
<td>National health system surveys</td>
<td>Longest running governmental public health survey with high participation rate (82% in 2010) among local health departments Harmonized with ASTHO State Public Health Survey since 2010 (National Association of City &amp; County Health Officials, 2011)</td>
<td>Organizational characteristics, programmatic activities, workforce data (size and discipline), leadership data (discipline and education)</td>
<td>Limited to the governmental public health workforce and provide little detail regarding the nature of the workforce aside from attributes of the lead executives and proportions of staff in various job classifications</td>
</tr>
<tr>
<td>National profile of local health departments (NACCHO)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State public health survey (ASTHO)</td>
<td>State level public health profiles. Harmonized with NACCHO Profile since 2010 (Association of State &amp; Territorial Health Officials, 2010)</td>
<td>Public health workforce and leadership</td>
<td></td>
</tr>
<tr>
<td>Data source</td>
<td>Description</td>
<td>Variables</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>State health workforce surveys</td>
<td>North Carolina public health workforce development system (WDS)</td>
<td>Workforce characteristics</td>
<td>Variation in data elements across states</td>
</tr>
<tr>
<td></td>
<td>Initially built from statewide learning management systems to identify areas of training need, but providing individual level health department staff data regarding education, experience, certification. (Hajat et al., 2009)</td>
<td>(discipline and education) and local health department performance</td>
<td></td>
</tr>
<tr>
<td>National population health surveys</td>
<td>National health interview survey (NHIS) has monitored the health of the nation since 1957. NHIS data on a broad range of health topics are collected through personal household interviews, providing data to track health status, health care access, and progress toward achieving national health objectives. (Centers for Disease Control &amp; Prevention, 2011a)</td>
<td>PHN-sensitive outcomes regarding health care access and health status</td>
<td>Difficult to verify meaning of variables, difficult to link by time and location to variables from other databases</td>
</tr>
<tr>
<td></td>
<td>The NIS is a list-assisted random-digit-dialing telephone survey followed by a mailed survey to children’s immunization providers to monitor childhood immunization coverage since April 1994. (Centers for Disease Control &amp; Prevention, 2011b)</td>
<td>PHN-sensitive outcomes regarding immunization coverage</td>
<td>Difficult to verify meaning of variables, difficult to link by time and location to variables from other databases</td>
</tr>
<tr>
<td></td>
<td>National Survey of Children’s Health was conducted by the Centers for Disease Control and Prevention’s National Center for Health Statistics in 2003 and 2007. (Data Resource Center, 2011)</td>
<td>PHN-sensitive outcomes regarding children’s health status</td>
<td>Difficult to verify meaning of variables, difficult to link by time and location to variables from other databases</td>
</tr>
<tr>
<td>Clinical data</td>
<td>PHN documentation data</td>
<td>Interface terminology standards commonly used in PHN clinical documentation</td>
<td>Health problems, PHN interventions, and outcomes of PHN services reliability and observer bias require attention to data quality, variation in number of users by state</td>
</tr>
<tr>
<td>Administrative data</td>
<td>Client and provider details</td>
<td>Provider ID, client demographics and address, claims data</td>
<td>Variables may not be standardized</td>
</tr>
</tbody>
</table>
Resource File (ARF) is a national county level health resource information database. Data on health professions includes nurses, but does not include PHN-specific data. Other data are available regarding health care facilities and population, economic, and environment factors (U.S. Department of Health & Human Services, Health Resources & Services Administration, Bureau of Health Professions, 2011). Similarly, national and state registries exist for numerous health conditions, and thus may provide important contextual variables for PHN workforce and outcome studies. Although no current PHN studies using the ARF or registries were found in the literature, these variables have potential to be used with PHN workforce data to examine correlations between PHN capacity and important health conditions, such as obesity, children with special health care needs, and immunization rates (Centers for Disease Control & Prevention, 2011c; State of New Jersey Department of Health & Senior Services, Family Health Services, 2011; U.S. Department of Health & Human Services, Health Resources & Services Administration, Bureau of Health Professions, 2011).

National health system surveys include variables indicating services provided by health departments, and number of staffs in job classifications, including public health nurses (PHNs). However, they do not include the numbers, types, or qualifications of staff providing a specific service. In addition, staff numbers are reported based on their workforce functions (not on their credentials); thus, PHNs serving in positions not classified as a nursing position would not be identifiable as nurses (Association of State & Territorial Health Officials, 2010; National Association of City & County Health Officials, 2011). In terms of public health agency leadership, the National Association of City and County Health Official’s (NACCHO) National Profile of Local Health Departments Surveys (Profile) includes data regarding the licensure and education of health department “top executives,” providing a longitudinal record of PHN leadership in local public health systems (National Association of City & County Health Officials, 2011). Outside of governmental agencies, PHNs provide services through organizations, such as nonprofit community-based agencies, hospital systems, and health policy groups. Studies using NACCHO Profile data for PHSSR include examining relationships between nurse leadership in local health departments and the sets of services an agency provides (Bekemeier & Jones, 2009) and relationships between nurse-led health departments and health department performance (Bhandari et al., 2010) and clinician (nurse and physician) local public health leaders and county-level mortality disparities (Bekemeier, Grembowski, Yang & Herting, 2011).

State level data sources may include variables regarding outcomes of PHN interventions. For example, Hajat et al. (2009) used North Carolina Public Health Workforce Development System (WDS) data depicting county totals for “decline in teen pregnancy over time” as one measure of a population-level outcome that was positively associated with services conducted by PHNs and other “health professionals” in the health departments under study. Issel, Bekemeier, and Baldwin (2011) identified a set of three indicators that were sensitive to PHN interventions across multiple health departments and two states. The indicators endorsed were rates of Chlamydia, first trimester prenatal care, and early childhood immunization. Data sources included birth certificates, the Illinois statewide patient record system, and local health department WIC data in Illinois; and birth certificates, STD/TB service, and immunization registry data in Washington State.

Other PHN-related variables may potentially be found in national population health survey databases, such as the National Health Interview Survey (Centers for Disease Control & Prevention, 2011a), the National Immunization Survey (Centers for Disease Control & Prevention, 2011b), and the National Survey of Children’s Health (Data Resource Center, 2011). However, no studies using these sources were found in the literature.

Clinical databases (i.e., administrative, billing, and interface terminology documentation data) have potential for studying the quality of population-focused PHN interventions (Monsen et al., 2006). Administrative variables include PHN and agency identifiers. Interface terminology variables describe population health status and related population-based PHN interventions at individual, family, and community levels. The Omaha System is an interface terminology that is commonly used in PHN clinical documentation software (Farri, Monsen, Westra & Melton, 2011; Martin, 2005; Melton et al., 2010). Examples of PHN research using clinical and administrative data include: a reliability
study (Monsen et al., in pressa); the examination of home visiting intervention effectiveness for high risk mothers (Monsen, Banerjee & Das, 2010a; Monsen, Radosevich, Kerr & Fulkerson, 2011b); intervention effectiveness for mothers with intellectual disabilities (Monsen, Sanders, Yu, Radosevich & Geppert, 2011c); benchmarking PHN outcomes (Monsen, Radosevich, Johnson, Farri, Kerr, & Geppert, 2011; Monsen et al., 2010b, 2011c); the development of a maternal risk index for PHN caseload management (Monsen et al., 2011b); and the development of a new interim improvement metric for family home visiting (problem stabilization) (Monsen, Farri, McNaughton & Savik, 2011a). Community level interventions were compared to individual level interventions in a study of PHN nurse manager interventions (Monsen & Newsom, 2011). Some states (e.g., Minnesota, Washington, Arizona, Maine) have implemented PHN program evaluation protocols using interface terminology data (Elsbernd, Barnhart, Stock, Monsen & Prock, 2011). An Omaha System data warehouse of PHN assessments, interventions, and outcomes exists within the University of Minnesota School of Nursing Center for Nursing Informatics. The purpose of this data warehouse is to advance PHN outcomes research (Omaha System Partnership, 2011).

Limitations of Public Health Nursing Data Sources

Public health nursing data exist in diverse forms and typically must be extracted, transformed, harmonized, merged, and cleaned before analysis can begin (Table 2). Availability of PHN data from resource files, surveys, registries, and clinical health records varies by state, agency, and program. Thus, there are many challenges that must be overcome to build a database that comprises all necessary variables to address a particular research question or to evaluate program effectiveness. All secondary data sources have limitations, such as missing data and inherent errors; thus data cleaning is critical. Observer bias issues must be identified and addressed as a limitation, and alternative explanations for findings must be considered. Clinical data have limitations common to all observational datasets, such as fidelity to documentation procedures and observer bias. Omaha System users have addressed these limitations through training and shared inter-rater reliability materials and standards (Monsen & Martin, 2002a,b; Monsen et al., 2006, in pressa; Minnesota Omaha System Users Group, 2011). Linking variables from multiple databases depend on many factors, such as identifying PHN variables that are similar across settings. Linking databases often requires significant time to obtain, merge, and extract the data. Collaboration is required to promote data sharing among these diverse data sources to build the data infrastructure, as data may be owned by individual investigators, government sponsored databases or public–private partnerships that facilitate linkage of data owned by private organizations. These challenges can be minimized by careful selection of the sample and variables, and subsequent management of the data (Bradley, Penberthy, Devers & Holden, 2010).

Implications and Recommendations

The PHN researchers should prioritize research that identifies minimum data elements relevant to PHN practice, the core variables essential in measuring PHN outcomes at various levels, the reliability and validity of PHN processes and outcomes, and how existing databases can be used, expanded, combined, or refined to meet these identified goals. The critical need for a uniform minimum dataset of PHN interventions, outcomes, and health system data drives this research agenda. The PHN researchers bring unique practical, conceptual, methodological, statistical, and informatics expertise that is requisite for the development of a PHN data infrastructure. The PHN researchers also bring expertise in public health systems to public health workforce interventions, finance, management, and administration studies. To conduct such research, further resources and antecedents that support the broader field of PHSSR are needed. These resources include funding to support the research effort and attention to the congruence between the PHSSR agenda and the PHN research agenda (Bhandari et al., 2010; Mays et al., 2006; Scutchfield et al., 2004).

Public health practice-based research networks (PBRN), formally established and funded by The Robert Wood Johnson Foundation (RWJF) in 12 states across the United States offer further opportunities for identifying reliable and valid outcome measures and additional existing and new data sources (University of Kentucky College of Public Health, 2011). In addition, the Omaha System
Partnership for Knowledge Discovery and Health Care Quality is an international partnership in which PHN scholars and practitioners are actively engaged in advancing PHN research using PHN-generated clinical datasets (Martin, Monsen & Bowles, 2011). The Omaha System Partnership has scientific teams from the University of Minnesota and the University of Pennsylvania working with affiliate members from practice settings in many countries. These investigators have initiated over 30 investigations of important clinical questions, with 16 scientific articles at the time of this publication. Capitalizing upon these partnerships between practitioners and researchers will strengthen the rigor and advance the utility of research that connects PHN practice and specific health outcomes, improving PHN accountability and measurement.

A comprehensive uniform PHN dataset is proposed using combined elements of PHSSR data (health services variables), population health data (surveys and registries), and clinical PHN documentation data (intervention and outcome variables for services at the individual, family, community, and systems levels). Development of this dataset would enable meaningful investigation of a broad range of practice, quality, evaluation, and safety questions. Several examples of possible studies enabled by the PHN data infrastructure are suggested below for future consideration by PHN researchers:

- Development of standardized metrics for PHN performance.
- Evaluation and monitoring of population health status across jurisdictions.
- Examination of links between public health system factors, PHN interventions, environmental and socioeconomic factors, and client outcomes.
- Effectiveness of PHN services and programs at individual, family, community, and systems levels.
- Ecological models incorporating contextual health systems data to improve understanding of PHN interventions and outcomes within the larger public health and health care systems.

Key questions to consider as the PHN data infrastructure is developed are

- Who or what agency might fund/finance such a dataset?
- Who or what agency would/should maintain oversight?
- What should the database infrastructure look like?
- What kind of electronic platform is recommended?
- What exact data should be in this uniform dataset?
- How could/should local/regional/state data be collected and fed into such a database?
- How frequently should this database be updated?
- How could this database be accessed?
- Who would be given access rights?

It is incumbent upon the PHN community of scholars, practitioners, and policy makers to address these issues to advance the PHN data infrastructure agenda.

The Institute of Medicine (2011) report on Accountability and Measurement in Public Health identified a critical need for better data to enable evaluation and improvement of the public health system. In alignment with this report, an invited group of national PHN scholars, practitioners, and stakeholders used a Delphi process to identify gaps in PHN data. Public health nursing data exist in multiple forms and settings, with major gaps, barriers, and limitations to their use in research and evaluation. Addressing these gaps will enable PHN researchers and practitioners to advance toward a more comprehensive uniform PHN dataset, combining elements of the PHSSR data and clinical PHN documentation data. Such a data infrastructure will improve accountability and measurement, demonstrate the value of PHN services, and improve population health.

Acknowledgments

Funded by AHRQ Conference Grant: 1 R13 HS01 8852-01.

References


Magee, T., Lee, S. M., Giuliano, K. K., & Munro, B. (2006). Generating new knowledge from existing data: The use of large data sets for nursing research. *Nursing Research, 55*(Suppl. 2), S50–S56.


