

12-18-2015

Strengthening Public Health through Web-Based Data Query Systems

Manik Ahuja

University of Connecticut, USA, manik.ahuja@uconn.edu

Follow this and additional works at: <https://opencommons.uconn.edu/dissertations>

Recommended Citation

Ahuja, Manik, "Strengthening Public Health through Web-Based Data Query Systems" (2015). *Doctoral Dissertations*. 993.
<https://opencommons.uconn.edu/dissertations/993>

Strengthening Public Health through Web-Based Data Query Systems

Manik Ahuja, PhD

University of Connecticut, 2015

Aggregate level local health data has become more easily available to the public through Web-Based Data Query Systems (WDQS). Local health data can be a powerful vehicle for improving the health of a community. When aggregated, local health data help monitor the incidence, trends, and patterns and disease in a given population. WDQS make local health data easily available over the internet. WDQS are interactive and have the capability for users to query off multiple datasets and to pre-select variables. Despite the advantages of WDQS, only 29 states have implemented them. States that have not implemented WDQS are using outdated technologies such as static reports to share their health data. We conducted a three part study to investigate the challenges state agencies face with their implementation of WDQS. The three part study included a systematic review of literature, a Delphi study, and a survey of state health coordinators in all fifty states.

We found that the high cost of system development, data sharing between state agencies, inadequate staffing, standardization of vocabulary between datasets and a lack of understanding of how consumers use their data as the most challenging. Website performance, poor website usability, the cost of hardware/software, privacy/security, data storage, and the ability large data sets are less of a problem. The contribution of this project was significant in developing an understanding of key gaps in knowledge on problems in the development and usage of WDQSs. In the long term, we anticipate that having more useful data will help lead to improved health surveillance and more informed and targeted interventions at the local level.

Strengthening Public Health through
Web Based Data Query Systems

Manik Ahuja, MA

B.A., University of Washington

M.A. University of Washington

A Dissertation

Submitted in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy
at the
University of Connecticut

2015

APPROVAL PAGE

Doctor of Philosophy Dissertation

Strengthening Public Health through
Web-Based Data Query Systems

Presented by

Manik Ahuja, MA

Major Advisor _____

Robert Aseltine Jr., Ph. D.

Associate Advisor _____

Susan Reisine, Ph..D.

Associate Advisor _____

Nick Warren, Sc. D.

ACKNOWLEDGEMENTS

I would like to thank the members of my doctoral committee, Dr. Robert Aseltine, Dr. Susan Reisine, Dr. Nick Warren, Dr. Andrew Cislo, and Dr. Pam Williams. Special thanks to Dr. Aseltine for all his time and effort as the major advisor and for his mentorship. I am fortunate to have received continuous support from my dissertation committee, who contributed to the success of my project. They were available to help with new ideas and provide useful feedback as my project evolved. I would like to thank Connie Cantor, Diane Burnat, Kathleen McDermott, and Allyson Moffo from the Center of Public Health and Health Policy and Barbara Case and Morgan Spencer from Community Medicine for all their support over the last five years. I am grateful to have strong mentors and support throughout this journey. I look forward to making a significant public health impact in the future.

TABLE OF CONTENTS**PAGE NUMBER**

Title Page	I
Approval Page	II
Acknowledgements	III
Table of Contents	IV
List of Tables	VI
List of Figures	VII
Chapter 1: Introduction	1
Chapter 2: Manuscript #1: The Dissemination of Local Health Data using Web Based Data Query Systems: A Systematic Review	11
Abstract	11
Background	13
Methods	15
Results	20
Discussion	29
Chapter 3: Manuscript # 2: Challenges Faced with the Implementation of Web Based Data Query Systems: a Delphi Consensus Study	33
Abstract	33
Background	35
Methods	38
Results	44
Discussion	48

Chapter 4: Manuscript #3: Barriers State Agencies Face in the Dissemination of their Health Data	54
Abstract	54
Background	55
Methods	57
Results	63
Discussion	70
Chapter 5: Summary and Conclusions	76
References	88

LIST OF TABLES

TABLE	PAGE
1. Quality Rating Indicators and Criteria	18
2. Quality Rating of Articles	20
3. References of Articles Included in the Review and Summaries of Their Basic	27
4. Summary of Findings	28
5. Overall Consensus from Round 1 and Round 3	46
6. Items that met Consensus Using a Five-Point Likert Scale	47
7. Items that did not meet Consensus after Round 3	48
8. Profile of Health Data Characteristics at the State Level	64
9. Mean Participant Ratings on System Aspects on a Five-Point Likert Scale	66
10. Participant Ratings of Challenges State Agencies Face on a Four-Point Likert	69
11. Relationship between Manuscript #2 and Manuscript #3	79
12. Comparison in Response between Delphi study vs BRFSS study	82

LIST OF FIGURES

FIGURE	PAGE
1. Process Flow of Dissertation	9
2. Flow Chart of Study Identification, Selection, and Rejection	19
3. Schematic of Delphi Study	39

Chapter 1:

INTRODUCTION

The Importance of Health Surveillance

Public health surveillance is defined by the World Health Organization as “the continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice” (Thacker, 2012). There is a need for public health surveillance as it is vital in understanding and ultimately improving population health (Smith, 2003). Surveillance information improves public health by targeting interventions and documenting their effect on the population (Nsubuga, 2006). For example, rates of smoking, obesity, and the number of children fully immunized can be measured over time. Public health surveillance goals include the assessment of population health, the identification of public health priorities, the evaluation of existing public health programs, and the development of new effective interventions and strategies to protect and improve public health (Nsubuga, 2006). Public health surveillance is not just gathering health data and generating reports; it is the continued watchfulness over the distribution and trends of incidence. Public health surveillance may be used to target or modify education, immunization, and other risk-reduction programs, including elimination of hazards in the environment or workplace. The ultimate purpose of public health surveillance is to drive public health action (Birkhead, 2015). In the U.S., local health surveillance is conducted at the state level. Each state’s health department is responsible for monitoring the health of its population. Their purpose is the same, which is to promote, protect, and maintain the health and welfare of their citizens. Each state is responsible for its own data collection, analysis, and dissemination of their health data

(Chambers, 2006). Data dissemination can be defined as the distribution or transmission of statistical data to users (Cheng, 2012).

Local Health Data for Health Surveillance and Dissemination

Local health data can be a powerful vehicle for improving the health of a community (Luck, 2006). When aggregated, local health data help monitor the incidence, trends, and patterns and disease in a given population (Diamond, 2009). Data are aggregated by variables, such as race, ethnicity, gender, education level, and county, which help identify populations at risk. These data can highlight both the existence of problems and opportunities for improvement within a community and lead to targeted interventions (Luck, 2006).

Data are also used to further support decision making, strategic planning, and evaluation of existing public health practices and programs for disease control and prevention (Portnoy, 2014). They are used by multiple stakeholder groups such as government agencies, community based organizations, and advocacy groups. Government agencies (state, federal, and local) use local health information to develop policies and legislation, allocate resources across departments and program areas, and plan and evaluate their activities (Luck, 2006). For instance, the legislation process may lead to funding that is allocated for an at-risk population. Community-based organizations (CBOs) use local health information to design, target, and evaluate programs that meet specific community needs. A nonprofit group may use health data to raise awareness for an emerging health issue. Advocacy groups use data to quantify the need for policy change and convey their message effectively to policymakers, the media, and the public. Researchers also depend on data to understand patterns of health and disease as well as to evaluate policies and design new health improvement strategies (Luck, 2006). Employers and business associations can use data to assess conditions affecting the health of their workforce and

estimate effects of their decisions on the community's health (Luck, 2006). Examples of data that are available include births, deaths, fetal deaths, infant mortality, hospital discharge data, and immunizations. State agencies collect these data and aggregate it based on geographic and socio demographic factors such as race, gender, and education.

Health surveillance data also contributes to understanding the epidemiology of disease, predicting future pattern of disease occurrence and developing prevention programs (Choi, 2012). For instance, if a particular ethnic group has a high incidence of disease combined with high population growth rate, we can estimate that rates are likely to increase for that group. Such forecasts are useful for planning resource needs.

The Use of the Internet to Disseminate Local Health Data

The most common method to disseminate local health data today is the internet. The internet has emerged as an essential public health tool (Thacker, 2012). It has become a critical medium for public health practitioners and multiple stakeholder groups seeking health information (Thacker, 2012). The internet helps facilitate this as data are easily available in significantly less time. The advancement of the internet has changed how public health organizations collect and disseminate their data. Data are now easily transmittable from one organization to another. They are disseminated to the public through multiple channels. There are three primary modes that state agencies use to release data on the internet that include ad-hoc system, a static interface, and Web Based Data Query Systems (WDQS).

Ad-hoc system allows users to submit data requests through a state's agency's website. An ad-hoc is defined as a request for one purpose, as the term 'ad hoc' is commonly used in the information technology industry. An ad-hoc interface does not contain any data, but allows the user to make a request for a report. For instance, through the Massachusetts Department of

Human Health and Services website, users are able to submit requests for basic vital statistics such as births, deaths, and infant mortality. A user can submit a request online and the request is reviewed by health agency staff. The request is either denied or approved. It was more common for state agencies to use ad-hoc systems in the early to mid-1990s. Ad-hoc requests require manual intervention such as requesting data by staff and formatting data so they are usable. Depending on staffing resources, ad-hoc requests can take weeks to complete. Most ad-hoc systems have been replaced by a static data interface or WDQS.

A static interface is a website administered through a state's health department. Within a static interface, users are able to view or download reports but cannot query off datasets. A static interface is not an application but a webpage that contains multiple reports that a user can view or download. These reports are prepared and generated in advance before they are published to the website. In this situation, the state of Indiana uses a static interface, *StatsIndiana*, to share their data. Users can select from a wide range of reports and select on indicators such as fertility rate, births by white mothers, and births by black mothers. Users are able to view a report on the internet, or download the report data in a Microsoft Word, or in a PDF format. Static interfaces are updated on time intervals (quarterly, annual, etc.). A static interface limits user choice to pre-calculated statistics and allows no choice of parameters for a query (Friedman, 2006). In most cases, a state agency staff will make the most commonly requested reports available. For instance, a state agency may make low birth weight rates across county available but not *both* county and race/ethnicity. An interactive based system, such as a WDQS, allows the user to create their own type of report, while a static report is limited to the discretion of state agency personnel.

Web-Based Data Query Systems are interactive based systems that use a dynamic interface and allow users to interact with the system. A dynamic interface enables users to customize data queries through choosing data sets, variables, measures, and the format for presenting query results (Friedman, 2006). Currently, there are 29 states that share their health data through WDQS. WDQS allow users to customize queries based on preference. To illustrate, the state of Pennsylvania developed *EPIQms*, which allows the user to query from a broad range of datasets. Unlike static reports, WDQS allow the user to directly query from a dataset. For instance, through *EPIQms*, users can query off datasets such as communicable disease, teen pregnancies, emergency medical deaths, and cancer incidence. *EPIQms* allows users to select a dataset and filter by year and county.

Advantages of WDQS

WDQSs offer greater breadth of data, depth of data, functionality, and provide data in a timely manner in comparison to a static interface or ad-hoc based system. While ad-hoc systems have the ability to provide users with customizable data, there are delays in processing. The user submits a request based on their data need, and state agency staff may either deny or complete the request. However, depending on staff resources, data may not be available for several weeks (Ver Ploeg, 2006). By the time data are available, the value to a researcher might be limited (Bernstein, 2012). Timely dissemination of data to those who make policy and implement intervention programs is critical to the usefulness of health data (Nsubuga, 2006). A static interface may be useful to provide overall trends in a population but lacks necessary data to answer a specific research question. A static interface is not customizable, and data available may not be relevant to the user's need. WDQSs have the capacity to generate tables as query outputs and to specify formats for those tables. The production of row and column percentage

and time trending are examples of basic functionalities. To illustrate, a user may be interested in comparing cardiovascular disease rates for Hispanics with periodontal disease across two counties. Such level of reporting is unlikely through a static interface but may be possible through a WDQS. WDQSs have capabilities for users to query off multiple datasets and to pre-select variables, which are advantageous (Friedman, 2006). WDQSs also offer greater breadth of data. For example, birth datasets typically allow users to query using demographic variables such as age, education, ethnicity, race, sex of the mother, maternal risk factors such as smoking, prenatal care, and the infant's birth-weight (Friedman, 2006).

Development of WDQS

The WDQS development process contains five major phases that include planning, design, development, testing, and implementation.

- *The planning phase* - The planning phase includes setting goals and objectives, identifying data sources, determining if there are adequate financial resources, and also gaining political support. Political support is necessary to get 'buy-in' that there is a need to allocate resources for WDQS (Friedman, 2006). WDQSs are expensive to develop, in which case it is important for state agencies to demonstrate that there is a public health need. Planning also involves data collection and collaboratively working with entities across the state to access their data. In this case, it may be necessary for a state agency to work with local hospitals so their emergency room discharge data can be accessed.
- *The design phase* - The design phase includes creating system requirements, setting timelines for deliverables, meeting business requirements, designing technical specifications, and understanding data sources. In this phase, the project management team works closely with software developers in designing the proposed system. Data

sources are profiled and validated to ensure that the flow of data fits the requirement of the system.

- *The development phase* - The development phase includes the software development of the system, which were specified in the design phase. Software development requires specific expertise and training. Software developers/programmers are generally trained in multiple programming languages such as C+, Java, Python, and other languages. Throughout the development phase, project managers and analysts work with the software developers to ensure that the project specifications are being met. In many cases, modifications from the original specifications are changed and discussed in which case, the project specifications submitted in the design phase can be updated or modified.
- *The testing phase* - The testing phase consists of testing functionality, data validation, and ensuring the system requirements are properly met. A testing lead develops a user acceptance testing (UAT) plan with test case scenarios. Testing is often conducted internally by state agency staff as they execute the UAT plan. The goal of this process is to identify problems prior to implementation and release.
- *The implementation phase* - The implementation phase includes making the system available to the public, so it can be accessible on the internet. Once the system is implemented, it is important to routinely maintain the system. This process of WDQS maintenance can cost state agencies several hundred thousand dollars. Having adequate financial support is vital to adequately maintain systems as data are continuously updated.

Barriers to Implementation

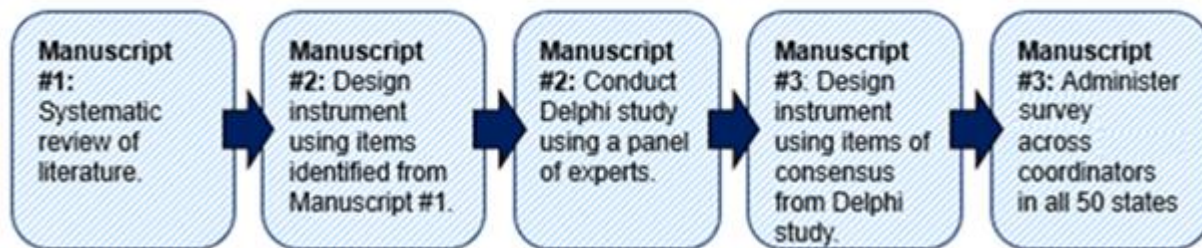
Over the last decade, computers and information technology have continued to improve. Computers and servers are faster and more powerful than ever. Speed and bandwidth have significantly improved, and the cost of hardware and software is much lower. Despite the advancements in technology and the advantages of WDQS, over 40% of the states in the U.S. have not implemented a WDQS. Many states with existing WDQS are using outdated technologies. Currently, 29 states have implemented WDQS, in comparison to 27 in 2005. Between 1999 and 2005, there were aggressive efforts with WDQS implementation as the number of states with WDQS increased from 2 to 27 (Friedman, 2006). Along with a lack of implementation, existing states with WDQS have had shortcomings in upgrading their systems in a timely manner. A system should be upgraded with improved functionality at least once a year in order to keep up with technological innovation (Chapin, 2001). This has not been the case with WDQS, as systems are not keeping up with the rapid improvements in technology. To illustrate, the widely adopted IBIS-PH was first released and developed in 1999 by Utah Department of Health. The next major upgrade, version IBIS-PH 2.0, was not released until 2010, and the next upgrade is due in 2016. IBIS has been adopted by five other states, making them dependent on Utah's progress. In this case, the slow progress of upgrading a system can affect multiple states.

Our research has indicated that despite the known advantages of WDQS, there are many barriers state agencies face with their implementation. Twenty-one states are not using WDQS, and are using outdated technologies such as a static interface or ad-hoc system. Amongst states with WDQS, less than five states have made significant enhancements since 2010. WDQS also face problems with usability, having complete data, and inoperable systems.

Process Flow of Dissertation

These three manuscripts are interdependent. The Delphi study was dependent upon the completion of our systematic review, and our survey across 50 states was dependent upon completion of our Delphi study (Fig. 1). Topics of importance found in our systematic review were used to design the Delphi study. Items from the literature (Manuscript #1) that were relevant to the dissemination of WDQS were used to develop a four topic/42-item questionnaire for our Delphi study. Items that were of importance based on consensus from a panel of experts in our Delphi study (Manuscript #2) were used to design our BRFSS coordinator survey (Manuscript #3).

Figure 1: Process Flow of Dissertation



Description of three manuscripts

1. In the first manuscript, we conducted a systematic review of literature on the challenges state agencies face with their dissemination of WDQS. This paper analyzes the current research available on WDQSs and identifies key barriers that state agencies face.

2. In the second manuscript, we applied the Delphi method to gather a consensus of the most important topics from a panel of experts. The Delphi method is a flexible, effective and efficient research method that can be successful in answering research questions in information systems (Skulmoski, 2007). We designed the survey instrument based on items identified from the first manuscript. Topics identified in the literature are used in the Delphi survey, and experts rated the importance of each item. The Delphi study consisted of three rounds.

3. In the third manuscript, we conducted a comprehensive survey using a broad based sample. We used the results from the second manuscript to build the questionnaire for this survey. Items of importance from our Delphi study were used in this questionnaire. We surveyed Behavior Risk Factor Surveillance System (BRFSS) coordinators from all 50 states and the District of Columbia on the challenges state agencies face with their dissemination. We invited BRFSS coordinators because of their significance in releasing health data and their role in their state's technology decisions (Friedman, 2006).

Chapter 2:

Manuscript #1: The Dissemination of Local Health Data Using Web Based Data Query Systems: A Systematic Review

ABSTRACT

Background: Public health departments throughout the United States have placed increasing emphasis on improving health at the local and state level. Several state agencies make public health data available through Web-Based Data Query Systems (WDQS). There is potential for public health benefit at the local level. However, relatively little is known about the characteristics of utilizing WDQS, barriers to implementation, and their impact. The purpose of this paper is to conduct a systematic review to analyze the research available on WDQS and to identify gaps in knowledge.

Methods: We searched online databases including PubMed, Scopus, and Web of Science (Search dates: January 1, 1994 – January 1, 2015). We identified barriers and conducted a systematic review of literature. Documents with an emphasis on the development, data collection, and implementation of WDQS were eligible for inclusion. We grouped articles based on identified focus areas that include data sharing, privacy/risk, and technology. Quality was rated on a 5-point Likert scale using a quality indicator scale that includes the following metrics: important, useful, relevant, and complete.

Results: Sixteen barriers were identified and classified in four categories: data collection, IT infrastructure, cost, and usability. Of the articles reviewed, 78% had a technology focus, followed by data sharing (56%), and privacy/risk (22%). Quality was rated and articles that rated ‘below average’ (38%) were highest, followed by average (33%), and above average (28%).

Conclusion: This review systematically highlights the barriers and challenges organizations face when sharing their data. The review also identified gaps in knowledge around the impact and utilization of WDQS. Further research needs to be conducted on the barriers highlighted.

BACKGROUND

Web-based Data Query Systems (WDQS) are a common mechanism that local and state agencies use to disseminate local level health data (Gjelsvik, 2006). WDQS are easily accessible on the internet, making local health data more easily available to a wider audience. Many of the challenging health problems facing the U.S. require high-quality aggregate level public health data at the local level (Myers, 2008). The availability of such data has considerable potential benefit for improving health as a method of identifying community health problems and monitoring the health of a community as well as targeting interventions and informing policy decisions (Luck, 2006). National and state health surveys are popular and provide critical information towards achieving state and national health objectives but rarely serve local needs (Healthy People, 1991). The local level can be defined as data at the county, town, city, or municipality level. They help provide the health status, resources, and health challenges of a community. For example, in 2001 The Fresno County Public Health Department used county-level data, which identified the most important public health issues and obtained additional resources to address them (Luck, 2006). It was determined that rates of smoking were significantly higher in certain subgroups of New York. This led to 35,000 free courses of nicotine patches to heavy smokers as data made it possible to distribute the nicotine patches in areas where they were most needed (Fielding, 2004).

Public health data is defined as data that are primarily collected by public health agencies for routine responses, such as disease surveillance or program monitoring, without primary intention of research (Van Panhuis, 2014). At the local level, data are used by public and private organizations in the community. Public organizations use local health data for activities such as resource allocation, community needs and policy decisions. Private organizations use it to assess

conditions affecting the health of their workforce, estimating effects of their decisions on the community's health, and other areas of interest (Narayan, 2006).

Health data have limited value unless they are disseminated in a timely manner to serve as a basis for public health action (Lee, 2012). WDQS make public health data available in a timely manner and in many cases in a matter of seconds. Historically, data requests made through ad-hoc and static processes were submitted directly to state agency staff. Such requests took as many as several weeks to complete due to limited staff, resources, and an overall cumbersome process (Ver Ploeg, 2006). Prior to the growth of the internet, data “stovepipes” or data storage silos were characteristic of public health and other government agencies. Gaining access to these “stovepipes” for local health data was a cumbersome process for agency staff (Warnecke, 2008). Data today are stored on secure servers that store and transmit data at rapid speeds.

WDQS are also useful for epidemiologists and researchers interested in future trends. Information technology also allows for the integration of diverse data sources that helps support predictive models for timely interventions (Kopp, 2002). Such predictive models help formulate trends and use high level statistics to help guide timely interventions. As an example, a population with a high growth rate may provide us with future projections of disease.

While public health data has great potential to ameliorate health problems, little is known about barriers organizations face and the impact and use of WDQS (Haggard, 2006). Efforts to expand the availability and use of local health information face major technical and institutional barriers, as well as health information privacy concerns (Luck, 2006). The objective of this study was to conduct a systematic literature review to evaluate the state of the literature and to identify gaps in knowledge. We aim to identify organizational and technical barriers. We also

aim to understand utilization by identifying which groups are primary users of WDQS, how the data are consumed, and their public health impact.

METHODS

We conducted a systematic review according to the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines (Moher, 2010) to identify document articles about development and dissemination of public health data using state level Web Based Data Query Systems (WDQS). We defined development as the designing of the prototype, gathering data, developing the product, testing the application, and releasing of the application. We defined dissemination as the distribution or transmission of statistical data to users (Smith, 2003). Various channels are used to disseminate data, such as software applications, the internet, CD-ROM, or through authorized files for specific users. We defined public health data as data that were primarily collected by public health agencies for routine responses, such as disease surveillance or program monitoring, without primary intention of research (Byass, 2009).

Prior to the literature search process, the author established specific inclusion criteria for eligible articles. For inclusion, an article was included if it met *all* of the following criteria: (1) published in the English language; (2) a United States based focus; (3) published in a peer-reviewed scholarly journal. In addition, articles were included if they met *at least one* of the following criteria: (1) primary emphasis is on the development of WDQS; (2) primary emphasis is on the dissemination of public health data at the local level; (3) a single or multiple evaluation of WDQS systems; (4) evaluation of WDQS programs at the local health or state agency level; (5) evaluation data sharing of health data at the local level. Excluded from the analysis were articles with a focus on query systems for electronic medical records (EMR), electronic health

records (EHR), a focus on the life sciences, genetics, population science, and other non-relevant topics.

We searched the PubMed, Scopus, and Web of Science databases in January 2015 to identify studies relevant to public health WDQS (Fig. 2). All studies were read independently by one investigator. We used the keywords WDQS, “Web Based Data Query Systems,” state public health query systems, state public health surveillance query systems, public health query systems, “disseminated” and “public health,” state public health aggregate level data, “state agency” AND “public health,” and “data” AND “public health agencies.”

Next, we identified articles from January 1, 1994 to January 1, 2015, the period that corresponds to the rapid growth of the internet. We identified a total of 3,680 articles (Figure 1); of these, 843 were duplicate articles and 2461 were excluded based on title and abstract review. We reviewed the full text of the remaining 376 papers; 358 studies were excluded because they focused on other topics, such as the dissemination of EMR/EHR/Clinical data (126 articles), biological/genetics/natural science focus (62 articles), environmental and occupational health/outbreaks (55 articles), and had an international focus (27 articles) or other miscellaneous topics (89 articles), that did not meet the selection criteria. Eighteen articles were finally included in this review. We grouped and generalized barriers specific to the four challenges identified including cost, data sharing, IT infrastructure, and usability. These challenges were identified based off common themes. For each barrier identified, we reported key challenges organizations faced from the literature reviewed.

We then categorized each article on relevance and categorized them by study type of and focus. Articles were classified into one of the four possible types, which include: evaluation, methods, survey, or commentary. The study type ‘evaluation’ included articles that evaluated

single WDQS systems, multiple WDQS systems, or WDQS programs at the state or agency level. The study type ‘methods’ included articles with proposed new methodologies to collect, disseminate, or evaluate WQQSs. The study type ‘survey’ included articles that conducted a survey or interviewed informants and analyzed the results to form conclusions. The study type ‘commentary’ included editorials and articles that were based on expert opinion or narratives on WDQS topics. We then classified each article one or more focus areas. The focus area of technology included articles with themes specific to information technology (IT), data dissemination, hardware software and the development of WDQS. The focus area of data collection included themes of data sharing, health surveillance, and the exchange of public health data between organizations. The focus area of privacy/risk included issues with privacy, security, small cell counts, and risk of sensitive health data.

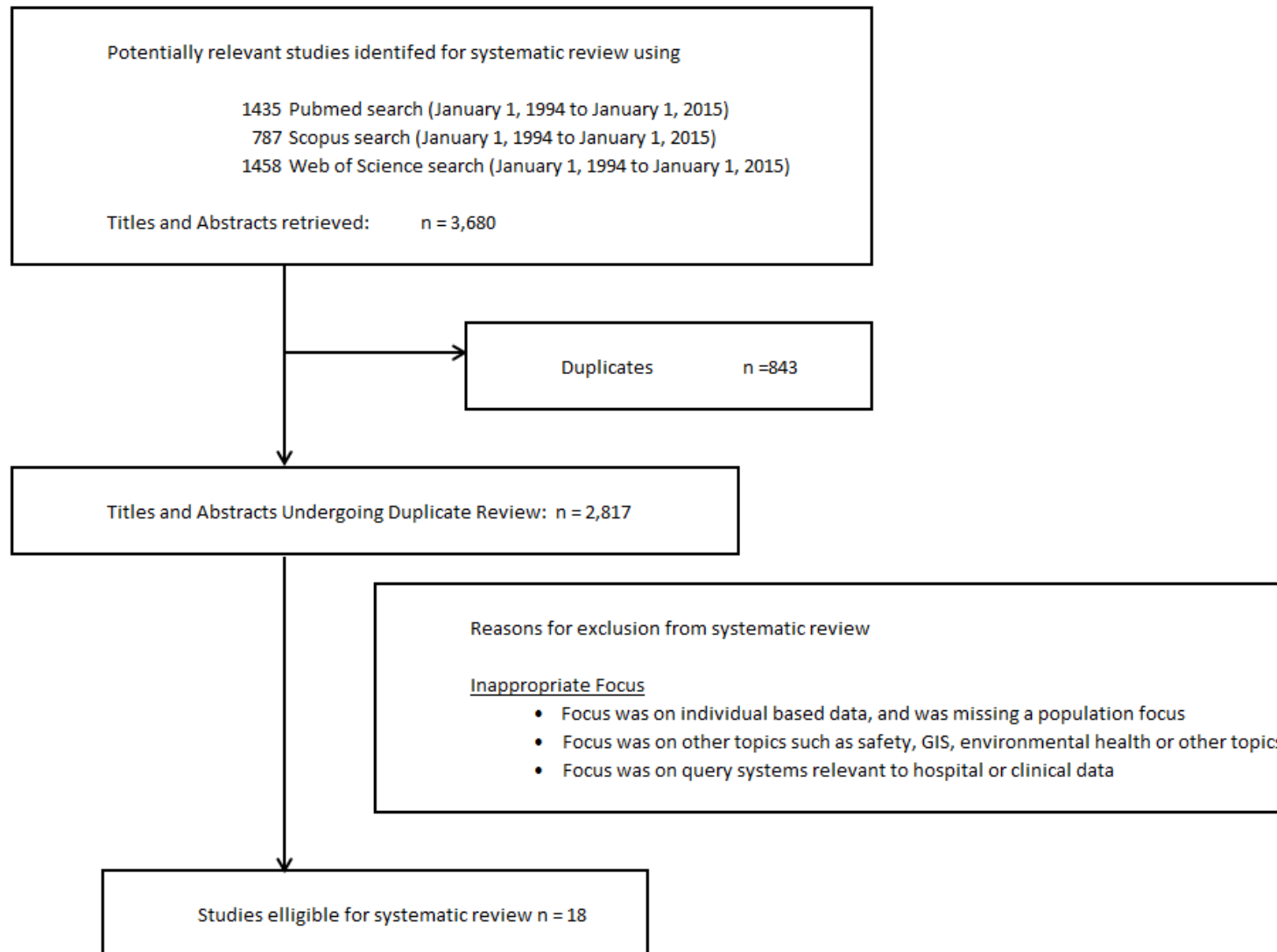
Articles were rated on quality using a quality indicator scale (Table 1) proposed by Jefferson (2008) (Jefferson, 2008). We chose the method proposed by Jefferson as it was suited for studies that are non-evidence based and focus on population health. Other popular quality indicator approaches such as Cochrane/GRADE were not used, as they are based on assessing the quality of evidence based studies. The Cochrane/GRADE approach is often used for evaluating the quality of evidence in systematic review. However, this approach is more common for evidence based or clinical studies (Schunemann, 2006). Articles were rated on a Likert scale (Table 2), based on criteria specified in Table 3. The quality rating ranged from 1 to 5, in which a rating of 5 was designated as ‘Excellent,’ while a rating of 1 was designated as ‘Extremely Poor.’ The quality indicator of ‘importance’ is defined as having a major impact on health. Indicators used to measure importance are potential change in health status or health care delivery. The quality indicator ‘useful’ is defined as the overall scientific contribution or

knowledge of a subject. An indicator that is used to measure usefulness is the article's contribution to the systematic review. The quality indicator 'relevant' is defined as the level of relevance to the journal's aims and readers. The quality indicator 'complete' is defined as having all relevant information presented. Indicators used to measure completeness are whether there is no selective presentation of data and if references are properly cited (Moher, 2010). Based on the indicators, a rating was applied. The overall quality of the article was determined by the mean of the four indicators: important, useful, relevant, and complete.

Table 1 Quality rating indicators and criteria.

Indicator	Criteria	Quality
Important	Primary findings have a potential for major public health impact	5 = Excellent 4= Above Average 3= Average 4= Below Average 5= Extremely Poor
Useful	Study contributes significantly to the scientific knowledge or debate	5 = Excellent 4= Above Average 3= Average 4= Below Average 5= Extremely Poor
Relevant	Topic is relevant and consistent with the aims of the journal and its readers	5 = Excellent 4= Above Average 3= Average 4= Below Average 5= Extremely Poor
Complete	All relevant information is presented	5 = Excellent 4= Above Average 3= Average 4= Below Average 5= Extremely Poor

Figure 2 Flow chart of study identification, rejection and selection



RESULTS

Quality Ratings

Overall, articles that were rated (Table 2) with a mean rating of ‘4 or higher’ were fewest (n=5; 28%), followed by average with a mean rating ‘between 3 and 3.99’ (n=6; 33%), and below average with a mean rating ‘between 2 and 2.99’ (n=7; 39%).

Table 2 Quality rating of articles

Citation	Importance	Useful	Relevant	Complete	Mean Rating
Braithwaite et al. [35]	2	2	2	5	2.75
Cohen et al. [37]	2	3	2	2	2.25
Diamond et al. [23]	2	2	3	4	2.75
Friedman et al. [30]	5	5	4	3	4.25
Friedman et al. [1]	5	5	5	4	4.75
Friedman et al. [34]	2	2	4	3	2.75
Gjelsvik et al. [6]	5	5	4	2	4
Haggard et al. [5]	2	2	4	3	2.75
Hesse [31]	5	3	3	5	4
Irani et al. [30]	3	2	4	3	3
Love & Shah [2]	5	4	3	3	3.75
Maylahn et al. [43]	3	3	4	4	3.5
Portnoy et al. [8]	2	2	3	3	2.5
Rudolph et al. [38]	4	3	3	2	3
Schumacher et al. [32]	3	3	3	3	3
Shah et al. [41]	2	2	4	5	3.25
Vest et al. [26]	4	5	5	5	4.75
Wang et al. [40]	2	2	4	3	2.75

Challenges Faced

From these articles, we identified four primary challenges, which include data collection, cost, IT infrastructure and usability. We identified a total of 16 barriers for the four identified challenges, which include data collection (n=6; 38%), IT Infrastructure (n=4; 25%), cost (n=3; 19%), and usability (n=3; 19%). Each challenge will be discussed separately.

Data Collection Challenges

The availability of data sharing is necessary to maximize knowledge and potential health benefits (Walport, 2011). We have identified six key barriers to data collection.

1. *Linking data to multiple sources.* State agencies face major challenges in collecting data from multiple sources. There are political barriers organizations face when sharing data between agencies within a state (Vest, 2013). State health agencies (SHA) and local health departments (LHD) must exchange data to have a complete picture of the health of the communities they serve (Vest, 2013). Local health data alone does not give a complete understanding of health problems in a population. Data on determinants of health, including health behaviour and access to health care services, are rarely available locally. In this case, linking data across multiple entities in a state gives a more complete representation of a community's health (Simon, 2001). A variety of sources can be integrated into one system. Examples include survey data, administrative claims data, program administration data, vital statistics, and private industry data (Studnicki, 1995). Vital statistics, such as birth and death, can be easily obtained. Birth and death data has often become a common starting point for WDQS (Gjelsvik, 2006). However, birth, death, and other key vital statistics, although important, may not give a complete picture of a population's health. Basic vital statistics are useful but are often missing the underlying social determinants of disease and are lacking a comprehensive description of an affected population (Solet, 1999).

2. *Human resources.* There are human resource challenges that contribute to data retrieval problems, which can lead to data sharing challenges. In many cases, data may not be retrievable because of high staff turnover or other administrative factors (Baldwin, 2009). Poor knowledge transfer to new IT staff can prevent access of data. Knowledge about the access, data structures, and data dictionary are provided by the documentation. The IT field is known for poor

documentation processes, which has an adverse impact. New employees with limited documentation may face challenges in understanding or interpreting data (Niu, 2009).

3. *Technical barriers.* Technical barriers also contribute to data sharing gaps. Fostering greater knowledge about IT/IS, public health informatics, and increasing awareness of data exchange issues can help narrow these gaps. Large data sets are typically not accessible and the quality of these data are often poor. The requirements for centrally cleaning and standardizing the data and data manipulation may delay the beginning of analysis. There is almost no orderly method for reporting findings back to the data holders. Data from many information systems might not be shared easily, or that might not have been a requirement of the system at the time of its development (Savel, 2012). Many systems are using old technologies which are not compatible with newer ones.

4. *Political barriers.* Political issues include having the necessary ‘buy-in’ from policymakers and governmental officials. Agency employees may have attitudes that inhibit data sharing, such as sense of general distrust and negative prior experiences (Keller, 2009). There are also issues of mistrust between data providers and users. Providers anticipate misrepresentation, misuse, or intentional abuse of the data (Anderson, 2009). Organizations that have invested time in data collection may be reluctant to share as organizations with greater technological capacity could gain the majority of credit. Such issues are a major challenge for agencies in a low resource setting (Chandramohan, 2008). In this case, sharing data may result in losing control and ownership over data.

5. *Data not compatible.* One of the key challenges is integrating data sources that had never before been combined (Zinn, 1994). Several data systems that currently exist are generally not compatible with one another (Love, 2006). Some legacy WDQS systems were developed by

early adopters in the 1990s that are not compatible with newly developed systems. Unless such data are manually manipulated, they will not transfer successfully from one system to another.

6. *Challenges in acquiring data that is meaningful and useful.* Data may be missing, incomplete or may not be relevant to the end users goals (Friedman, 2006). One of the key challenges is acquiring complete and usable data (Savel, 2012). Data have little utility unless they are meaningful or tied to an objective (Lawler, 2004). Incomplete or missing data significantly lowers the utility of WDQS.

IT Infrastructure

Technology provides countless opportunities for improvement in efficiency and productivity in many activities of public health surveillance (Lee, 2012). It can streamline health data by making data more readily available and having the capacity to store and maintain large data sets. While information technology is useful to public health, there are many challenges. Technology challenges with the dissemination of local health data are a big problem (Savel, 2012). We have identified four key barriers that contribute to IT Infrastructure challenges.

1. *Challenges with privacy/sensitive data.* WDQS pose a unique challenge to protect identifiable health information. Because the results cannot be anticipated, agencies take special precaution to protect the underlying data or aggregate tables (Love, 2006). There are potential risks of small cell counts in particular sub-groups. Rates of small cell sizes may not be reliable, and small cell numbers may lead to violation of individual privacy. Disclosure risks may be high when there are relatively few people with demographic characteristics, such as sex, age, and race, in small communities (Rudolph, 2006).

2. *Technology options.* There is no single approach to implementation as public health agencies differ in structure, governance, budgets, and information technology (Love, 2006). There are

many different ways to deploy a WDQS, and each state has taken a different path (Gjelsvik, 2006). A state may design and develop its own system, adopt a system from another state, purchase a web-query software package, or purchase commercial off-the-shelf software from a vendor. Initially, organizations aspire to develop their own system that are fully integrated and are linked to multiple data sources. A limited number of states have the organizational buy-in, political support or funding to support such efforts (Savel, 2012).

3. *Insufficient data storage.* While there is valuable health data available, it is important to have the hardware and software capabilities to store and maintain the data. Health data can consume large amounts of data storage space, and health information technology is not always keeping pace (Manos, 2014).

4. *Browser compatibility issues.* Some systems must be operated with a specific web browser, leading to browser incompatibility issues (Hui, 2014). In other cases, their options may be dictated by the different operating systems they are using. Some legacy operating systems are using older versions of Internet Explorer and are not compatible with current WDQSs. In other cases, some WDQS are designed for optimal use with browsers such as *Firefox*, *Google Chrome*, or *Safari*. Therefore, browser incompatibility is highly problematic for end users (Hui, 2014).

Cost Challenges

State agencies face challenges in receiving adequate funding for WDQS projects. WDQS primary costs consist of development, implementation, maintenance, and innovation (Friedman, 2006). Generally, spending is lower on health IT projects, about 50% less than other sectors. Overall, it is difficult to establish a return on investment (ROI) and measure the impact (Friedman, 2006).

1. *Cost of software development.* The cost of software development is expensive, as wages for IT professionals are high. The average hourly wage in the United States for experienced software developer is in excess of \$60 per hour (Melber, 2013). Public health software development is more complex and may be more costly than other industry and professions (Studnicki, 1995). It is important to hire highly skilled technical workforce to ensure the optimal result.

2. *Cost of hardware/software.* Hardware, software, and networks can be expensive to maintain and support (Hui, 2014). Licensing is expensive, which can be a financial burden for state agencies (Gjelsvik, 2006). Newer technologies, such as cloud computing and virtualization, may save costs in the long term but require a substantial investment in technology, which can be difficult for public health agencies.

3. *Cost of agency staff/headcount.* It is necessary to have adequate staff for data analysis, dissemination, community health assessment, and program evaluation activities (Friedman, 2006). While there may be enough funding for initial cost for development staff, there is still a need for maintenance. In the long term, staffing resources are important to support and maintain the systems.

Usability Challenges

Usability testing informs us of key problems with the functionality, the ability to retrieve data, and the usefulness of the data. Usability is determined by user-computer interactions and by the degree of successful degree to complete an intended task. Through usability testing, evaluators detect problems that suggest improvements. Little is known about factors that contribute to usability problems with WDQS (Bennett, 1984). We have identified three barriers as contributing factors to usability problems.

1. *Lack of end user evaluation.* There needs to be greater attention to technology and specific needs of its end users. End users should be consulted to ensure their system and data needs are met. One of the key problems found in the literature is the lack of connection between the WDQS program provider and the participants or end users (Haggard, 2006). This lack of connections has made WDQS programs difficult to evaluate.

2. *Insufficient system evaluation.* Systems should be evaluated periodically to ensure that they are serving a useful public health function and are meeting their objectives (Thacker, 1988). One of the key objectives of WDQS is to have an impact on public health (Haggard, 2006). This process (Figure 2) begins at the data collection phase and continues with the development, dissemination, and intervention phase. The long term goal or outcome is to have a public health impact.

3. *Lack of usability testing.* Usability testing is a process in which a small group of evaluators rates a user interface (UI) against a checklist of standards to find usability problems (Monkman, 2013). Through usability testing, evaluators detect problems that suggest improvements. WDQS have had problems with usability as they are difficult to use and navigate (Love, 2006). Given that UI designs vary from state to state, there is not a uniform method available to conduct usability testing. Overall, usability testing is important, as a poor usability contributes to low adoption and users not returning to the site (Rose, 2005).

Study Type and Focus

We categorized our findings (Table 3) into study types and focus areas. Each article was classified into one of four study types that include commentary (n=7; 38%), evaluation (n=5; 28%), survey (n=4; 22%), or methods (n=2; 11%). Articles with a technology (n=14; 78%) focus were highest, followed by data sharing (n=10; 56%) and privacy/risk (n=2; 11%). Each article analyzed was summarized individually (Table 4).

Table 3 References of articles included in the review and summaries of their basic characteristics.

Citation	Topic	Type	Focus
Braithwaite et al. (2006)	An evaluation of Utah's IBIS-PH Web Based Data Query System	Evaluation	Technology
Choi et al. (2012)	Past and future of Health Surveillance	Commentary	Data Sharing, Technology, Privacy/Risk
Cohen et al. (2006)	An evaluation of the MASSCHIP, a web based system from Massachusetts	Evaluation	Technology
Diamond et al. (2009)	Proposed method to collect health data	Methods	Data Sharing
Fegan et al. (2011)	The potential of data sharing	Commentary	Data Sharing
Friedman et al. (2006)	Desired characteristics of WDQSs	Survey	Technology
Friedman et al. (2006)	Envisioning the future of WDQSs	Commentary	Technology, Data sharing
Friedman et al. (2010)	The importance of using the PopHR record system for health data	Evaluation	Technology, Data sharing
Gjelsvik et al. (2006)	States unique approaches to WDQS development	Survey	Technology
Gjelsvik et al. (2009)	An evaluation of Rhode Island's WDQS system	Evaluation	Technology
Haggard et al. (2006)	Proposed methodology for WDQS evaluation	Methods	Evaluation
Hesse (2011)	The growth in the sources of health data	Commentary	Privacy/Risk, Data Sharing
Irani et al. (2011)	The greater need for community level health data	Commentary	Data Sharing, Technology
Love & Shah (2006)	Organizational issues with WDQSs	Commentary	Data Sharing, Technology
Maylahn et al. (2005)	The Need for health surveillance data for the aging populating	Survey	Data Sharing, Technology
Portnoy et al. (2014)	Responding to the needs of local health data	Survey	Data Sharing/Technology
Rudolph et al. (2014)	Risks of small cell counts and disclosure	Commentary	Privacy/Risk
Schumacher et al. (2006)	The evolving changes in assessment for health data	Commentary	Technology, Data sharing
Shah et al. (2014)	Proposed design of an open source based WDQS system	Commentary	Technology
Vest et al. (2013)	Data sharing gaps between agencies	Evaluation	Data Sharing
Wang et al. (2008)	Development of the New York health surveillance system	Evaluation	Technology

Table 4 Summary of Findings

Citation	Summary
Braithwaite et al. [35]	This article summarizes the design and development process of the open source Utah IBIS-PH WDQS system. The system is easily adoptable by other states and can be modified by a broad base of software developers.
Cohen et al. [37]	The Massachusetts based MASSChip WDQS links multiple sources within the state. One major disadvantage is that it requires software to be downloaded, and cannot be used on the web.
Diamond et al. [23]	To protect the privacy of health data and increase efficiency new models to collect health data should be implemented. <i>The networked model for analyzing population health data should be considered.</i>
Friedman et al. [30]	WDQS functionalities from 27 states were evaluated by the author and fourteen participants were surveyed. Overall, advanced and basic functionalities are found infrequently in WDQS systems.
Friedman et al. [1]	States have also developed WDQSs with innovative functionalities but need to further evaluate its use and how the data is consumed.
Friedman et al. [34]	The <i>PopHR</i> health population health record format corrects basic problems in the comprehensiveness, timeliness, and use of population health information. Existing WDQSs could be further developed using PopHRs.
Gjelsvik et al. [6]	Ten states were surveyed on the path selected in WDQS development. Overall, there was no dominant method to implementation as states took unique approaches.
Haggard et al. [5]	A logic model was developed to evaluate WDQSs. The model used inputs, constraints, program activities, program outputs, and outcomes to determine the effectiveness of WDQS programs.
Hesse [31]	Methods need to be used to ensure that health data is credible, trustworthy, and evolve over the next century. Data needs to be integrated rather than isolated, and we need to 'connect the dots'.
Irani et al. [30]	WDQSs are valuable for community health assessments and for state and county health departments to quickly share their findings with multiple stakeholder groups. Taking these measure can lead to interventions.
Love & Shah [2]	Funding and staffing do not guarantee successful implementation of WDQSs. It is important that states have the passion, commitment, drive, and are organizationally strong.
Maylahn et al. [43]	Greater emphasis need to be made on the dissemination of health data relevant to the elderly and aging population.
Portnoy et al. [8]	A national and state national system of data collection which would improve WDQSs and the overall health of the population.
Rudolph et al. [38]	This article discusses the importance of cell suppression, aggregation of intervals, and other methods to reduce the risk of small cell counts.
Schumacher et al. [32]	The CDC's assessment initiative, a program aimed at strengthening WDQS systems and practices used to improve data accessibility is a valuable program that has had little influence.
Shah et al. [41]	This article demonstrates the functionalities of <i>Profile-IQ</i> , an open source WDQS software built on "cloud computing". Cloud computing technologies are primarily enabled by virtualization, which is of lower cost.
Vest et al. [26]	This paper explored data sharing gaps amongst state and local health agencies. For vital records and reportable diseases, data sharing gaps were the norm.
Wang et al. [40]	This health surveillance system developed in New York, matches hospital cases to vital records. Such a system helps conduct trends analysis on birth defect prevalence and mortality with data query and visualization capabilities.

DISCUSSION

Using a systemic review of literature, we identified sixteen unique barriers that contribute to key challenges with WDQS dissemination. Eighteen articles found in the systematic review were used in the analysis. None of the eighteen articles were excluded from analysis, including articles with lower quality scores. All of the eighteen articles analysed provided some form of scientific contribution to this review. Four of the primary challenges are further addressed based on findings from this review.

Data Sharing

Progress in data sharing in public health has been slow in comparison to other science disciplines. Disciplines such as immunology, genomics, and environmental science have been more receptive to data sharing. This concept has not been fully embraced by public health, as there are many data sharing barriers and challenges. Data sharing barriers limit both the progress of research and public health benefit (Walport, 2011). Unless data sharing barriers are better understood, solutions will remain ineffective (Van Panhuis, 2014). There is no uniform framework or operational guidelines for data sharing in public health. There is currently no uniform policy for the ethical collection, storage and use of public health data (Lee, 2012). Public health is highly interdisciplinary and includes a wide range of data sources always evolving in size and complexity. Most data are derived directly from populations, which are monitored by health agencies. These include data such as vital statistics, demographic information and other records (AbouZahr, 2007).

Portnoy (2014) recommended a national and state national system of data collection which would improve population health (Portnoy, 2014). However, a survey conducted with several states concluded that such an undertaking would not be feasible since dissemination needs,

planning, and practices varied widely. Hesse (2011) emphasized that public health surveillance should shift from the idea of data isolation to data integration. New methods to converge public health data should replace segregated thinking (Hesse, 2011). Overall, it is difficult to establish uniformity when there are no commonly accepted standards or guidelines (Hesse, 2011).

IT Infrastructure

There are many technical barriers such as a lack of standardization of common protocols across agencies, variance in data quality, and incompatibility between databases. Some of these barriers prevent agencies from linking across multiple data sources. However, there are a limited number of states that have been able to overcome these challenges. As an example, *MassCHIP* provides access to 36 data sets from the Massachusetts Department of Public Health. Its user base includes stakeholders such as state employees, epidemiologists, agencies, community advocates, researchers, and journalists. *MassCHIP* also has four unique ‘entry points’ based off user experience. In this case, a novice user is directed to a broader sample of data, while an advanced user can query on a specific topic of interest (Cohen, 2006). Illinois developed the *IQuery* system in 2012, which connects 17 health categories and more than 150 different health indicators such as communicable diseases and births. The Missouri Information for Community Assessment (MICA) provides users with age-adjusted rates of hospitalizations, procedures and emergency room visits. Georgia’s OASIS system provides users with advanced statistics such as Health Adjusted Life Years (HALY), Years of Life Lost (YLL), and premature mortality. Rates ratios, such as odds ratios and relative risk, and statistics, such as years of life lost and premature mortality, are enhanced functionalities. Only a few states offer such functionality. In the future, more improved technological solutions will continue to replace less-efficient/high maintenance solutions. There is now software available that is less programming intensive, and that trend will

continue (Love, 2006). Such software packages are expensive, but may overcome various technical barriers. Despite this trend, states need to ensure that they have knowledgeable staff and personnel to support and maintain systems that are less programming intensive.

Cost Barriers

Cost barriers restrict organizations from maintaining and implementing WDQS. An organization may lack financial resources for development and skilled human resources for interpretation of data. The skill set of available personnel and cost can significantly impact a state's implementation decision. When a state designs and develops their own system, it is vital to have both the resources and knowledgeable IT staff. States that do not have knowledgeable IT staff or resources have the option of partnering with another state. Historically, states have mentored and frequently shared knowledge with one another for building solutions. For example, Utah's *MATCHIIM* system and their *IBIS-PH* system have been shared with multiple states (Love, 2006). The *IBIS-PH* system was written using open source technology, which is of little or no cost. If a state adopts such software, they must also have a budget to hire IT staff that is familiar with the development and maintenance of open source coding. Also, *IBIS-PH* runs on Unix and Linux servers, which requires specific expertise. Open source refers to source code that is available for any modifications, as developers see fit. The maintenance of open source software is labor intensive and requires a strong level of technical expertise (Braithwaite, 2006). Commercial software is purchased and owned by an organization and is typically licensed for an organization's own use. In most cases, commercial software runs on the *Microsoft Windows* platform, the most widely used operating system (Melber, 2013). Although commercial packages are expensive, the software manufacturer generally provides helpdesk and technical support as part of their license agreement.

Usability

Researchers have developed models for evaluation, but little is known if they are being used. To illustrate, Haggard (2006) developed a logic model to evaluate outcomes that includes inputs, activities, and outputs. Inputs include elements such as money, staff, and computer equipment. Activities include elements such as designing a software interface, defining system requirement, and gathering data. Outputs include elements such as health indicator profiles, data availability at the community level, and timely public health data on priority measures. Outcomes include activities such as improved awareness and understanding of specific public health issues, community motivation to change agent, and better health status (Haggard, 2006). Friedman (2006) recommended that we attempt to answer questions such as who utilizes WDQS, for what purpose and if employing them improves the health of the community (Friedman, 2006).

Conclusion

The late 1990's until 2005 can be seen as an initial innovation phase, as several states were in the process of implementing and adopting WDQS. Following that period, expectations for robust systems increased and these expectations have not been met. States are now faced with the challenge of upgrading and improving their health data systems. The review highly emphasized technical challenges state agencies face with implementation. Further research needs to be conducted on the importance and magnitude of these challenges. In the long term, we anticipate further research can lead to the implementation and best practices of systems that are robust, easy to use, complete, and have a public health impact.

Chapter 3:

Manuscript #2: Challenges to the Implementation of Web Based Data Query Systems: a Delphi Consensus Study

ABSTRACT

Background: State health agencies (SHA) and local health agencies (LHA) face several challenges with the dissemination of local health data using WDQS. A recent systematic review looking at these challenges identified cost, data collection, IT infrastructure and usability as major challenges for SHA and LHA. This study aimed to utilize expert consensus to identify the most important items that contribute to these challenges to help guide future research.

Methods: A total of 17 experts, including researchers and public health professionals, participated in a three-round Delphi process. In the first round, four topics were represented on a 42 item questionnaire using a 5-point Likert scale, along with free text responses. Free text responses were analyzed thematically leading to a series of items for a second Delphi Round. Participants were given an opportunity to revise results in the third Delphi round. Consensus was defined at interquartile range (IQR) ≤ 1 . Results from this study will be used to develop a questionnaire for a subsequent study that surveys state health coordinators.

Results: A total of 13 (31%) of the 42 items presented in the initial questionnaire reached consensus. Of the 15 items presented from the free text responses, 11 (73%) of the items met consensus. In the final questionnaire, 21 (50%) of the 42 items presented reached consensus.

Conclusion: Experts confirmed that software development costs, inadequate human resources, data sharing gaps, a lack of political support, and poor data quality contribute significantly to dissemination challenges. We must seek solutions to overcome data sharing gaps between SHA

and LHA in order to have more complete data that is useful for end users. There is a need to understand how data are consumed and utilized and their impact on public health.

BACKGROUND

There is clear evidence to support that the availability of high quality data at the local level can lead to targeted interventions, impact policy decisions (Van Panhuis, 2014), reduce health disparities, and improve health care delivery systems (Portnoy, 2014). Web Based Data Query Systems (WDQS) make local health data easily available to the public. Implementation is limited, as many states faces use outdated technologies such as static reports and ad-hoc based systems. A recent systematic review identified four primary challenges organizations face in the dissemination of local health data. These challenges identified include cost, data collection, IT infrastructure, and software usability.

Cost Challenges

WDQS are expensive to develop and maintain. There are several cost challenges that include staffing, IT development, hardware, software, and maintenance. Cost is a major barrier as many states have struggled to obtain support and funding for health related IT projects (Ortiz, 2003). State budgets are generally year to year, which prevents long term planning. Overall, health information technology (IT) spends about 50% less on IT infrastructure than other sectors (Bates, 2002). It is necessary to have ‘buy-in’ from policymakers and governmental officials for such projects. Having political support and ‘buy-in’ is important in order for agencies to receive financial and human resource commitments (Nsubaga, 2006). Evaluating cost enables us to understand how cost may be a barrier to improved IT development and helps determine if state agencies are effectively spending their resources.

The cost of software development is high as wages, as the average hourly wage in the United States for an experienced software developer is over \$60 per hour (Clemens, 2013).

Some states have taken the path of hiring consultants in their software development process, rather than hiring permanent staff. Software consultants or independent contractors can benefit the organization in the short term without a long term solution. There may be a budget for development but not a long term plan that includes benefits and sustainable employment. States have limited IT budgets, making the cost for software development a major challenge.

Data Collection Challenges

One of the key challenges is acquiring complete and usable data (Savel, 2010). Data collection with WDQS development includes activities such as data acquisition, data sharing, and data manipulation. Data acquisition consists of collecting data from multiple sources such as state health agencies (SHA), local health agencies (LHA), and private organizations. Data sharing are necessary in order to have a complete picture of the health of the community (Vest, 2013). Incomplete data in local practice create inefficiencies and limit the ability to use the information for decision-making. There are many data sharing gaps that exist. Data sharing between state and local agencies is important in having usable data through WDQS as well as to increase IT capabilities at the state level. Barriers to data sharing are caused by both technical and non-technical factors (Vest, 2013). Examples of technical factors include missing primary identifiers, disaggregation of indicators, and the inability to identify data elements. Examples of non-technical factors include reluctance of agencies or organizations/hospitals to release data, challenges in institutional review board (IRB) approval, legal and political issues (Nsubaga, 2006).

IT Infrastructure Challenges

While there are valuable health data available, it is important to have the hardware and software capabilities to store and maintain the data. Health data can consume large amounts of

data storage space, and health information technology is not always keeping pace (Manos, 2014). Poor bandwidth and IT infrastructure can limit the capabilities of a software application, in particular over the internet. Poor IT infrastructure can contribute to poor software latency for the end user. Latency is the amount of time it takes information to travel from its source to its destination. Poor latency can cause delays in retrieving data and decrease the utility of a software application. Along with technical challenges, there are communication gaps between the project staff and software developers. Communication problems between them are a major factor in the delay and failure of software projects (Curtis, 2010).

Usability Challenges

Usability informs us of key problems with the functionality, the ability to retrieve data, and the usefulness of the data. Usability is determined by user-computer interactions and by the degree of successful completion of an intended task (Hornbaek, 2006). There have been usability problems reported with WDQS, as they have been reported as difficult to use (Luck, 2006). Data may be missing or incomplete or may not be relevant to the end users goals. Websites that are user-centric and easy to use have higher user satisfaction. Websites with poor usability lead to poor perception and satisfaction, resulting in users not returning in the future (Hornbaek, 2006).

The Delphi Study

This study aimed to ascertain the importance of barriers organizations face in the dissemination of local health data and their usability from the perspective of experts. The Delphi method was chosen, due to its suitability for areas of inquiry where only uncertain or incomplete knowledge exists (Verhagen, 2001). It is an iterative, multi-stage, group facilitation process designed to transform opinion into group consensus that involves a series of structured questionnaires (Hasson, 2000). Using experts, the Delphi method seeks to gain a group

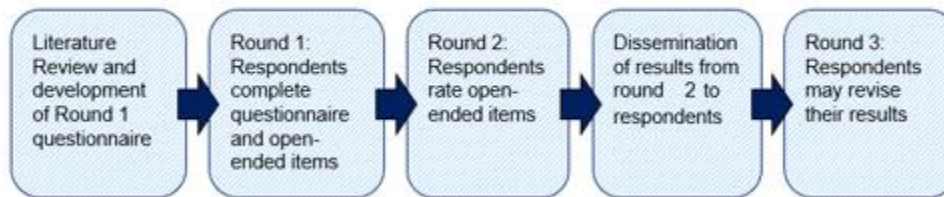
consensus via individuals on a specific topic. Consensus in a Delphi study is defined as a “general agreement of a substantial majority” (Verhagen, 2001). The Delphi study involves a series of questionnaires using multiple rounds. We chose to use a controlled feedback method known as ‘quasi anonymous feedback,’ in which names of the participants are only known to the researcher (Lofmark, 2004). It is known as ‘quasi’ anonymous because complete anonymity cannot be guaranteed as the researcher knows the name of the panel member and their responses. Anonymity amongst the participants allows a true consensus to be reached and eliminates many problems that arise from bias and peer influence. Anonymity assures that specious persuasion does not occur, since anonymity reduces the effect of dominant individuals (Fischer, 1978). It is termed “controlled” because the facilitator decides on the type of feedback and its provision (Gratch, 2012). The Delphi study was driven to answer the following question: What findings of the systematic review are important and which topics should be prioritized for future decision making regarding best practices of WDQS implementation? Items of importance from this study will be used for a subsequent study that surveys state health coordinators from all fifty states.

METHODS

Delphi Process

We conducted a three round Delphi study (Fig. 3) to confirm findings of the systematic review and to prioritize key barriers and challenges that exists for WDQS implementation at the state and local levels. We selected three rounds as it is efficient and typical of most Delphi studies (Jones, 1992). Although the possibility of more than three rounds is offered, there is a need to balance time and possible participant fatigue (Jones, 1992). The goal was to reach expert consensus.

Figure 3. Schematic of Three Delphi Rounds



Selection of Panel Members and Recruitment

A purposive sample of 17 (n=17) experts were selected for this study. Purposive sampling is non-probability sampling that uses the judgement of the researcher to recruit participants. Purposive samples are often used in Delphi studies (Hasson, 2000). Since expert opinion is sought, a purposive sample is necessary where people are selected not to represent the general population, rather their expert ability to answer the research questions (Fink, 2009). It is recommended that 10 to 18 participants are used on a Delphi panel (Okoli, 2004). An expert has been defined as a group of ‘informed individuals’ and as ‘specialists’ in their field or someone who has knowledge about a specific subject (McKenna, 1994). Careful selection of the panel of experts is paramount to a successful Delphi study (Jairath, 1994). It is recommended that participants have either published articles, taught courses about the topic, or their primary job responsibility is related to the area (Gibbs, 2001). In this case we required that panel members had met at least *one* of the following criteria: 1) have published literature in peer reviewed scholarly journals, or 2) a significant portion of their job responsibility involves the dissemination of local health data using WDQS. We searched online for contact information, for authors with relevant publications, and the authors were contacted by email. Those with experiences working with the dissemination of WDQS were identified through multiple channels. We searched on websites including state and local health department websites and *Naphsis*. *Naphsis* is an organization that has a mission is to provide health information to

improve the public's health. *Naphsis* also manages a publically available list indexed of active WDQS by state. We also searched the Department of Health website from states that have deployed WDQS and attempted to find knowledgeable staff. All participants were contacted through email.

Ethical consideration

The Delphi study participants were informed that their participation in the study was entirely voluntary, and they implicitly consented to participate by completing the questionnaire. They were also informed that their results were confidential as well as their names. Their names are not provided to other participants in the study. Ethics approval for the study protocol was received from the Human Subjects Protection Office at the University of Connecticut Health Center.

Procedure

Participants were invited to participate through email. All rounds of the Delphi study were conducted through email, and each questionnaire was administered in a Microsoft Word format. Participants were given approximately one week to complete each round, and a reminder e-mail was sent following the deadline. To ensure strong retention of expert involvement, the study was set at three rounds. It is acknowledged that having a planned number of rounds is an indicator of good quality in designing a Delphi study (Diamond, 2014).

Overview

In the first round, we administered a 42 item questionnaire. In addition, there was an open text response section. In the second round, we administered a 15 item questionnaire based on a filtered list of open ended responses from Round 1. Items that did not meet consensus in Round 1 or Round 2 were presented to the participants in Round 3. In the third Round,

participants were given an opportunity to revise their results from Round 1 and Round 2 in order to come to consensus with the group.

First Round Questionnaire

The first round of Delphi study consisted of a questionnaire based upon extensive review of the literature (Hsu, 2007). There were two components, which included a quantitative based questionnaire, and a qualitative open-ended response section. It is recommended that quantitative and qualitative data are used in a Delphi study in order to improve group judgement (Rowe, 1999). Once completed, experts were required to return the questionnaire via email for statistical aggregation and review.

Quantitative Component

A four topic 42 item questionnaire was administered. The four topics included cost, data collection, IT infrastructure, and usability. Within each topic were a series of items that were relevant to that topic. Participants were asked to rate the importance they would give to each item on 5-point Likert scale of 1 to 5 (Unimportant – 1; Little importance – 2; Moderate importance– 3; Important – 4; Very important - 5). They rate importance based on their experience in research or practice. Rating importance helps prioritize topics for future research and which items are important based on expert opinion.

Data Analysis

Quantitative data from the 42-item questionnaire was compiled in Microsoft Excel and then imported to SPSS v 19.0 for analysis. The mean, median, and interquartile range (IQR) were calculated for each item in SPSS. The IQR is calculated to assess the extent of agreement between the experts on the changeability of each item (Jones, 1995). An IQR is a measure of statistical dispersion, being equal to the difference between the upper and lower quartiles, and

consists of the middle 50% of the observations (Sekaran, 2003). An IQR of less than 1 indicates that more than 50% of all opinions fall within 1 point on the scale (De Vet, 2000). It is a frequently used measure in Delphi studies, and it is generally accepted as an objective and rigorous way of determining consensus (Von der Gracht, 2012). An IQR of 1 or less can be considered a good consensus on a 5-point Likert scale. If an item reaches consensus in Round 1, the process stops there as the research question is answered. If consensus is reached, theoretical saturation is achieved (Skulmoski, 2007). Theoretical saturation is the phase in data analysis in which all concepts in the theory are well-developed. In this case, when an item reaches consensus, no further data needs to be collected for that item. Items that do not meet consensus are carried over to Round 3. It is also possible that there are items in a Delphi study that have high consensus but are not considered important by the experts (Vendelanotte, 2010). In this case, we report the item as being in consensus but not being important. In this situation, an item with a low IQR and low mean would indicate that experts are in agreement that an item was unimportant or of low importance.

Qualitative Component

In the optional open text response sections, participants listed additional items of importance. A first round questionnaire with open-ended questions is consistent with most Delphi studies (Farmer, 1998). The open-ended component of the study reduces the chances of excluding items that the researcher may have omitted. In a Delphi study, open-ended questions are recognized to increase the richness of the data collected (Bond, 1982). In this case, items not found in our systematic review may have been reported by participants.

A cumulative list of all open text items were analyzed using Microsoft Excel. *Qualitative content analysis* was used to identify and interpret themes in the qualitative material. Qualitative

content analysis is a technique for systematic text analysis, which uses themes to identify qualitative responses. Researchers regard content analysis as a flexible method for analyzing open ended, text based responses (Cavanagh, 1997). This type of design is usually appropriate when existing theory or research literature on a phenomenon is limited. Themes are not pre-categorized and instead are formed from the flow of data (Kondracki, 2002). Data was sorted, and common themes were found amongst the responses. These themes were filtered and interpreted. Fifteen new items were developed from panel members' suggestions from the open ended section. The final list of 15 items were compiled and carried over to Round 2.

Second Round Questionnaire

A compiled list of 15 participant generated items from Round 1 were distributed to participants. Participants rated each item on 5-point Likert scale, in which 1 indicated 'unimportant' and 5 indicated 'very important.' Quantitative data from the Round 2 questionnaire were entered into Microsoft Excel and then imported to SPSS v 19.0 for analysis. Items that achieved an IQR of ≤ 1 met consensus, and items with an IQR > 1 were carried over to Round 3.

Third Round

In the final round of the Delphi process, respondents were provided a list and median score of items from Round 1 and 2 that did not meet consensus. Participants were given an opportunity to revise results from Round 1 and 2. We calculated the mean, median and IQR after Round 3. Items with an IQR ≤ 1 met consensus. Items with an IQR > 1 were not in consensus. Participants also were given an opportunity to qualitatively provide a reason for not coming to consensus for items they chose not to revise. We use this feedback and report these findings in our discussion sections.

RESULTS

Response Rate

Forty-two experts were invited to take part in the study, and 17 experts agreed to do so. Reasons for refusal included time constraints, given there are multiple rounds and steps involved in a Delphi study (Robinson, 2014). It is also known that experts tend to be busy, and the multiple rounds can be cumbersome (Wentholt, 2010). Of the 17 participants that agreed to participate, 16 (94%) of them completed Round 1. Two participants dropped out of the study after Round 1, and 14 participants (82%) completed both Round 2 and Round 3.

Participants

Of the 17 experts, 7 experts published literature on the development, evaluation, or the dissemination of WDQS. Those who published literature included researchers, college/university professors, and private consultants. The remaining 10 experts are public health professionals. This group consisted of experts such as epidemiologists, health directors, and other members of the public health community.

Summary of rounds

The summary for each round is presented in Table 1. This table presents an overview of the scoring for each round and outlines the following. Items that met consensus are presented in Table 2. Items that did not meet consensus are presented in Table 3.

Rounds 1 and 2

In Round 1, 13 of the 42 (31%) of the items met consensus, and 29 items and were carried over (69%) to Round 3 because consensus was not reached. In the second round, 10 out of the 15 items (67%) met consensus and were retained. The five items (33%) that did not meet consensus in Round 2 were presented in the final round (Round 3) for an opportunity to revise.

Final Round

All items that met consensus are shown in Table 2. Overall, consensus was reached in 21 of the 42 items (50%) that were originally presented in Round 1d. Overall, 60% of the usability items presented met consensus, followed by cost (54%), data collection (46%), and IT infrastructure (33%). Of participant generated/qualitative items, 11 of the 15 items (73%) met consensus.

Results by Topic

Cost

Overall, the cost of staffing and cost of system development were rated as the most important items (Table 2). This includes the cost of adequate public health staff (mean=4.33, IQR = 1), the cost of system development (mean=4.14; IQR = 1), and IT staff (mean=4.07; IQR = 1). Participants were in consensus that staffing and development are the most significant barriers they face. The cost of IT technical support for state agency staff (mean=3.46; IQR=1), IT technical support for end users (mean=3.13, IQR = 1), and hardware/servers (mean=3.00; IQR = 0.25) are less of a challenge. The cost of data storage (mean=2.40; IQR = 0.5) was a less important barrier.

Data Collection

The most significant barriers to data sharing reported were the quality of data, the ability to connect to multiple sources, and the utility maintained for end users. The most important barriers are acquiring data that are useful and meaningful (mean = 4.63; IQR = 0.75), acquiring data that have been requested by end users (mean = 4.42; IQR = 1), acquiring data from multiple sources (mean = 4.43; IQR = 1), and collecting data in a timely manner (mean=4.21; IQR = 1).

IT Infrastructure

IT infrastructure barriers were rated of lower importance in comparison to cost and data collection. Two items were in consensus, including collaboration with software developers to ensure systems needs are met (mean = 3.57; IQR= 1) and challenges in decision making on software options (mean = 3.53; IQR = 1). Collaboration amongst IT staff consists of working with project managers, IT staff, business analysts, and other personnel to ensure that the design specifications are met. Challenges in decision making consisted of deciding which path a state chosen in their development. Some design their own system, while others hire consultants or adopt a system from another state.

Usability

The most important barriers to usability pertained to accessibility of data. Participants agreed that having meaningful and useful data (mean = 4.40; IQR = 1), missing of incomplete data (mean = 3.80; IQR = 1), and the quality of data output (mean=3.71; IQR = 0.75) are important. Participants agreed that difficulty in interpreting drilldowns (mean=3.5; IQR = 0.75) and a website that freezes up (mean=3.33, IQR = 1) were not as challenging as problems with accessibility of data.

Table 5: Overall Consensus from Round 1 and Round 3

Topic	Round 1 Consensus			Round 3 Consensus	
	Total # of items (n)	Items in Consensus (n)	%	Items in Consensus (n)	%
Cost	13	5	(38)	7	(54)
Data Collection	13	4	(31)	6	(46)
IT Infrastructure	6	6	(33)	2	(33)
Usability	10	3	(30)	6	(60)
Total	42	13	(31)	21	(50)

Table 6: Items that met participant consensus using a five-point Likert Scale

Topic	Item	Mean Score	Median	IQR	Round consensus was achieved
Cost					
	Cost to have adequate state agency (public health) staff	4.33	4	1	Round 1
	Cost of system design/software development	4.14	4	1	Round 1
	Cost to have adequate staff/headcount for IT staff (internal)	4.07	4	1	Round 1
	Cost of IT technical support for state agency staff	3.46	3.5	1	Round 3
	Cost of technical support to end users	3.13	3.5	1	Round 3
	Cost of servers/hosting Applications	3.00	3	0.25	Round 1
	Cost of data storage	2.40	2	0.5	Round 1
Data Collection					
	Challenges in acquiring data that are useful and meaningful	4.63	5	0.75	Round 1
	Challenges in acquiring data that have been requested by relevant stakeholders/end users	4.42	5	1	Round 1
	Challenges in acquiring data from multiple data sources across the state	4.23	4	1	Round 1
	Challenges in working with private hospitals and clinics to release data	4.21	4.5	1	Round 3
	Collecting data in a timely manner	4.21	4	1	Round 1
	Challenges in working with public hospitals and clinics to release data	4.01	4	1	Round 3
IT Infrastructure					
	Collaboration with software developers and IT staff ensuring that requested system design needs are met	3.57	4	1	Round 1
	Challenges in decision making on technology options (open source, commercial, or adopt from another state)	3.53	4	1	Round 1
Usability					
	Data are meaningful and is useful for the end user	4.40	5	1	Round 1
	Data are missing or incomplete for end user	3.80	4	1	Round 3
	Quality of user data output (Excel, csv, pdf, html etc.)	3.71	4	0.75	Round 1
	Drill Downs/Data Filers are difficult to understand	3.50	4	0.75	Round 1
	Website freezes up	3.33	3.5	1	Round 3
	Data do not go far back enough in time	3.21	3.5	1	Round 3
Qualitative Items (Participant Generated)					
	Evaluation of end users	4.57	5	1	Round 2
	Standardization of vocabulary	4.50	5	1	Round 2
	Providing context in a way which makes a 'story' of the data	4.50	4.5	1	Round 2
	Hidden costs associated with development	4.42	5	1	Round 2
	A greater understand of how the consumer consumes the information	4.42	5	1	Round 2
	Need for 'user centric' design	4.35	4.5	1	Round 2
	Helpdesk support for end users	4.28	4	1	Round 3
	Data from the private sector	4.07	4	1	Round 2
	Using existing public health surveillance systems and mandated hospital discharge reporting maintained by state department of health	3.72	4	0.5	Round 2
	Rigorous validation of data and statistical algorithm	3.20	3	0.25	Round 2

Table 7: Items that did not meet Consensus after Round 3

Item	Mean	Median	IQR
Collecting data in a timely manner	4.14	1.75	4
Navigation and website buttons are clear, concise, and easy to understand	4.14	5	1.75
Challenges in having 'buy in' from state governments	4.00	4	2
Challenges in linking across multiple data sources	4.00	4.5	2
Cost of system maintenance after deployment	3.71	4	1.75
Availability of IT support staff by email or phone for technical questions	3.71	3.5	2
Collaboration with software developers	3.64	4	1.75
Resources for improving and updating systems	3.57	4	1.75
WDQS links/URL's do not work or links within website do not work	3.57	4	2
Challenges in having 'buy in' from local governments	3.53	3.5	2.25
Cost of software testing and QA/QC control testing	3.47	3.5	1.75
Privacy issues with small cell counts with aggregate data	3.47	3	2
Reliability of systems	3.35	3	3
Systems are not compatible with end user web browser	3.35	1.5	3
Data types mismatch when importing data	2.97	3	3
Cost of commercial software	2.92	3	2.25
Time to process queries is long	2.92	3	2
Challenges in acquiring vital statistics	2.85	3	1.75
Speed and bandwidth in processing user queries	2.57	2.5	2
Capability of hardware/servers to handle large volumes of data	2.50	2	2.5

Discussion

This study aimed to gather first-hand knowledge held by public health professionals and researchers who have expert knowledge on WDQS. Using the Delphi method, we were able to determine the importance of several items found in our systematic review.

In addition, the participant generated section from the qualitative section in Round 1 was beneficial as it presented challenges not emphasized in prior literature. Several participant generated items are rated with strong importance. Items generated by participants, such as standardization of vocabulary, providing context to make a story of the data, and hidden costs with development, were met with strong consensus. Of the items in consensus, 73% of the

participant generated items met consensus in comparison to 38% that we presented from the literature review.

Cost Barriers

This study confirmed that cost is one of the most significant barriers. Good data may be available, but they are lacking trained staff that understood how to interpret the data. Staffing resources are expensive, as many states do not have the budget for increasing headcount. Adequate staffing is important in the development process of WDQS, along with having trained staff that understand the utility of the data. It is important for policymakers and stakeholders in the government to understand the value and importance of WDQS.

Our review of the literature also indicated that the cost of hardware and software was a major barrier. This study found the contrary, as participants did not agree that the cost of hardware and software was a problem. The cost of hardware and servers is mainframes and servers have dropped significantly over the last two decades. The price for servers and mainframes today are a fraction of what they were in the 1990s (Cho, 2011). Similarly, there was consensus that the cost of data storage is *not a challenge*. The cost of hardware, software, and data storage has decreased significantly over the last ten years. The cost of data storages drops an average of 50% every 18 months (Haq, 2010). Data storage chips continue to increase in capacity and can store greater amount data at lower costs. Computer processors speeds and bandwidth has increased significantly over the last decade. Processer speeds have doubled an average every 18 months since 1975 (Roberts, 2002). In the future, we can anticipate that the cost of hardware, software, and data storage will continue to drop in price. This will further reduce the cost burden for state agencies.

Additionally, the study informed us that there is a need for collaboration amongst states, which would help reduce costs. This includes stronger collaboration in development, strategy, and forming partnerships. While states do make their own decisions in technology options, some have adopted systems from other states (Gjelsvik, 2006). Using open-source technologies, states can share their software cost and reduce the level of planning, design, and development costs.

It was qualitatively suggested that that receiving grant money is a challenge, but this item did not meet consensus. Responses vary as some states have greater funding and support than others for WDQs. A median of '4' demonstrated that grant funding was an important topic; however, participants could not agree on this topic. It is likely they could not agree as the availability of grants varies from state to state.

Data Collection Challenges

Data acquisition challenges were reported to be paramount. Acquiring data from private and public hospitals along with other sources across a state were a major barrier. It was reported with consensus that data such as births, deaths, and basic vital statistics were easy to obtain. It may be necessary to have more information than basic vital statistics to understand the overall health of the population.

Participants reported that state agencies faced major challenges in having access to data from multiple sources. Linking data from multiple sources can provide information regarding social determinants of disease, which helps give us a comprehensive description of an affected population (Solet, 1999). According to the Centers for Disease Control and Prevention (CDC), one of the biggest challenges is to find effective ways of combining multiple sources of complex data and information into meaningful and actionable knowledge (Savel, 2012). Participants agreed that one of the key problems is having an understanding of what data are meaningful and

useful. This was reported from the qualitative components in Round 1, which met consensus in Round 2.

IT Infrastructure Challenges

Participants reported that one of the key challenges with IT infrastructure is the retention of staff and knowledge transfer. There is high turnover in the field of information technology as staff is highly transient. IT employees generally have greater allegiance to the industry rather than the organization (Husion, 1992). High employee turnover is a problem that organizations cannot ignore because of the financial burden and adverse impact on deliverables (Abii, 2013). When IT staff leaves, organizations have to attend to the high cost of training and development, burnout of existing staff, and decreased quality of products and services due to the shortage of staff (Abii, 2013).

Interestingly, several items that were highly emphasized in the literature were found to be of little importance in our study. The ability of systems to handle large data sets and small cell counts did not meet consensus. There were also items that met consensus with a low mean indicating low importance. For example, data storage met consensus, with low importance. Overall, there was consensus on two of the six items IT infrastructure items surveyed, of which only one of the items considered to be an ‘important’ challenge. The collaboration in working with IT staff to ensure needs are met was the item reported as an important challenge. The IT development process involves multiple stakeholders that include project managers, software developers, and business analysts. The role of the business analyst is critical as this person acts as a liaison between the software development team and the project staff. The business analyst is responsible for establishing the set of requirements, managing the smaller details that are involved in IT projects, and establishing timelines for deliverables. When organizations do not

account for the small details that very large IT projects entail, it can lead down a difficult and bumpy road (Barton, 2014). This bumpy road can lead to the development of systems that are missing key software functionalities.

Challenges with Usability

This study demonstrates that there needs to be greater efforts in usability testing, in particular with users outside the organization. It has been common practice to conduct usability testing within organizations. The maximum benefit is achieved when usability testing is conducted with typical end users and not the designers or clients (Lee, 1999). Participants also agreed that there needs to be further evaluation of end users' data needs. Such evaluation will help determine if systems are useful and meet their needs.

Many usability issues emphasized in the literature were found of lower importance. For example, issues such as browser compatibility and poor bandwidth were insignificant. As bandwidth, data processing, and quality of hardware continues to improve, such issues will become of even lower importance in the future. There was a lack of consensus on issues such as website navigation and drop down filters. Participants that rated navigation issues as important indicated that are confusing and difficult to interpret for novice users. However, there was consensus on the standardization of vocabulary, in which case, it was reported that users have a difficult time interpreting data from one WDQS system to the other. As an example, one system may use 'ethnicity' and 'race' as one variable, while another system may use them as two separate variables. This item was reported by the participants but was not found in the literature and should be further explored.

There was disagreement amongst the participants on the needs for health desk support. Some participants reported that a lack of help desk support was a major barrier for end users and

their system usability. They indicated it was necessary to have adequate phone support for users who are ‘stuck’ or cannot interpret the queries. Others disagreed and indicated that systems should be self-functioning and should not require helpdesk support.

Conclusion

Past research highly emphasized technology centered problems, such as usability, bandwidth, and slow computer processors, as barriers to implementation. However, information technology has substantially improved over the last decade, making technology centric issues less of a problem. This research confirmed that we can build robust systems that are easy to use, provided state agencies have the financial resources, staffing, and support. While they can build robust systems, data are necessary. Organizations must share their data and overcome political barriers. Results of this study add to past research and demonstrate that barriers to implementation have shifted to a person centric problem. There are a lack of people resources and funding available to implement systems. Both aspects involve humans and not the technology itself. If we do not place greater investment into health data systems, we will be lacking access to data that may help combat health disparities at the local level. This may lead to at-risk groups being left untreated. Greater emphasis needs to be placed on how we can improve resource allocation with WDQS programs that will ultimately lead to public health impact.

Chapter 4:

Manuscript #3: Barriers State Agencies Face in the Dissemination of their Health Data

Target Journal: *American Journal of Public Health*

ABSTRACT

Objective: States agencies face many challenges in their dissemination of local health data. The purpose of this study is to identify the most significant challenges they face from the perspective of Behavioral Risk Factor Surveillance System (BRFSS) coordinators. We also seek to find an association between perceived system aspects, challenges faced, contextual factors, and overall satisfaction with state level health data systems.

Methods: We surveyed Behavioral Risk Surveillance System (BRFSS) coordinators from 45 states. We surveyed participants about contextual factors and asked them to rate system aspects and challenges they face with WDQS on a Likert scale. We used a t-test to compare means on participant ratings for states with and without WDQS. We created a satisfaction summary scale to understand the relationship between system status and overall satisfaction. Qualitative feedback was collected at the conclusion of the survey.

Results: Overall, BRFSS coordinators from states that have implemented Web Based Data Query Systems (WDQS) perceived their systems more favorably than states without WDQS. They reported that the cost of system development, staffing, and interpreting their data in a meaningful way are the most significant challenges they face. System satisfaction is highly dependent on the type of system the state has employed.

Conclusion: Securing adequate financial and staffing resources are vital in the dissemination of local level health data. Technology and usability problems are less of a problem, provided there are adequate resources.

BACKGROUND

Virtually everyone in public health acknowledges that the progress in information technology has paved the way for exciting opportunities to disseminate local health data more efficiently (Thacker, 2012). Despite the advances in information technology, state and local health agencies have not kept pace. The need to have ready access to health data at the local level has been recognized as crucial to fulfill the broad responsibilities for monitoring the health status of populations and for planning interventions (Simon, 2001). Currently, 29 out of 50 states make health data available interactively using Web Based Data Query Systems (WDQS). WDQSs are advantageous as they make local health data more easily and readily available. They present data using a dynamic interface, which allows users to interact with the system, and customize data queries through choosing data sets, variables, measures, and the format for presenting query results (Friedman, 2006). Despite the advantages of WDQSs, implementation has not reached its potential. Of the remaining 21 states without WDQS, 19 still employ a static interface. A static interface limits user choice to pre-calculated statistics and allows no choice of parameters for a query (Friedman, 2006). Static interfaces are updated on time intervals (quarterly, annual, etc.) and make reports available for download in formats such as PDF or Microsoft Word. There are advantages, as static interfaces cost significantly less to develop and require fewer staff to maintain. However, they do not offer the level of depth in data and have limited functionality in comparison to WDQS. For instance, users can query off multiple data sources with WDQS. This cannot be achieved through a static interface, as users are unable to submit queries and are dependent on only the reports that are made available.

In the mid 1990's, several states began sharing their data through a static interface. As technologies improved, some states upgraded to WDQS. Between 1999 and 2005, several states

had implemented and developed WDQS. There was an explosive growth period, as many states were interested in making health data more easily available. Within that time period, legacy WDQS such as IBIS-PH/Utah and EPISQMS/Pennsylvania were first implemented. IBIS and EPISQMS were interactive and linked data from multiple sources across their states. Since 2005, implementation of new systems has been limited, as only a few new states have developed WDQS. Along with limited implementation, existing systems are not being frequently upgraded or enhanced. A system should be upgraded with improved functionality at least once a year in order to keep up with technological innovation (Chapin, 2001). To demonstrate, the widely adopted IBIS-PH developed by the state of Utah was first released in 1999. The next major upgrade, version IBIS-PH 2.0, was not released until 2010. The next version is not expected to be released until 2016. The slow progress in upgrading can impact multiple states, as in the case of IBIS-PH, which has been adopted by four other states. In the initial phases of WDQS development in the late 1990s, there was a strong push towards collaboration amongst states (Simon, 2001). This was achieved within the first decade as states such as Utah used open source technology to develop their WDQS and encouraged system adoption. Open source software makes its source code made available freely, which can be distributed to anyone and for any purpose (St. Laurent, 2008). While open source systems make software freely available, they are expensive to maintain; 90% of the costs of a typical system arise in the maintenance phase (Brooks, 1995).

Adequate funding and resources are vital to implement WDQS (Simon, 2001). These costs include software development, hardware costs, and staffing. The overall cost for implementation on average is between \$250,000 and \$500,000 (Friedman, 2006). A significant portion of this cost is allocated for software development. The wages for experienced software

developers are high, and demand continues to increase. The recent demand for experienced software developers has led to higher wages that most state governments cannot afford to pay (Gunasekaran, 2001). The challenges are expected to continue, as the projected demand for software developers is expected to rise 22% until 2020 (USBLS, 2011).

It was found in our prior literature review that the primary challenges to dissemination of state level health data include cost, data collection, IT infrastructure, and usability. We investigate these challenges from the perspective of BRFSS coordinators from each state. We invited BRFSS coordinators because of their significance in releasing health data and their role in their state's technology decisions (Ver Ploeg, 2006). The overall goal of this study is to identify the most challenging aspects state agencies face in their dissemination and to identify primary factors that influence their system satisfaction.

METHODS

This study is the third phase of a mixed-method study evaluating barriers state agencies face with the dissemination of their local health data. The first phase included a systematic review of the literature available on state level health data systems. The second phase consisted of a three round Delphi consensus study that surveyed a panel of experts that included researchers and knowledgeable public health professionals. Participants evaluated the importance of several items identified in our literature review, which included the four topics of cost, data collection, IT infrastructure, and usability challenges. The goal was to identify items of importance, based on expert consensus. Items of importance that met consensus were used to develop the survey instrument for the third phase of the study.

In the third phase, we designed and administered a web-based questionnaire, which was administered to BRFSS coordinators. Each state is represented by one BRFSS coordinator.

The BRFSS, established in 1984, is a health-related telephone survey that collects state data in all 50 states about U.S. residents regarding their health-related risk behaviours, chronic health conditions, and use of preventive services. The first component of the questionnaire surveyed on contextual factors, such as software used and path chosen in development. The second component of the questionnaire asked participants to rate system aspects. The third component of the questionnaire asked participants to rate the challenges state agencies face with their health data system. The survey was designed using *Qualtrics* software.

Development of the Survey

The survey instrument was designed using three question types that include: contextual factors, system aspects ratings, and rating the challenges state agencies face. Contextual factors were designed using a categorical multiple-choice based format. Contextual factors surveyed on questions such as data sources used, types of software used, path used in system development, and headcount. Responses of these questions were instrumental in conditional branching logic, or ‘skip logic,’ within the survey. In this situation, the response indicated that a state is using an interactive query system; questions relevant to those with an interactive query would follow.

For the system aspects ratings, participants were presented a list of items to rate. They rated how they perceived each of the system aspects presented. They rated each item on a 5-point Likert scale of 1 to 5 (Poor – 1; Fair – 2; Good– 3; Very Good – 4; Excellent - 5). System aspect questions were based on items such as website performance, data quality, and other aspects with the development and dissemination of health data. Items in the ‘challenges faced’ component were also derived from the phase 1 Delphi study. For the ‘challenges faced’ component, participants rated the level of challenges they face on a 4-point Likert scale

of 1 to 4 (Not at all challenging – 1; Not very challenging– 2; Somewhat challenging– 3; Very Challenging – 4). At the conclusion of the survey, participants were given an opportunity to provide qualitative feedback and list additional items of importance not presented in the questionnaire.

Study Sample and Recruitment

We aimed to recruit Behavioral Risk Factor Surveillance System (BRFSS) coordinators from all fifty U.S. states and the District of Columbia to participate in the study (n=51). In January 2015, each prospective participant was sent a letter in the U.S. Mail inviting them to participate. We found the names and contact information of all BRFSS coordinators from the Center of Disease Control and Prevention (CDC) website. The letter provided details of the study, indicated that their participation was confidential as results would not be disclosed to anyone except study staff, specified that their participation was voluntary, and stated that they could withdraw at any point of the survey. Each participant was sent a follow-up email with a link to the survey with a secure username and password within 1-2 weeks upon receipt of the letter.

Ethical consideration

The study participants were informed that their participation in the study was entirely voluntary and they implicitly consented to participate by completing the questionnaire. They were also informed that their results were confidential as well as their names. Ethics approval/IRB determination for the study protocol was received from the Human Subjects Protection Office at the University of Connecticut Health Center.

Measurements

Dependent variables

- *Challenges state agencies face* were measured on a 4-point Likert scale from 1 (not at all challenging) to 4 (very challenging). Respondents were presented nine items and asked to independently rate challenges based on their experience. Higher values were indicative of an item being a greater challenge. They rated the following challenges:
 - Cost of system development
 - Cost of hardware/software
 - Cost of vendors/consultants
 - Lack of internal IT staff
 - Helpdesk support
 - Lack of trained staff that understand data
 - Receiving data in a timely manner
 - Privacy
 - Political Opposition

- *System aspects* were measured on a 5-point Likert scale from 1 (Poor) to Excellent 5 (Excellent). Respondents were presented 12 items as asked to independently rate each system aspect of their health data system that include:
 - User friendliness
 - Website performance
 - Standardization of Vocabulary
 - Breadth of data
 - End User Satisfaction
 - Quality of data
 - Ability to link to multiple data sources
 - Availability of race, gender, and other social determinants
 - Accessibility to researchers
 - Accessibility to non-researchers
 - Timeliness of support requests
 - Knowledge of how consumers consume their information

- *System Satisfaction* is a scale constructed from a group of eight system aspects that were rated by participants. Satisfaction is considered to be one of the most widely

used measures of IT implementation success (Daramawan, 2005). We measured system satisfaction to assess how participants perceive their systems. We selected items from the survey that were specific to their *system attributes*. For example, items such as timeliness of support and knowledge of how consumers consume their information were excluded. We used the following items from the survey that include: user friendliness, breadth of data available, quality of data available, website performance, system ability to link to multiple data sources, system satisfaction, accessibility to researchers, and accessibility to non-researchers. The index was calculated based on the mean response for these items, of which the maximum score is 5 and the minimum is 1. Cronbach's alpha was used to test reliability. The closer the Cronbach's alpha coefficient is to 1.0, the greater the internal consistency of the items in the scale. Internal consistency is important as it measures whether several items that propose to measure the same general construct produce similar scores. A Cronbach's alpha of .8 or higher is recommended (Gliem, 2003).

Independent variables

WDQS status is the primary independent variable of interest. It is known that WDQSs offer greater functionality and are more useful for researchers and other key stakeholders. However, we use this variable to assess differences in system aspects, satisfaction and participant ratings. States that reported they employ a WDQS are coded 1 and states without a WDQS are coded 0.

Data Analysis Plan

Data analysis was conducted in three steps. First, we presented descriptive statistics of contextual factors (Table 8). In the second phase, we analysed the ratings of each of each

item in the survey. Third, we analysed differences in participant ratings between states with WDQS and without WDQS.

In the first step, *Contextual Factors* were assessed analysing the distribution of the reported responses (Table 8). We use frequency tables and examined the distribution of reported items. In the second phase, *Participant Ratings* of their satisfaction with various aspects of their health data system (Table 9), and the challenges they faced in implementing and maintaining this system (Table 10) were analysed. In the third phase, we assessed differences in participant ratings of satisfaction and challenges among those with WDQS and without WDQS. We compared the means on these items using an independent samples t-test to determine whether a significant mean difference existed between states with WDQS and without WDQS. We individually conducted a t-test for each dependent variable of interest to determine statistical significance. Finally, we also calculated a summary scale as an indicator of overall system satisfaction by summing participant ratings of:

- User friendliness
- Breadth of data available
- Quality of data available
- Website performance
- System ability to link to multiple data source
- End user satisfaction
- Accessibility to researchers
- Accessibility to non-researchers

This was calculated by summing responses to each item and dividing it by the number of items in the index (n=8). We compared the mean on this summary scale among states with and without WDQS using an independent samples t-test. In all analyses, we defined items with a $p < .05$ as statistically significant. We used SPSS version 21.0 software (IBM, Armonk, NY) to perform the analyses.

RESULTS

We received completed surveys from 45 states (44 through a web-based questionnaire and 1 telephone response), with an overall response rate of 89%. Details of the characteristics by state are shown in Table 8. Participants rated system aspects, which are shown in Figure 2. Participants also rated the challenges they face; these are shown in Figure 3.

Step 1: Contextual Factors

As can be seen in Table 8, 43 of the 45 (95%) states surveyed share their health data over the internet in some form. Twenty-eight states share their data through an interactive query system/WDQS, representing (62%) of the states responding to this question (Table 8). The remaining 15 states (38%) share their data using a static interface.

Participants reported that the most common data that states make available include BRFSS (73%), birth (68%), and death (66%). Only a few states make lead screening (11%) and hospitalization (9%) data available.

We surveyed participants on who they perceived is their largest consumer of data; 48% of the participants reported the largest consumers are state agency personnel, followed by the general public (18%), researchers (13%), and policymakers (8%). Evaluation of end users is limited as 44% of participants confirmed that they do *not* evaluate end users. End user evaluation is a process that evaluated system satisfaction from an end-user perspective (Vosbergen, 2012).

Decision making in software development has varied from state to state. Participants reported the path(s) chosen in development. More than one response was allowed, as states may have more than one path to development. For example, a state may develop their own system with one data source and use a commercial off-the-shelf product for another data source. The most common path chosen in development were states that developed their system ‘in-

house' (n=13), followed by hiring an outside contractor/vendor (n=8), adopted from another state (n=6), and a commercial off-the-shelf product (n=4).

Table 8 Profile of health data characteristics at the state level

		N (Number of states)	Percent
Present health data over the internet (n=45)			
	Yes	43	(95)
	No	2	(5)
Present health data using an interactive web based query system (n=43)			
	Yes	28	(62)
	No	15	(38)
Types of data that are publically available (n=43)			
	BRFSS	33	(73)
	Birth	31	(68)
	Death	30	(66)
	Lead Screening	5	(11)
	Hospitalization data	4	(9)
* Path taken in software development development (n=28)			
	Developed in-house	13	
	Developed by outside vendor or contractor	8	
	Adopted from another state	6	
	Developed using off the shelf commercial software	4	
	Other	2	
Largest consumer of data (n=40)			
	State Agency Personnel	19	(48)
	General Public	7	(18)
	Researchers	5	(13)
	Policymakers	3	(8)
	Other	6	(15)

* For the path chosen in development, more than one response was allowed. In some cases, states chose multiple paths to development.

Step 2: System Aspects

Overall, participants rated (Table 9) their system's ability to link data to multiple data sources and standardization of vocabulary across datasets the lowest. Standardization of vocabulary across datasets refers to the ability to connect one attribute to another from one system to another. For example, in one state the metric is labeled 'obese or overweight,' while in another state it is labeled as 'obese.' Quality of data and system security were rated the highest. Data are of high quality if they are fit for their intended uses in operations, decision making and planning (Juran, 1999). System security is protection of information systems against

unauthorized access to or modification of information. In our analysis, we categorized survey items into three distinct groups that include: *website usability*, *accessibility and support*, and *data quality*.

Website usability

User friendliness (mean=3.10; 95% CI = 2.80, 3.41) and end user satisfaction (mean=3.14; 95% CI = 2.85, 3.43) were lowest. A user friendly website is one that is easy to use, reliable, has interactive features, and is customizable. End user satisfaction is the extent to which users believe the system available to them meets their informational requirements. Website performance (mean=3.34; 95% CI = 3.02, 3.66) rated slightly higher. Factors that influence website performance include efficiency, trust, system availability, fulfillment, and responsiveness (Park, 2007). These elements are common drivers of website usability. System security was rated highest (mean=3.82; 95% CI = 3.40, 4.24). This is an indication that there is trust in security measures and unauthorized access with WDQS.

Accessibility and Support

Participants gave their lowest ratings to timeliness of response to support requests (mean=3.53; 95% CI = 3.20, 3.87). Common problems that require technical support include website navigation issues, data retrieval problems, or how to interpret data output. In these cases, a support request may be raised by the end user through e-mail, telephone, or through a ticketing system. There was minimal difference when differentiating accessibility between researchers and non-researchers. Participants rated accessibility to researchers (mean=3.67; 95% CI = 3.35, 3.99) and accessibility to non-researchers (mean=3.58; 95% CI = 3.26, 3.90) slightly higher.

Data Quality

Participants rated the ability to link across multiple data sources (mean=2.80; 95% CI = 2.20, 3.41) and standardization of vocabulary (mean=3.03; 95% CI = 2.65, 3.42) lowest. They rated breadth of data (mean=3.33; 95% CI = 2.99, 3.68) and quality of data (mean=3.80; 95% CI = 3.60, 4.19) higher.

Overall Satisfaction Scale

The satisfaction scale was comprised of a summary scale that consisted of eight system aspects. System satisfaction was rated higher for those with WDQS (mean=3.57; 95% CI = 3.07, 3.87) in comparison for states without WDQS (mean=3.03; 95% CI = 2.72, 3.34). However, this comparison was statistically significant only at the .10 alpha level.

Table 9 - Mean Participant Ratings of System Aspects on a Five-Point Likert Scale

Grou Item	Overall Mean Score (n=43)	With WDQS (n=28), (mean, 95% CI)	Without WDQS (n=15), (mean, 95% CI)	P value
<i>Website usability</i>				
User Friendliness	3.1	3.15 (2.73, 3.56)	3.00 (2.69, 3.32)	.636
Website performance	3.34	3.54 (3.15, 3.93)	2.92 (2.47, 3.36)	.064
* Standardization of vocabulary	3.03	3.37 (2.89, 3.84)	2.40 (2.08, 2.71)	.012
End User satisfaction	3.14	3.32 (2.94, 3.69)	2.85 (2.47, 3.22)	.111
<i>Data Quality</i>				
Quality of data	3.9	4.00 (3.62, 4.37)	3.69 (3.28, 4.09)	.327
Breadth of data	3.33	3.42 (2.96, 3.87)	3.15 (2.71, 3.58)	.462
* Ability to link to multiple data sources	2.81	3.28 (2.57, 3.98)	1.75 (1.26, 2.24)	.013
Availability of race, gender, and other social determinants	3.91	4.05 (3.63, 4.46)	3.64 (3.24, 4.03)	.114
<i>Accessibility and Support</i>				
Accessibility to researchers	3.67	3.92 (3.56, 4.27)	3.15 (2.66, 3.63)	.059
Accessibility to non-researchers	3.58	3.76 (3.39, 4.12)	3.23 (2.69, 3.76)	.114
Timeliness of support requests	3.53	3.68 (3.28, 4.07)	3.27 (2.73, 3.80)	.234
Knowledge of how consumers consume information	2.34	2.33 (1.89, 2.76)	2.36 (1.96, 2.75)	.931
<i>System Satisfaction</i>				
Satisfaction Summary Scale	3.39	3.57 (3.27, 3.87)	3.03 (2.72, 3.34)	.056

Note: Overall mean score represents the full sample. An independent sample t-test was used to compare the means between states with WDQS and without WDQS. Statistical significant is noted by *, where $p < .05$.

Step 2: Challenges Faced

Overall, participants rated (Table 10) the costs of system development, consultants/vendors and hardware/software as the most challenging. The cost of system

development includes the design, development, testing, and implementation of their system. The cost of consultants/vendors includes outsourcing software development to third party vendors. The cost of hardware/software includes the cost of servers, mainframes, software/software licensing, and other applicable costs. Participants reported that the lack of political support and issues with data privacy as less of a challenge. Political support refers to the level of commitment from relevant political stakeholders for data sharing (Savel, 2012).

Staffing and Support

Table 10 includes the mean ratings from the ‘challenges faced’ component. Participants reported a lack of internal IT staff (mean=2.97; 95% CI = 2.55, 3.39). Examples of internal IT staff include the roles of business analyst, systems analyst, technical project manager, and project management staff. Participants reported concerns that there was a lack of trained staff who understood health data (mean=2.62; 95% CI = 2.26, 2.98). Examples of trained staff that understand health data include epidemiologists, public health professionals, or researchers that can analyze health data in a meaningful way. Their role is to communicate a set of specifications for software developers to interpret and understand. IT staff are responsible for establishing the set of requirements, managing the smaller details that are involved in IT projects, and establishing timelines for deliverables.

Participants rated the lack of IT support (mean=2.60; 95% CI = 2.09, 3.11) lower than internal staffing and their understanding of health data. IT support is available for technical issues such as website navigation, system problems, and data retrieval. It was reported that receiving data in a timely manner (mean=2.41; 95% CI = 2.11, 2.69) was less of a challenge. However, states still need the staffing resource to make meaning of the data, despite the timeliness of receiving data.

Cost Barriers

The cost of system development (mean=3.23; 95% CI = 2.92, 3.24) reported as the most challenging. The process of system development includes the design, development and implementation of a software application. Participants reported that the cost of hiring outside consultants or vendors (mean=3.03; 95% CI = 2.92, 3.24) is challenging. Participants reported that the cost of hardware/software (mean=2.89; 95% CI = 2.57, 3.22) as a slightly lower challenge. The cost of computers, hardware and servers is much less of a barrier than it once was (Luck, 2006). The price for servers and mainframes are a fraction today compared to the 1990s.

Other challenges

Participants reported that concerns over privacy (mean=2.55; 95% CI = 2.24, 2.87) as less of a challenge in comparison to other items. Over the last decade, there has been greater awareness on privacy of sensitive health data. Organizations are taking measures to prevent inadvertent disclosure of data at the individual level. There have been vast technological improvements and techniques to suppress data with small cell counts. Participants reported that political barriers are less challenging (mean=1.77; 95% CI = 1.34, 2.20). A political barrier may be defined as structural barrier between or within organizations.

Table 10 - Participant Ratings of Challenges State Agencies Face on a Four-Point Likert Scale

Group	Item	Overall Mean Score (n=43)	With WDQS (n=28), (mean, 95% CI)	Without WDQS (n=15), (mean, 95% CI)	P value
<i>Cost</i>					
	Cost of system development	3.33	3.05 (2.68, 3.41)	3.6 (3.28, 3.91)	.063
	* Cost of hardware/software	2.89	2.63 (2.25, 3.00)	3.44 (3.09, 3.78)	.013
	Cost of vendors/consultants	3.05	2.93 (2.44, 3.41)	3.25 (2.63, 3.86)	.433
<i>Support and Staffing</i>					
	Lack of internal IT staff	2.97	2.90 (2.37, 3.42)	3.10 (2.48, 3.71)	.663
	Helpdesk support	2.60	2.41 (1.80, 3.02)	3.00 (2.25, 3.74)	.272
	Lack of trained staff that understand data	2.62	2.42 (1.98, 2.85)	3.00 (2.49, 3.50)	.117
	Receiving data in a timely manner	2.40	2.36 (2.07, 2.76)	2.42 (1.98, 2.85)	.848
<i>Other Challenges</i>					
	Privacy	2.55	2.37 (2.06, 2.67)	2.9 (2.28, 3.51)	.101
	Political Opposition	1.77	1.6 (1.10, 2.11)	2.14 (1.47, 2.80)	.231

Note: Overall mean score represents the full sample. An independent sample t-test was used to compare the means between states with WDQS and without WDQS. Statistical significant is noted by *, where $p < .05$.

Step 3: WDQS vs Non-WDQS

We analyzed the mean results of the reported challenges between states *with* interactive query systems (WDQS) with states that are *without* WDQS. Overall, states *with* WDQS rated systems more favorably. We compared the means for reported responses and found that states *with* WDQS rated systems aspects higher in 11 out of 11 (100%) survey items. Overall, states *with* WDQS found items less challenging than states *without* WDQS. States *without* WDQS rated 8 of out 9 (88%) of the items presented as a greater challenge. The cost of hardware/software ($p = .013$), standardization of vocabulary ($p = .012$) and ability to link to multiple data sources ($p = .013$) were reported as statistically significant.

DISCUSSION

To our knowledge, this was the first study to investigate barriers state agencies face with the dissemination of their local health data in a nationally represented sample. This study found that state agencies are hindered by the high cost of development, hardware/software, lack of adequate staffing, and lack of trained staff that understand their data. We found that most states develop their own systems, while few adopt systems from other states. In addition, only a few states evaluate the system needs of end users or conduct usability testing on their systems. Overall, it was advantageous for a state to have employed a WDQS. BRFSS coordinators in states with a WDQS were significantly more satisfied with their systems than those without WDQS.

Contextual Factors

Contextual factors provided us information on items such as development, technologies used, personnel, and other key attributes at the state level. Most states surveyed indicated that they developed their own system in-house or hired consultants/vendors. Only few states collaborate with other states. Collaboration can help reduce the cost burden for states with limited resources. A limited number of states take advantage of open source technologies. Only four respondents indicated they are using open source technology. The free and open source approach is a promising alternative for delivering cost-effective information technology (Shaame, 2003). There has been considerable growth in the last 5 to 10 years for open source technology in other fields such as healthcare delivery and medical research (Raghupatchi, 2014). Despite the growth of open source technology in other sectors, state agencies have not followed.

In addition, evaluation of systems and end users is limited. Participants reported that only a few states evaluate end users on how they perceive their systems. Since a large amount of

money is spent on information technology systems, it is important to evaluate them (Seddon, 1999). Without evaluating users, it is unknown if systems are being used by their intended users (Friedman, 2006). Participants also reported that only a few states are conducting usability testing. Usability testing informs us if systems are providing accurate data, are easy to use, and are useful for the end user.

Participants confirmed that the largest user group for WDQS are internal state agency personnel. They use this data for surveillance, reporting, and monitoring the health of their community. This is a problem, as health data systems such as WDQS have potential to be useful for a variety of stakeholders. Since large resources are being spent on these systems, a broad user base can justify expenditures (Wessels, 2003).

Participant Ratings of System Aspects and Challenges

Taken as a whole, our findings suggest that BRFSS coordinators from states with WDQS are more satisfied with their systems. Those with WDQS rated all system aspects higher and found challenges such as cost of hardware/software and system development less challenging. This is despite the higher cost and staffing burden of implementing and maintaining WDQS, which is an indication of favorable assessment of the cost-benefit ratio of WDQS relative to static systems.

A lack of *internal* IT staff was reported as a major challenge. Internal IT staff includes the role of business analyst, systems analyst, project manager, and other relevant staff. Their roles are instrumental in the maintenance and development of new systems. They ensure the system is developed to user requirement and understand the business needs (Iyamu, 2011). A lack of internal IT staff can lead to missing requirements and introduce large risks, which contribute to failure or delayed projects. As an alternative, several state agencies hire short term

vendors/consultants when they do not have the requisite IT staff or software developers in-house. This study found that the cost of hiring vendors/consultants is a major challenge, as their hourly billable rates are known to be high. Although vendors/consultants are able to complete short term projects, there is high turnover. IT staff turnover escalates the cost of software development by 40-60%, and contributes to gaps in knowledge transfer (Chang, 2003). It can take several months for new consultants/vendors to become familiar with the system and its processes. This is costly and inefficient for state agencies.

This study found some items that contradict the literature we reviewed and the Delphi study. For example, it was found in the literature review and Delphi study that data sharing and political barriers are major barriers for state agencies (Vest, 2013). This study found the contrary, as data sharing and political barriers were less challenging. Participants confirmed qualitatively in the feedback section that they have access to high quality data but lack the capacity and trained staff to understand, maintain, and interpret the data. The lack of political barriers indicates that state agencies are willing to share their data but face challenges to make data useful and meaningful. Issues of data privacy and system security were also reported to be less challenging. Because of significant high profile breaches of data systems in both public and private organizations, data privacy concerns have been on the forefront of media and academic literature. Privacy concerns have had a significant negative impact on important public health research, as data stewards have been reluctant to release their data (Wartenberg, 2010). This study found the contrary, as privacy concerns are not rated as highly challenging.

Satisfaction with system security was amongst the most highly rated aspects of state data systems. This is an indication that there is trust in their system's infrastructure and there is less concern of access from unauthorized users. Along with trust, data quality was highly rated. In

the context of health data, data quality can be defined as “features and characteristics of a data set that bear on its ability to satisfy the needs that result from the intended use of the data.”

(Arts, 2002). While data quality was rated favorably, the ability to link to multiple data sources was rates less favorably. Systems with a single data source are missing data on social determinants of disease and are lacking a comprehensive description of an affected population (Narayan, 2006). Incompatibility of IT systems and a lack of skilled staff contribute to this problem (Vest, 2013).

Implications

The lack of availability of local health data prevents local health departments from targeting populations at risk and evaluating interventions (Brownson, 2011). Local health is easily available through WDQS and can lead to immediate action and priority setting. As an illustration, 2001 BFFSS data in Texas found that populations near the US-Mexico border had a three times higher rate of diabetes than the general population (Brownson, 2011). Public health interventions led to greater consumption of fruits and vegetables in the subsequent years. The consumption of fruits and vegetables and reducing unhealthy foods reduces the risk of obesity, Type 2 diabetes, and other non-communicable diseases (Knowler, 2002). Rates of diabetes and obesity have continued to rise in the US over the last decades, as diabetes is now amongst the leading causes of death (Knowler, 2002).

There is also a financial burden as preventable diseases such as diabetes, and obesity leads to billions in Medicare and Medicaid expenditures (Finklestein, 2004). The United States spends more in health care expenditures than any other nation. Despite the high expenditures, rates of non-communicable diseases are rising and life-expectancy is ranked 34th in the world (Murray, 2013). Greater effort needs to be made monitoring the health of our population. This

can be achieved through health surveillance efforts and using data to track and monitor disease. It is advantageous for state and local health agencies to invest resources in WDQS. In the long term, investing resources in WDQS may help save billions in healthcare expenditures costs, help save lives, and lead to a healthier population.

Limitations

The results may not be generalizable beyond the perspective of only BRFSS coordinators in each state. Perspectives from technology professionals are not captured in this study. We selected BRFSS coordinators because of uniformity, as each state is equally represented with one per state. It would be difficult to achieve this level of uniformity if we had interviewed technology professionals. Each state's organizational structure in information technology varies from state to state.

Second, our study may reflect bias, as BRFSS coordinators with WDQS may rate their systems higher. Those with BRFSS may have a positive bias towards their system and perceive their system as favorable. In addition, sample size for states without WDQS was smaller (n=15) than those with WDQS (n=28). Those without WDQS had a lower response rate in comparison to those with WDQS. Of the 21 states without WDQS, 15 (71%) had responded to the survey. Those with WDQS had a higher response rate, as 28 out of 29 states (96%) had completed the survey.

This study identified state agency personnel as the primary user group. We did not investigate why a broad user base is lacking. Further research should be conducted to understand why use is limited outside state agency personnel. WDQS can be useful to multiple stakeholder groups, such as researchers, policymakers, and community organizations, which help influence policy and make decisions. Little is known if there is a lack of awareness that these systems are

useful and are available. Future research can help determine if systems are too complex for users without a public health background to understand. To demonstrate, some systems use epidemiological terms, such as adjusted odds-ratios, or by default display 95% confidence interval error bars. Little is known as to which user group the target audience is for WDQS.

Conclusion

Forty-five BRFSS coordinators shared their perspectives on factors that contributed to challenges they face with their data dissemination. Overall, they reported strong system satisfaction and rate their systems favorably. However, they reported that funding for WDQS is lacking, which have impeded, slowed progress or halted their efforts with their dissemination. In the current environment, it may be challenging for state agencies to receive funding for WDQS development as there is limited knowledge of their utility. This study has provided us with evidence that the technology and usability of systems is less of a problem. This study also found that data sharing and political barriers are limited, as organizations are more likely to share their data with one another. Our results indicated that priorities need to be shifted to the end user perspective and the importance of these systems. Directions for future research include further evaluating potential uses and identifying the data needs of end users. Providing evidence indicating their resourcefulness and their potential impact on public health will likely help justify expenditures. It is important to highlight the utility and importance of these systems to policymakers that control funding.

Chapter 5

Summary and Conclusions

In its landmark report of 1988, the IOM recommended that every public health agency systematically collect, assemble, analyze, and *make information available* on the health of the community (Thacker, 1998). This mission includes statistics on health status, community health needs and epidemiological information available at the community level. These activities are known as *public health assessment*, which is one of three core functions of public health. Our study has demonstrated that the availability of health data at the state and community level falls short of goals recommended by the IOM. Data are often incomplete, missing, or lack substantive information to address the needs at the community level, making it difficult to address emerging health problems. Making timely, high quality health data available to researchers, practitioners and policymakers is critical to improving the overall health of the population as well as meeting the goals for initiatives such as *Healthy People 2020*. Information technology has vastly improved since the IOM released this report in 1988; however, state agencies have not been able to keep pace. While technological innovation has brought change to the landscape, state agencies struggle with the basics, such as ensuring that they provide reliable, high quality data.

There have been efforts to improve health surveillance at the national level. For example, in the 1990s, the CDC developed the BRFSS, a model for state population health surveys. The National Cancer Institute surveillance, epidemiology, and end results (SEER) system was developed in the 1990s. SEER operates in eleven population-based cancer registries and three supplemental registries covering 14 percent of the U.S. population (National Cancer Institute, 2003). National and state health surveys are popular and provide critical information towards achieving state and national health objectives but rarely serve local needs (Healthy People,

2000). The goals of making data available at the local and community level are not being fully met. Our comprehensive study informed us in-depth of problems that exist as well as solutions to improve the landscape of health surveillance at the local level. In order to satisfy the core public health function of assessment, we must not only track health data but make it *available* to the community.

Process of this dissertation: The Systematic Review of Literature

We reviewed the available literature (Manuscript #1) to help formulate a strong baseline understanding of public health data systems. We obtained knowledge of the challenges state agencies faced such as website usability, and poorly designed systems. We were also informed about political and data sharing barriers. Our literature review also revealed that WDQS are not being evaluated by their end user base, and little is known of their utilization. We also found that data are often missing and incomplete and systems are difficult to use (Luck, 2006). Based on findings from our review, we categorized the study into four major challenges, which included: cost, data collection, usability, and IT infrastructure. The literature review was useful as it provided us with a fundamental understanding of problems that exist but yet required further research. This information was then used to design our Delphi study questionnaire.

The Delphi Study

The results from the Delphi study provided us with knowledge of the issues and challenges confronting public health data systems, based on expert consensus. Some topics that were deemed by our experts as important were consistent with what we observed in our review of the literature, such as cost of system development, data sharing challenges, and lack of staffing/support, which were all found to be in consensus. However, some issues cited in the literature, such as speed/bandwidth, system reliability, privacy, cost of hardware/software, and a

system's ability to handle large datasets, were not in consensus. This may be indicative of certain problems having been resolved over the past decade; as technology has vastly technical problems pose much less of a challenge.

The qualitative/participant generated section of the study identified problems that were not found in the literature. In particular, participants reported that standardization of vocabulary across datasets was challenging and that there was little understanding of how users consume data, both of which were not found in the literature review. This component of our study was key for the development of the BRFSS coordinator survey, as participant generated items would have not been found had we relied solely on survey questions derived from the literature review.

Table 11: Relationship between Delphi Study (Manuscript #2) and BRFSS coordinator survey (Manuscript #3)

Topic	Manuscript #2: Delphi study (Items in Consensus)	Manuscript #3: BRFSS survey
Cost		
	Cost to have adequate state agency (public health) staff	Approximately how many full-time staff currently work on maintenance and support for your health data system?
	Cost of system design/software development	Rate the challenge on the cost of system development
	Cost to have adequate staff/headcount for IT staff (internal)	How much of a challenge has it been to have adequate IT Internal staff
	Cost of IT technical support for state agency staff	Please rate the challenge for the following item: Lack of IT Helpdesk support for end users
	Cost of technical support to end users	Rate the following: Timeliness of response from support
	Cost of servers/hosting Applications	How much of a challenge has the following been? The cost of hardware/software
	Cost of data storage	Item met consensus, however it was rated as 'not important'
Data Collection		
	Challenges in acquiring data that are useful and meaningful	Please rate the following: Our understanding of how the consumer uses the information
	Challenges in acquiring data that have been requested by relevant stakeholders/end users	Please rate the following challenge: Acquiring data from outside sources
	Challenges in acquiring data from multiple data sources across the state	Please rate the following: Your system's ability to link to multiple data sources
	Challenges in working with private hospitals and clinics to release data	Which of the following data sources do you currently make available to users through your health data system
	Collecting data in a timely manner	Please rate the following: Timeliness of data from outside sources
	Challenges in working with public hospitals and clinics to release data	Which of the following data sources do you currently make available to users through your health data system
IT Infrastructure		
	Collaboration with software developers and IT staff ensuring that requested system design needs are met	Please rate the following: IT staffing and support within your organization
	Challenges in decision making on technology options (open source, commercial, or adopt from another state)	Which of the following path did you choose in your development?
Usability		
	Data are meaningful and is useful for the end user	Please rate the following: End user satisfaction
	Data are missing or incomplete for end user	Please rate the following challenge: Data are missing or incomplete for the end user
	Quality of user data output (Excel, csv, pdf, html etc.)	Please rate the following: Quality of data
	Drill Downs/Data Filers are difficult to understand	Please rate the following: User Friendliness
	Website freezes up	Please rate the following: Website performance
	Data do not go far back enough in time	Please rate the following: Receiving data in a timely manner
Qualitative Items (Participant Generated)		
	Evaluation of end users	Does your state currently perform a comprehensive evaluation of who utilizes its system
	Standardization of vocabulary	Please rate the following challenge: Standardization of vocabularies across datasets
	Providing context in a way which makes a 'story' of the data	How much of a challenge has it been to have knowledgeable staff that understand your data?
	Hidden costs associated with development	Please rate the following challenge: Cost of outside vendors/consultants
	A greater understand of how the consumer consumes the information	Please rate the following challenge: Our understanding of how the consumer uses the information
	Need for 'user centric' design	Please rate the following: Website performance
	Helpdesk support for end users	Please rate the following: Timeliness of support for end users
	Data from the private sector	Please rate the following challenge: Acquiring data from outside sources
	Using existing public health surveillance systems and mandated hospital discharge reporting maintained by state department of health	Please rate the following: Breadth of data available
	Rigorous validation of data and statistical algorithm	Please rate the following challenge: Data are missing or incomplete for the end user

The State Health Coordinator Study

Issues and topics deemed to be important by our expert panel from the Delphi study were used to develop questions for our State Health Coordinator study (Table 11). The survey of state health coordinators provided insight into the perspective of those who are systems owners and are in working in public health practice. BRFSS coordinators from 45 states responded to the survey, which gave us a broad-based sample. Given their role in their state's technology systems (Bloom, 2000), they were uniquely positioned to serve as the population for this study. We examined how participants rated their challenges in presenting data over the internet, as well factors that influence satisfaction with their data systems. It was evident that funding for development and staffing was the biggest challenge. Barriers such as data quality, website performance, and usability can be overcome if states have adequate funding and staff. We found that whether a state had employed a WDQS was also a strong predictor of how they rated their systems, despite the high costs associated with development and staffing.

The Delphi Study vs. the State Health Data Coordinator Study

Although one of the key objectives of the Delphi study was to develop an instrument for the State Health Data Coordinator Study, the results obtained from the two studies were also analyzed (Table 12). There were both similarities and differences in what was observed in the two studies. Both studies found that the cost of system development, staffing, IT support, understanding how end users consume data, and having data that are meaningful and useful constitute a challenge for states. There were several items found to be a problem in the Delphi study that were found to be less of a problem by BRFSS coordinators. For example, data sharing, data collection, and access to high quality data were found to be a problem in the Delphi study. The State Health Coordinator Study found data sharing to be a less of a problem and that

good data are available, but states are lacking adequate staff to maintain and interpret data. The State Health Coordinator Study also found political barriers as less of a problem, as a BRFSS coordinator did not perceive access to data as a big problem. The Delphi study and literature review found that political barriers are a big problem, and entities within a state do not share their data. However, the State Health Coordinator Study rated political barriers as the least challenging aspect from all items surveyed. While BRFSS coordinators work with state level health data systems on a regular basis and apply these systems in practice, their perspective may be different than other state agency personnel. Given their role and authority within state public health agencies, BRFSS coordinators may face fewer political challenges in accessing data and managing political opposition.

Usability problems were widely cited in the literature but found to be less of a challenge in the Delphi and State Health Coordinator Study. Information technology has improved significantly over the last two decades and is now ubiquitous in most aspects of life, both of which have made usability challenges for end users less of a problem. However, the Delphi and State Health Coordinator Study confirmed that states lack the resources to modernize and continuously update their systems. This study found that if there are adequate resources, states can develop systems that are more robust with improved functionality.

Table 12 : Comparison in response (Delphi study vs BRFSS study)

Topic			
Cost	Delphi Study	BRFSS Study	Agreement (Y/N)
Cost to have adequate state agency (public health) staff	Very challenging	Very Challenging	Yes
Cost of system design/software development	Very challenging	Very Challenging	Yes
Cost to have adequate staff/headcount for IT staff (internal)	Challenging	Challenging	Yes
Cost of IT technical support for state agency staff	Challenging	Challenging	Yes
Cost of technical support to end users	Challenging	Challenging	Yes
Cost of servers/hosting Applications	Challenging	Less Challenging	No
Cost of data storage	Less Challenging	N/A	N/A (Item of low importance)
Data Collection			
Challenges in acquiring data that are useful and meaningful	Very challenging	Very Challenging	Yes
Challenges in acquiring data that have been requested by relevant stakeholders/end users	Very challenging	Challenging	No
Challenges in acquiring data from multiple data sources across the state	Very challenging	Challenging	No
Challenges in working with private hospitals and clinics to release data	Very challenging	Challenging	No
Collecting data in a timely manner	Very challenging	Challenging	No
Challenges in working with public hospitals and clinics to release data	Very challenging	Challenging	No
IT Infrastructure			
Collaboration with software developers and IT staff ensuring that requested system design needs are met	Challenging	Challenging	Yes
Challenges in decision making on technology options (open source, commercial, or adopt from another state)	Challenging	Less Challenging	No
Usability			
Data are meaningful and is useful for the end user	Very challenging	Very Challenging	Yes
Data are missing or incomplete for end user	Challenging	Challenging	Yes
Quality of user data output (Excel, csv, pdf, html etc.)	Challenging	Less Challenging	No
Drill Downs/Data Filers are difficult to understand	Challenging	Less Challenging	No
Website freezes up	Challenging	Less Challenging	No
Data do not go far back enough in time	Challenging	Challenging	Yes

Discrepancies in findings between literature review and our two studies

There were vast differences in what was found in the systematic literature review, Manuscript #2 and Manuscript #3. In the literature we reviewed, most of the conclusions authors arrived at were based on general observations rather than evidence. As an example, observations were made that systems were difficult to use and emphasized problems with usability, but such observations were from the perspective of the author(s). Our State Health Coordinator Study found that information technology is not the problem, but state agencies are lacking the staff and resources to produce modern systems without high levels of functionality. None of the literature we reviewed included quantitative surveys of end users or state agency personnel. The literature

review included a heavy emphasis on information technology, the methods used in system development, and challenges with system development. It is clear from the results of the qualitative and quantitative studies we conducted that the existing literature on this topic is incomplete, and that the additional challenges and limitations we observed need to be addressed to produce a more comprehensive, robust, and functional public health data system.

Recommendations

Our study provides the basis for a number of recommendations related to strengthening and improving public health data systems in general and WDQS systems in particular.

Recommendation #1: Implementation of a centrally managed cloud based solution funded collaboratively by states

A centrally managed system has been proposed, but its efforts have not been successful, as there are challenges in transmitting large datasets. However, with the growth and emergence of cloud based technology, there is potential. Cloud based solutions are configurable computing resources which allow capabilities to store and process their data in a third-party data center (Hassan, 2011). There are privacy risks, given the sensitive nature of health data. If we are able to overcome risks of security, cloud based technologies could help centralize data. Our recommendation is to design a centrally managed system where state agencies are able to upload their surveillance data. The short term goal in this process is to create the infrastructure of a centrally managed system. This includes the managing of large data sets, and defining requirements. The long term goal is to create a centrally managed system that includes data from all states and is available for end users. As states have limited resources, this would reduce the cost of developing and maintaining systems. A centrally managed system would be efficient, as it would prevent duplicate development efforts across states. Such a comprehensive system

would allow state agencies to increase their efforts with data collection and other public health priorities. States that participate would collaboratively pay for this system and require a pay-in by each state, depending on size/population of the state. This solution would be efficient as states would be sharing costs rather than implementing their own solutions. However, such efforts would require standardization of platforms and vocabulary, so data can be easily shared between systems.

Recommendation #2: Standardization and interoperability of platforms

It is recommended that we aim for the standardization of platforms, as this can improve the flow of information from one system to another. Interoperability is the ability of different information technology systems and software applications to communicate, exchange data, and use the information that has been exchanged (Bloom, 2014). Other areas of healthcare have standardized and interoperable platforms, such as EMRs that use health information exchanges (HIE). Installing common platforms within states and between states may help increase interoperability.

As a first step in pursuing standardization and interoperability, we should systematically research available technologies to ensure the best path is taken. This can be accomplished by investigating existing platforms and interviewing system owners. Those platforms that are well perceived should be considered. In parallel, we must also investigate newer platforms and technologies not being currently used. The proposed project can be administered through the CDC, since they directly work with BRFSS coordinators at the state level. Conducting a study of this magnitude may have a long term benefit, as we can aim towards standardization over the next five to seven years.

Recommendation #3: Standardization of data and vocabulary

It is recommended that common data standards and coding are adopted both within states and between states for population level health data. For health information systems to be effective at improving population surveillance functions, standardization is necessary (Dixon, 2014). Standardization would help improve the flow of information from system to another, and reduce human intervention. Data standards help ensure consistency across multiple sources (Dixon, 2014). Currently, diagnosis codes in health surveillance datasets vary from organization to organization, creating extra time and resources which are spent on interpretation. Standardization has been effective in other areas of healthcare. For example, hospitals claims have been standardized by consistently updating and modifying ICD codes. Secondly, health surveillance systems must work towards uniformity when working social determinants such as race/ethnicity. In this case, if a dataset reports on race and ethnicity as two independent variables, while another only uses race, it is difficult to align or ‘match’ datasets, in which case human effort is then required to reclassify data so they can be used and published. A large percentage of staff time is currently allocated to data manipulation and other manual tasks that could be automated. State agency personnel resources are limited, with an average of only 2-3 people on staff to support public health data systems. The standardization of social determinants would reduce human efforts, increase efficiency, and reduce the chances of publishing incomplete data.

This study also found that there is a lack of standardization of vocabulary between systems. We must create consistency in measures and terminology used. To illustrate, some states use the term ‘percent obese’ while others use ‘percent overweight.’ Both imply a condition of excess weight in a human being. However, the ambiguity in terminology may confuse a system user, which could lead to misreporting. Significant time and resources are

spent on interpreting terminology from one dataset to another. We must aim at creating common data standards that would help identify discrepancies from one system to another. This would help end users, particularly if they are comparing data across multiple states. States must first work together to agree on terminology for reportable health measures. It is recommended that this process should involve creating a steering committee of system owners across multiple states. States must be aware of the importance of adopting these standards as the ultimate goal is for all states to adopt.

Recommendation #4 Implementation of a data sharing framework for population level health data

A data sharing framework can benefit state agencies as restrictive data access policies, bureaucratic hurdles, lack of commitment to data sharing, and a lack of trust between organizations have been a problem (Vest, 2013). Building trust between agencies is important, as organizations must work together to fulfill the mission of health surveillance and public health. A data sharing framework should work to ensure that states monitor the health of the population and share data with one another to fulfill their responsibilities. For example, EHR/EMR and other types of hospital/clinical level data have worked towards a national data sharing framework (Diamond, 2009). Such a framework has not been implemented for population level health data. It is recommended that a larger scale organization such as the CDC active be involved in implementation of this framework, so it can be applied across all states.

Future Directions

The importance of disseminating population level health data must be acknowledged. The growth of the internet has created new opportunity for expanding the availability of this data. Data storage, hardware/software is cheaper than ever, and processors are faster now that

ever. There is potential for using local health data to combat many health problems at the community level. We must work towards overcoming the barriers highlighted in this dissertation. Expansive efforts must be directed at designing, standardizing, and implementing new systems, as the next generation must continue in the path that was employed between 1998 and 2005. Over that time period, a large number of states received funding, as the growth of the internet was in its infancy.

Implementation and upgrading existing systems have stalled, despite the known benefits and importance of them. There has been selective and limited funding that has been available for WDQSs. As an example, nine states received 5-year federal funding in 2007 through the CDC Assessment Initiative. Using those funds, the Illinois Department of Public Health developed the robust *IQuery* system that links multiple data sources across the state. The Rhode Island Assessment Initiative for Data Dissemination (RI AIDD) also received funding through the CDC Assessment Initiative for WDQS development. With the funding they received, they were able to develop a static data interface but were unable to design an interactive based system. WDQS funding has been selective, and only a few states are able to maximize their return with the funding they receive. Some states have greater staffing resources, a strong infrastructure and are able to make better use of federal funds. Since funding is currently limited, greater efforts need to be made towards collaboration between states. Currently, there is an imbalance of technology and resources between states, and adopting systems from other states and sharing resources may benefit many states. As stated in this dissertation, it is important that states with funding should aim towards creating scalable solutions that are adoptable by multiple states. States must work in unison for a common public health goal, which is the improvement of the overall health of our population.

REFERENCES

- Abii, F. **Effects of Individual and Organizational Factors on the Turnover Intentions of Information Technology Professionals.** *International Journal of Management* 2013, (2):6.
- AbouZahr C. **From data to policy: good practices and cautionary tales.** *Lancet* 2007, 369(9566):1039-1046.
- Aggelidis, A. **Hospital information systems: Measuring end user computing satisfaction (EUCS).** *Journal of Biomedical Informatics* 2012, 45:3:566-579.
- Allen D. K. **Information sharing and interoperability: the case of major incident management.** *European Journal of Information Systems* 2014, 23(4).
- Anderson M, Seltzer W. **Federal statistical confidentiality and business data: twentieth century challenges and continuing issues.** *J Privacy and Confidentiality* 2009, 1(1):7-5.
- Arts, D. **Defining and Improving Data Quality in Medical Registries: A Literature Review, Case Study, and Generic Framework.** *Journal of the American Informatics Association* 2002, 9(6):600.
- Attieh, R. **Organizational readiness for knowledge translation in chronic care: a Delphi study.** *BMC Health Services Research* 2014, 14:534.
- Baldwin W, Diers J. **Demographic Data for Development in Sub-Saharan Africa.** New York: Population Council, New York; 2009.
- Barton, B. **Learned in Healthcare Information Technology.** *The Journal for Quality and Participation* 2014, 7(33): 22-24.
- Bates DW. **The quality case for information technology in healthcare.** *BMC Med Inform Decis Mak* 2002, 2:7.
- Bennett, J. L. **Managing to meet usability requirements: establishing and meeting software development goals.** Prentice Hall, Englewood Cliffs, NJ, 1984, 161-184.
- Bernstein A. **Public Health Surveillance Data: Legal, Policy, Ethical, Regulatory, and Practical Issues,** *Morbidity and Mortality Weekly Report.* 2012, 61(03):30-34.
- Birkhead G. **Public Health Surveillance Using Electronic Health Records: Rising Potential to Advance Public Health.** *Frontiers in Public Health* 2015, 4:5.
- Bloom Y. **Data uses, benefits, and barriers for the behavioral risk factor surveillance system: a qualitative study of users.** *J Public Health Manag Pract.* 2000. 6(1):78–86.

- Bond J. **A Delphi study of clinical nursing research priorities.** *Journal of Advanced Nursing* 1982, 7, 565–575.
- Braithwaite G, Haggard LM. **Utah’s IBIS-PH: an innovative user interface solution for Web-based data query systems.** *J Public Health Manag Pract* 2006, 12(2):146–54.
- Brooks, F. **The Mythical Man-Month- Essays on Software Engineering.** Addison-Wesley, Boston, 1995.
- Brownson RC, Petitti DB. **Applied Epidemiology: Theory to Practice.** 2nd ed. New York. Oxford University Press, 2011.
- Busch P. **Organisation Design And Tacit Knowledge Transferral: An Examination Of Three IT Firms.** *Journal of Knowledge Management Practice* 2006, (7):2.
- Byass P. **The unequal world of health Data.** *LoS Med* 2009, 6(11).
- Cavanagh S. **Content analysis: concepts, methods and applications.** *Nurse Researcher* 1997, 4(3), 5-16.
- Chambers LW. **Health surveillance: an essential tool to protect and promote the health of the public.** *Canadian Journal of Public Health* 2006, 97:3.
- Chandramohan D. **Should data from demographic surveillance systems be made more widely available to researchers?** *PLoS Medicine*. 2008; 5(2):0169-73.
- Chang, C.J., & Hwang. **Accounting education, firm training and information technology,** 2003.
- Chapin, J. **Types of software evolution and software maintenance.** *Journal of Software Maintenance and Evolution: Research and Practice* 2001, 13,1: 3–30.
- Cheng C. **Electronic school absenteeism monitoring and influenza surveillance,** *Emerging infectious diseases* 2012, 18: 885–887.
- Cho P. **An empirical model of mainframe computer investment.** *Journal of Applied Econometrics*, 2011, (1), 122–150.
- Choi BC. **The Past, Present, and Future of Public Health Surveillance.** *Scientifica* 2012.
- Clemens M. **Why do programmers earn more in Houston than Hyderabad?: evidence from randomized processing of US visas.** *The American economic review* 2013, 103: 3.
- Cohen BB, Franklin S, West JK. **Perspectives on the Massachusetts Community Health Information Profile (MassCHIP): developing an online data query system to target a variety of user needs and capabilities.** *J Public Health Manag Pract* 2006, 12(2), 155-60.

Curtis B. **A Field Study of The Software Design Process For Large Systems, Communications of the ACM.** *Interacting with Computers*, 2010, 22(5).

Daramawan, G. **Creating a Scale as a General Measure of Satisfaction for Information and Communication Technology User.** *Applied Rasch Measurement: A Book of Exemplars Education in the Asia-Pacific Region: Issues, Concerns and Prospects* 2005, 4; 271-276.

Diamond CC. Mostashari F, Shirky C: **Collecting and sharing data for population health: a new paradigm.** *Health Affairs* 2009, 1(2):454-66.

Diamond IR, Grant RC. **Defining consensus: a systematic review recommends methodological criteria for reporting of Delphi studies.** *J Clin Epidemiol* 2014, 67(4):401-409.

De Vet E. **Determinants of forward stage transitions: a Delphi study,** *Health Educ. Res.* 2005, 15:195–205.

Dixon, B. **Towards Interoperability for Public Health Surveillance: Experiences from Two States.** *Journal of Biomedical Informatics* 2014, (49):3-8.

Dixon BE and Grannis SJ. **Public Health Informatics and Information Systems.:** Chapter 5, *Public health informatics infrastructure*; Springer-Verlag. 2014, p. 69-88.

Ekemeier B,. **Local public health resource allocation: limited choices and strategic decisions.** *American journal of preventive medicine.* Dec 2013;45(6):769-775.

Farmer EA. **A Delphi study of research priorities in Tech-Prep.** *Journal of Vocational and Technical Education* 1998, 15([1](#)):42-49.

Fielding JE. **Local knowledge to enable local action.** *Am J Prev Med* 2004;27:183–184.

Fink, A. **How to conduct surveys: A step-by-step guide.** Los Angeles: 2009, SAGE Publications.

Finkelstein, E. **State-level estimates of annual medical expenditures attributable to obesity.** *Obesity Res*, 2004 12:18-24.

Fischer, R. G. **The Delphi method: a description, review and criticism.** *The Journal of Academic Librarianship* 1978, 4(2), 67-70.

Friedman DJ, Parrish RG. **Characteristics and desired functionalities of state web-based data query systems.** *J Public Health Manag Pract* 2006, (1)2:119–29.

Friedman DJ. **Envisioning futures for state Web-based data query systems.** *J Public Health Manag Pract* 2006;12:196–200.

- Friedman, DJ: **The population health record: concepts, definition, design, and implementation.** *J Am Med Inform Assoc.* 2010, 17(4): 359–366.
- Gibbs W, Graves P. **Evaluation guidelines for multimedia courseware.** *Journal of Research on Technology in Education*, 2001, 34(1), 2-17.
- Gjelsvik A. **States approaches to implementing a Web-based data query system.** *J Public Health Management Practice* 2006,12(2):170-5.
- Gliem, J. **Calculating, Interpreting, and Reporting Cronbach’s Alpha Reliability Coefficient For Likert-Type Scales.** *Midwest Research to Practice Conference* 2003, 82-88.
- Gratch H. A. **Consensus Measurement in Delphi Studies.** *Technology Forecasting and Social Change* 2012, 79, (8).
- Gunasekaran, A. **A model for investment justification in information technology projects.** *International Journal of Information Management.* 2001, 21, 349-364.
- Haggard L.M. and Burnett S.J. **Measuring the Impact of a Web-based Data Query System: The Logic Model as a Tool in the Evaluation Process.** *Journal of Public Health Management and Practice.* 2006, 12(2),189-195.
- Hassan Q. **Demystifying Cloud Computing.** *The Journal of Defense Software Engineering* 2011, 16–21.
- Hasson F. **Research guidelines for the Delphi survey technique.** *Journal of Advanced Nursing* 2000, 32: 1008-1015. 52.
- Haq U. **Validation in layered cloud infrastructures,** *Economics of Grids, Clouds, Systems, and Services* 2010, 153-164.
- Healthy People 2000: National Health Promotion and Disease Prevention Objectives.** Washington, DC: US Dept of Health and Human Services; 1991.
- Hesse B. **Public Health Surveillance in the Context of Growing Sources of Health Data: A Commentary.** *American Journal of Preventive Medicine* 2011, (41)6: 648-649.
- Hornbæk K. **Current practice in measuring usability: Challenges to usability studies and Research.** *International Journal of Human-Computer Studies* 2006, 64(2), 79-102.
- Hui, L. **GIS-Based Public Health Surveillance Systems: A Systematic Review.** *ISPRS Int. J. Geo-Inf.* 2014, 3(2), 481-50
- Husion A. **Fighting turnover,** *Computing Canada* 1992, 11:11.
- Hsu C, Sandford B. **The Delphi technique: Making sense of consensus.** *Practical Assessment, Research and Evaluation* 2007, 12, 1—8.

- Irani P. **Community health assessment: driving the need for current, easily accessible population health data.** *J Public Health Manag Pract* 2006, 12(2):113–118.
- Iyamu, T. **The Impact of Network of Actors on the Information Technology.** *Actor-Network Theory and Technology Innovation: Advancements and New Concepts.* 2011, 234-246.
- Jairath N, Weinstein, J. **The Delphi methodology: a useful administrative approach.** *Canadian Journal of Nursing Administration* 1994, 7:29-42.
- Jefferson, T, E. **Measuring quality of editorial peer review.** *The Journal of the American Medical Association.* 2002, 287(21):2786-2790.
- Johnson P. **Four decades of population health data: the integrated health interview series as an epidemiological resource.** *Epidemiology.* 2008;19(6):872–875.
- Jones J. **What will happen to the quality of care with fewer junior doctors? A Delphi study of consultant physicians' views.** *Journal of the Royal College of Physicians* London 1992, 26, 36–40.
- Juran, J.M. **Juran's Quality Handbook.** 5 ed. 1999, New York: McGraw Hill.
- Kanouse DE. **Dissemination of effectiveness and outcomes research** *Health Policy* 1995, 34:167–192.
- Keller M, Blench M. **Use of unstructured event-based reports for global infectious disease surveillance.** *Emerg Infect Dis* 2009, 15(5):689-95.
- Knowler, W. **Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,** *N England Journal of Medicine,* 2002. 346:393-403.
- Kondracki N, Wellman, N. **Review of methods and their applications in Nutrition Education.** *Journal of Nutrition Education and Behavior* 2002, 34, 224-230.
- Kopp S. **Telemedicine/telehealth: an international perspective. Public health applications:** *Telemed J E Health.* 2002 Spring;8(1):35-48.
- Lawler, E. E. **Human resources business process outsourcing: Transforming how HR gets its work done.** San Francisco: Jossey-Bass 2004.
- Lee LM, Heilig CM. **Ethical justification for conducting public health.** *Am J Public Health.* 2012, 102(1): 38–44.
- Lee, Sung Heum. **Usability testing for developing effective interactive multimedia software: Concepts, dimensions, and procedures.** *Educational Technology & Society* 1999, 2 (2)
- Lofmark. A. **An assessment form for clinical nursing education: a Delphi study.** *J Adv Nurs.* 2004, 48(3):291-8.

- Love D, Shah GH. **Reflections on organizational issues in developing, implementing, and maintaining state web-based data query systems.** *J Public Health Manag Pract.* 2006;12(2):184-18.
- Luck J, Chang C. **Using local health information to promote public health.** *Health Affairs.* 2006, 25(4):979-9.
- Mahon BE. **Implications for Registry-Based Vaccine Effectiveness Studies from an Evaluation of an Immunization Registry: A Cross-Sectional Study.** *BMC Public Health.* 2008, 8, 160.
- Manos, D. **Health IT not keeping pace with big data.** *Healthcare IT News* 2014, (2), 12.
- Maylahn C, Alongi S, Alongi J, Moore MJ, Anderson LA. **Data Needs and Uses for Older Adult Health Surveillance: Perspectives From State Health Agencies.** *Prev Chronic Dis.* 2005, 2(3):A02.
- McKenna HP. **The Delphi technique: a worthwhile research approach for nursing?** *Journal of Advanced Nursing* 1994, 19: 1221-1225.
- Melber, D. **MCSE, MVP.** *Internal Auditing*, 2013, (28).4: 45-47.
- Moher D. **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** 2009, *PLoS Med* 6(7).
- Monkman, H. Kushniruk, A. **Applying usability methods to identify health literacy issues: an example using a Personal Health Record.** *Sty Health Technol inform* 2013;183-179-85.
- Murray, CL. **The State of US Health, 1990-2010 Burden of Diseases, Injuries, and Risk Factors.** *JAMA*, 2013 310(6):591-606.
- Myers J, Frieden TR. **Privacy and Public Health at Risk: Public Health Confidentiality in the Digital Age.** *American Journal of Public Health.* 2008, (5):793–801.
- Narayan K. **Disease Control Priorities in Developing Countries**, 2nd ed. Edited by Jamison DT, Breman JG, Measham AR. Washington, DC, The World Bank; 2006.
- National Association of County & City Health Officials. **National Profile of Local Health Departments**, 2011: Washington, DC.
- Niu, J. **Overcoming inadequate documentation.** *Proceedings of the American Society for Information Science and Technology.* 2009, 46:1.
- Nsubuga P. **Public health surveillance: a tool for targeting and monitoring interventions.** Washington, DC: World Bank Publishers 2006, 997–1015.

- Okoli C. **The Delphi method as a research tool: an example, design considerations and applications.** *Information & Management* 2004, 42(1), 15–29.
- Ortiz E. **Use of Information Technology to Improve the Quality of Health.** *Health Serv Res.* 2003, Apr; 38(2):
- Park, Y.A. and Gretzel, U. **Success Factors for Destination Marketing Web sites: A Qualitative Meta-Analysis.** *Journal of Travel Research.* 2007, 46(1), 46-63.
- Portnoy B. **Independent state health surveys: responding to the need for local population health data.** *J Public Health Manag Pract.* 2014, 20(5):E21-33.
- Raghupathi W. **Big data analytics in healthcare: promise and potential.** *Health Information Science and Systems* 2014, 2:3.
- Roberts J. **Entrepreneurs, Investors Can Find Opportunities in Lengthening downturn.** *Electronic Business* 2002, 7:74.
- Robinson KR. **Developing the principles of chair based exercise for older people: a modified Delphi study.** *BMC Geriatrics* 2014, 14:65.
- Rose A. **Using qualitative studies to improve the usability of an EMR.** *Journal of Biomedical Informatics,* 2005, 38(1): 51–60.
- Rowe G, Wright, G. **The Delphi technique as a forecasting tool: Issues and analysis.** *International Journal of Forecasting* 1999, 15, 353-375.
- Rudolph B.A. **Small numbers, disclosure risk, security, and reliability issues in web-based data query systems.** *J Public Health Management Practice* 2006, 12(2) 176-183.
- Savel T. **The role of public health informatics in enhancing public health surveillance.** *CDC's Vision for Public Health Surveillance in the 21st Century. Morbidity and Mortality Weekly Report* 2012, 61(3), 20-24.
- Savel TG, Foldy S. **The role of public health informatics in enhancing public health surveillance.** *MMWR Surveill Summ* 2012, (2)61, 20-24.
- Schofield J. **Developing a Quality Plan.** *Health Care Quality Activities in Practice in Australia,* 4th Edition, 2012.
- Schumacher P. **Responding to the changing needs of public health assessment in the information age: the evolution of a program.** *J Public Health Manag Pract.* 2006, (2)2:109-12.
- Seddon, P. **Dimensions of information systems Success.** *Communications of AIS* 1999, 2, 20.

Sekaran, U, **Research Methods for Business. A Skill Building Approach**, 4th ed. John Wiley & Sons, New York, 2003.

Setel P, **Who Counts (1): A scandal of invisibility: making everyone count by counting everyone.** *Lancet* 370 2007, 1569–1577.

Shaame, A. **An educational framework for free and open source software**, *International Journal of Innovation Management and Technology*, 2003, 4(1), 348-355.

Shah GH1, Leep CJ. **Profile-IQ: Web-based data query system for local health department infrastructure and activities.** *J Public Health Manag Pract* 2014, 20(2):168-74.

Simon, P. **Meeting the Data Needs of a Local Health Department: The Los Angeles County Health Survey.** *Am J Public Health*. 2001 (12): 1950–1952.

Smith, P. F. **Updating public health surveillance for the 21st century.** *J Public Health Manag Pract* 2003, 19(3), 231-239.

Skulmoski,J. **The Delphi Method for Graduate Research** *Journal of Information Technology Education* 2007, 6:1.

Solet, D. **VISTA/PH software for community health assessment.** *J Public Health Manag Pract*. 1999, 5(2):60–63.

St. Laurent, AM. **Understanding Open Source and Free Software Licensing.** O'Reilly Media, 2008.

Studnicki J. **Evaluating the Performance of Public Health Agencies: Information Needs.** *American Journal of Preventive Medicine*, 1995 6:74-80.

Thacker SB. **A method for evaluating systems of epidemiological surveillance.** *World Health Statistics Quarterly*. 1988, 41(1):11–18.

Thacker SB. **Public health surveillance in the United States: evolution and challenges.** *Morb Mortal Wkly Rep Surveill Summ* 2012, 61:3-9.

United States Bureau of Labor Statistics, 2011.

Van Panhuis WG. **A systematic review of barriers to data sharing in public health.** *BMC Public Health* 2014, 1(4):1144.

Vandelanotte C. **The development of an internet-based outpatient cardiac rehabilitation intervention: a Delphi study.** *BMC Cardiovascular Disorders* 2010, 10:27.

Ver Ploeg M, Perrin E. **Eliminating Health Disparities: Measurement and Data Needs.** *National Research Council Report*. Washington, DC: National Academies Press 2006.

Verhagen AP. **The art of quality assessment of RCTs included in systematic reviews.** *J Clin Epidemiol.* 2001, 54:651–654.

Vest JR, Issel D. **Data Sharing between Local Health and State Health Departments: Developing and Describing a Typology of Data Gaps.** *J Public Health Manag Pract* 2013, 19(4):357-65.

Von der Gracht, HA. **Consensus measurement in Delphi studies.** *Technological Forecasting and Social Change* 2012, 79: 1525–1536.

Vosbergen S. **Evaluation of end-user satisfaction among employees participating in a web-based health risk assessment with tailored feedback.** *J Med Internet Res* 2012, 14(5):140.

Walley J. **Public Health: An action guide to improving health (second edition),** Oxford University Press, Oxford, UK, ISB.

Walport M, Brest P. **Sharing research data to improve public health.** *Lancet* 2011, 377(9765):537-53.

Wang Y1, Tao Z. **Development of a web-based integrated birth defects surveillance system in New York State.** *J Public Health Manag Pract* 2008, 14(6):E1-E1.

Ward, L. **Establishing key components of yoga interventions for musculoskeletal conditions: a Delphi survey.** *BMC Complementary and Alternative Medicine* 2014, 14:196.

Warnecke RB, Oh A, Breen N. **Approaching health disparities from a population perspective: the National Institutes of Health Centers for Population Health and Health Disparities.** *Am J Public Health* 2008, 98:1608–15.

Wartenberg D, Thompson WD. **Privacy versus public health: the impact of current confidentiality rules.** *Am J Public Health* 2010, 100(3):407-412.

Wartenberg D, Thompson WD. **Privacy vs. Public Health.** *American Journal of Public Health.* 2010, 100(3): 407-412.

Wentholt, A. **Effective identification and management of emerging food risks: results of an international Delphi survey** *Food Control* 2010, 21(12): 1731-1738.

Wessells, P. **Justifying the investment in information Systems.** *South African Journal of Information Management* 2003, 5(2).

Zinn TK. A CHIN primer. *Health Management Technology* 1994;15:28.