Evidence-Based Public Health: Using Data Visualization for Improving the Understanding of Data and Information

Seung Eun Park

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Betty Bekemeier, Chair

Abie D. Flaxman

Brock Craft

Program Authorized to Offer Degree:

School of Nursing

University of Washington

Abstract

Evidence-Based Public Health: Using Data Visualization for improving the understanding of data and information

Seung Eun Park

Chair of the Supervisory Committee:

Betty Bekemeier

Department of Psychosocial and Community Health

Despite the potential value of data visualization for improving understanding of data in a manner that results in better decision-making in health and healthcare, there is little understanding of its applications in public health practice, particularly for public health program planning and resource allocation. This dissertation explores what can be learned from the literature through systematic review and from interviews with public health professionals as a means to inform improvements in decision-making among public health professionals by developing an understanding and use of data through data visualization.

In the first paper, I provide a systematic review exploring the literature that examines how visualization can impact decision-making for general populations. Even though the evidence is limited due to a deficit of theoretical and methodological strength, the studies suggest that

interventions that include data visualization have a positive impact on cognitive and behavior change such as decision-making, attitude, motivation or perception.

In the second paper, I explore using qualitative data about how public health professionals use data, information, and evidence for their practice, especially for their decision-making, and their current use of and preferences for data visualization. Public health leaders use data, information, and evidence from various resources on a daily basis for communication with co-workers, stakeholders, and the public and for decision-making regarding their programs and services in various settings. They also have a high interest in applying visualization skills or tools in their public health practice as well as preferences regarding features or types of data visualization they want to have.

In the third paper, I examine how data presented to public health professionals via visualization was understood differently compared to data presented in a more traditional table format. Even with the small sample participating in this study, I found that there are apparent benefits to using data visualization, such as making finding information from data easier and reducing errors in tracking data. This paper focused on how data visualization improves the "making sense of data" step – a step followed by decision-making processes, analyzing and summarizing data, translating data into information, and then synthesizing the information into knowledge.

In conclusion, this dissertation presents the current evidence of data visualization and its application to public health professionals. This work suggests that data visualization could be an effective approach for improving decision-making and communication among public health professionals.

TABEL OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my dissertation committee, Dr. Betty Bekemeier, Dr. Abie Flaxman, Dr. Brock Craft, Dr. Jenny Tsai, and Dr. Darryl Holman, for sharing their time and expertise in helping me complete my dissertation research.

I particularly and heartily thank my advisor, Dr. Betty Bekemeier, who spent countless hours advising and supporting my research from the beginning. Her guidance and encouragement throughout the entire project gave me a strong foundation for becoming a researcher. She graciously let me become involved with her research team, the Public Health Activities and Services Tracking (PHAST), as a Research Assistant for over five years, and throughout this period, I have gained uncountable and valuable research experience above what I gained through my Ph.D. program and dissertation research. I also thank Dr. Flaxman, who was always willing to help me when I encountered questions and to suggest a better direction for research. I was always aspired by his sharp comments that were essential in filling the holes in my research. I am also grateful for the support of Dr. Craft who infused perspectives from data science and has introduced me to the theory and research of data visualization that was not within the traditional scope of nursing and public health. His encouragement gave me the confidence to continue my research related to data visualization. Dr. Tasi was also a thoughtful committee member who provided different views and provided the right amount of critique and affirmation. Dr. Holman as a Graduate School Representative (GSR), made sure my dissertation process and exam were fair as well as provided brilliant ideas for my research. Finally, I express my gratitude to Dr. Karen Schepp, who was my academic adviser in the first year of my Ph.D. program as well as a committee member but had to leave my committee due to her retirement.

She was not only a great academic adviser but also a mentor who made sure I was on the right track in my Ph.D. life.

Sincere thanks are extended to the Hester McLaws Nursing Scholarship Awards who provided funding for my dissertation. I could not have accomplished any of these goals without the support of faculty members, my student colleagues, and friends, as well as my family.

CHAPTER 1. INTRODUCTION

BACKGROUND

Public health leaders are encouraged to use "data and information, and evidence systematically" when making decisions for resource allocation and public health planning for disease prevention and health promotion programs in order to incorporate the concept of Evidence-Based Public Health (EBPH).[1–3] Evidence-Based Public Health (EBPH) is defined as "conscientious, explicit, and judicious use of current best evidence in making decisions about the care of communities and populations in the domain of health protection, disease prevention, health maintenance and improvement (health promotion)."[3] Brownson and others suggest that public health professionals should adopt the EBPH framework to optimize their decision-making.[1,2,4–6] They are to do this based on their understanding of what factors influence which population and what intervention works to address health issues in certain populations as that understanding is informed by available scientific evidence collected through systematic uses of data and information systems.[1–3,6]

One of the eight domains of core competencies for public health professionals is analytical and assessment skills, which include collecting, understanding, analyzing, and interpreting valid and reliable data, and making evidence-based decisions to address community health needs.[7] In a study assessing how urban US city health departments use data to inform their work, local public health leaders responded that they believed using their local data were good for informing decisions for public health programs and policies and educating decision-makers about certain community health problems. For example, the study explains how using data to cite notably higher smoking rates in their district compared to ones in neighboring districts drew more attention to the problem from decision-makers.[8]

The value of adopting EBPH strategies has been discussed in the literature, including its numerous benefits such as supporting access to better information, efficient public health planning and the implementation of successful policies in a public health environment characterized by limited resources.[9] In a survey research study conducted to understand perceptions about EBPH, directors, and managers from Local Health Departments (LHDs) strongly agreed that positive changes in their agencies could occur as a result of evidence-based decision-making, which is a process of EBPH. [10] Making the best decisions on prevention programs and policies is crucial considering their broad impact on a range of populations.[6]

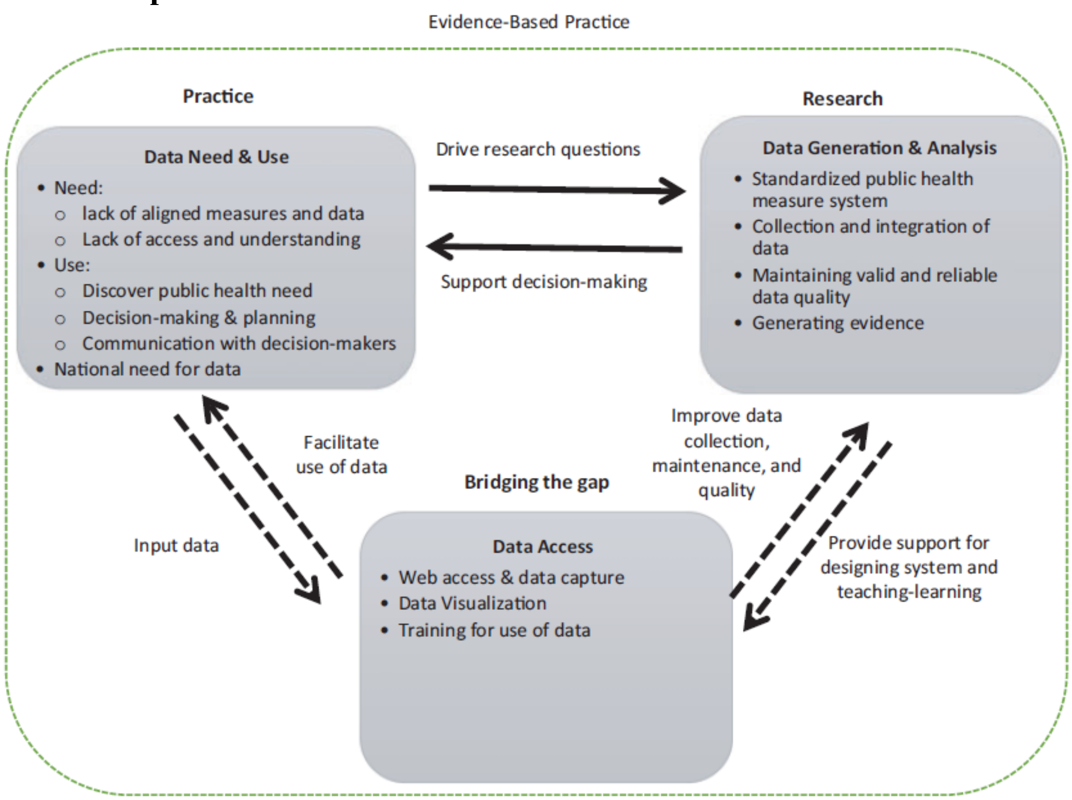
Public health professionals, however, underutilize data, information, and evidence for assessment and analysis of public health issues as guides in decision-making,[11] and decisions are often not based on a systematic review of the best evidence, but on many other factors.[2,12] Various barriers that prevent the translation of research evidence into public health practice are discussed in the literature. These include limited and inflexible funding and rigid bureaucratic government systems and leadership, as well as a lack of incentive and expertise within the workforce to explore and use data effectively. [1,11–14] In 2010, only 34% of LHDs in the U.S. employed an epidemiologist—a position that can contribute data and analyses to support evidence-based decision-making. [14(ps193)] Even before the 2008 U.S. economic recession, LHDs had faced severe cuts to budgets, programs, and staff, [15]and desired a means to be more strategic in allocating limited funding to effective programs. [6] Even with limited and reduced budgets, these agencies have continued to need to maintain critical public health functions to assure high priority health care services such as communicable disease control, environmental health, and chronic disease prevention for vulnerable populations.[13]

The literature consistently indicates the need for public health practitioners to have accessible, "timely, easy to digest, and up-to-date information that is filtered, summarized, and synthesized from authoritative content sources."[16] A recent National Academy of Medicine (NAM) report further emphasized this need by describing the importance of information and information systems for monitoring community health needs while expressing concern about "inadequate access to information systems and communication tools" for the public health infrastructure.[17]

Data visualization can be an approach to overcoming some of those barriers to implementing EBPH, such as lack of time or expertise to analyze data and synthesize information by improving access to data for public health professionals.[2] In a previous study describing the development of a Public Health Activities and Services Tracking (PHAST) Model as a means to help address the gap between public health research and practice (Figure 1.1),[18] data visualization was discussed as a means to make data more accessible and useful for public health professionals in their planning and decision-making. In the development of the PHAST Model, insight regarding the benefits of data visualization was initially derived from public health professionals through focus groups and then used to help build the model.[18]

Data visualization refers to visual representations such as graphs and maps used to understand, communicate, and support collaborative engagement and decision-making with data and information.[19–22] By using data visualizations, people can better understand data, gain insight, answer specific questions, and discover underlying facts.[19,21–23] Humans process visual representations, which are usually made with computers, through a cognitive system.[22] According to a recent computer science study, visualizing information was even effective in changing attitudes regarding certain topics among general audiences when compared with tabular displays of the same information. [24]

Figure 1.1. The Public Health Activities and Services Tracking (PHAST) model for standardized public health data



One of the first innovative visualizations related to health can be traced back to 1858.[9] Florence Nightingale, known today as a statistician and the founder of modern nursing, used a graph to demonstrate that deaths from preventable factors such as sanitation were far greater than deaths from combat wounds.[10] Her graph was persuasive to Queen Victoria and the British Government, who were unlikely to read her entire 830-page report, and led to improved sanitary systems in military hospitals.[9]

In recent years, numerous studies from various fields, especially computer science and information science, have developed data visualization tools and described the utility and usability of the tools in their field. As the example of Florence Nightingale and the literature suggest, data visualization can be a particularly useful and effective tool for improving public health by helping to generate insight through graphs that reflect current problems or simulate predictions of future problems or solutions based on current data.[27] Yet there are very few studies of data visualization used in public health, particularly for public health program planning and resource allocation, assessing how the tools help develop a "deeper level of understanding" of data and information.[11,12] Furthermore, little is known about what the needs and preferences of public health practitioners are for visual data and information that address local public health issues.

OBJECTIVES

My dissertation addresses this research gap through the following three objectives:

1. Systematically reviewing the available literature on the use of data visualization in order to assess the status of current science and evidence regarding its impact on decision-making related behavior as informed by cognitive processes such as understanding, attitude, and perception, and identify key elements of appropriate study designs for testing this impact.
2. Conducting a qualitative study to assess how public health leaders use data, information, and evidence and their current use of and preference for data visualization in their practice.

3. Conducting a descriptive mixed method study using an online survey to examine how data were understood differently when displayed in table format versus visualized formats by public health practitioners. This study also examined practitioners' confidence in their understanding and perceived ease of use with table and visualization presentations.

This dissertation consists of three papers. The first paper (Chapter 2) of the dissertation reports findings from a systematic review that summarizes theories, methodological details, and results of the experimental and quasi-experimental studies that have aimed to examine how data visualization impacts attitude, motivation, perception, or decision-making.

The second paper (Chapter 3) describes findings from a qualitative analysis of 14 individual interviews conducted with public health leaders to assess their use of data, information, and evidence for their decision-making as well as to understand their current use of and preferences for data visualization. This research was conducted to inform the design of a data visualization tool and to examine how to further public health practitioners' engagement with and use of data for decision-making through data visualization.

The third paper (Chapter 4) presents findings from a descriptive mixed method study exploring how data presented via visualization was understood differently compared to data presented in a more traditional table format. LHD participants were asked to answer open-ended questions by summarizing what they understood through looking at data in both table and visualization formats. The data from the open-ended questions were used for qualitative analysis. Quantitative analyses were used to describe perceived ease of use and scaled confidence between a table and visualized presentation.

Finally, a summary of all three papers is included in Chapter 5. This chapter concludes with a discussion of results and implications for future practice and research in using data visualization to support understanding data for decision-making and communication among public health professionals.

REFERENCES

1. Brownson RC, Fielding JE, Maylahn CM. Evidence-Based Public Health: A Fundamental Concept for Public Health Practice. *Annu Rev Public Health*. 2009;30(1):175-201. doi:10.1146/annurev.publhealth.031308.100134

2. Brownson RC, Gurney JG, Land GH. Evidence-based decision making in public health. *J Public Health Manag Pract*. 1999;5(5):86–97.

3. Jenicek M. Epidemiology, evidenced-based medicine, and evidence-based public health. *J Epidemiol*. 1997;7(4):187–197.

4. Armstrong R, Waters E, Dobbins M, et al. Knowledge translation strategies to improve the use of evidence in public health decision making in local government: intervention design and implementation plan. *Implement Sci*. 2013;8(1):121.

5. Association of State and Territorial Health Officials. Evidence-Based Public Health Implementation Toolkit. http://www.astho.org/Evidence-Based-Public-Health/Toolkit/. Accessed February 7, 2019.

6. Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. *Evidence-Based Clinical and Public Health: Generating and Applying the Evidence*. Office of Disease Prevention and Health Promotion; 2010. https://www.healthypeople.gov/sites/default/files/EvidenceBasedClinicalPH2010.pdf.

7. Council on Linkages Between Academia and Public Health Practice. *Core Competencies for Public Health Professionals*. Public Health Foundation; 2014:25.

8. Castrucci BC, Rhoades EK, Leider JP, Hearne S. What Gets Measured Gets Done: An Assessment of Local Data Uses and Needs in Large Urban Health Departments. *J Public Health Manag Pract*. 2015;21: S38-S48. doi:10.1097/PHH.0000000000000169

9. Lhachimi SK, Bala MM, Vanagas G. Evidence-Based Public Health. *BioMed Res Int*. 2016;2016:1-2. doi:10.1155/2016/5681409

10. Harris JK, Erwin PC, Smith C, Brownson RC. The Diffusion of Evidence-Based Decision Making Among Local Health Department Practitioners in the United States: *J Public Health Manag Pract*. 2015;21(2):134-140. doi:10.1097/PHH.0000000000000129

11. Baum NM, DesRoches C, Campbell EG, Goold SD. Resource allocation in public health practice: a national survey of local public health officials. *J Public Health Manag Pract*. 2011;17(3):265–274.

12. Sosnowy CD, Weiss LJ, Maylahn CM, Pirani SJ, Katagiri NJ. Factors Affecting Evidence-Based Decision Making in Local Health Departments. *Am J Prev Med*. 2013;45(6):763-768. doi:10.1016/j.amepre.2013.08.004

13. Bekemeier B, Chen A, Kawakyu N, Yang Y. Local Public Health Resource Allocation. *Am J Prev Med*. 2013;45(6):769-775. doi:10.1016/j.amepre.2013.08.009

14. Lovelace KA, Aronson RE, Rulison KL, Labban JD, Shah GH, Smith M. Laying the Groundwork for Evidence-Based Public Health: Why Some Local Health Departments Use More Evidence-Based Decision-Making Practices Than Others. *Am J Public Health*. 2015;105(S2): S189–S197.

15. National Association of County & City Health Officials. *2013 National Profile of Local Health Departments*.; 2013. HTTP: //www.naccho.org/topics/infrastructure/profile/upload/ 2013-National-Profile-of-Local-Health-Departments-report. pdf. Accessed December 8, 2015.

16. Revere D, Turner AM, Madhavan A, et al. Understanding the information needs of public health practitioners: A literature review to inform design of an interactive digital knowledge management system. *J Biomed Inform*. 2007;40(4):410-421. doi:10.1016/j.jbi.2006.12.008

17. Institute of Medicine (US) Committee on Assuring the Health of the Public in the 21st Century, ed. *The Future of the Public's Health in the 21st Century*. Washington, D.C: National Academies Press; 2003. https://www.ncbi.nlm.nih.gov/books/NBK221239/.

18. Bekemeier B, Park S. Development of the PHAST model: generating standard public health services data and evidence for decision-making. *J Am Med Inform Assoc*. 2018;25(4):428-434. doi:10.1093/jamia/ocx126

19. Telea A. *Data Visualization: Principles and Practice*. Second edition. Boca Raton: CRC Press, Taylor & Francis Group; 2015.

20. Few S. *Now You See It : Simple Visualization Techniques for Quantitative Analysis*. 1st edition. Oakland, California: Analytics Press; 2009. http://www.tableau.com/blog/stephen-few-data-visualization. Accessed December 16, 2015.

21. Few S. What Is Data Visualization? Visual Business Intelligence. https://www.perceptualedge.com/blog/?p=2636. Published May 4, 2017. Accessed January 23, 2019.

22. Ware C. *Information Visualization Perception for Design*. 3rd ed. Waltham, Mass: Morgan Kaufmann; 2013.

23. Tufte E. *The Visual Display of Quantitative Information*. 2nd ed. Cheshire, Connecticut: Graphics Press; 2001. http://www.edwardtufte.com/tufte/books_vdqi. Accessed January 20, 2016.

24. Pandey AV, Manivannan A, Nov O, Satterthwaite M, Bertini E. The Persuasive Power of Data Visualization. *IEEE Trans Vis Comput Graph*. 2014;20(12):2211-2220. doi:10.1109/TVCG.2014.2346419

25. Meirelles I. *Design for Information*. Beverly, Massachusetts: Rockport Publishers; 2013. http://isabelmeirelles.com/book-design-for-information/. Accessed January 20, 2016.

26. Speyer P, Pagels B, Park N. Communicating data for impact. *Commun Data Impact – Seattle Inst Health Metr Eval Forum One*. 2015.

27. Ryan GW, Bloom EW, Lowsky DJ, et al. Data-Driven Decision-Making Tools To Improve Public Resource Allocation For Care And Prevention Of HIV/AIDS. *Health Aff (Millwood)*. 2014;33(3):410-417. doi:10.1377/hlthaff.2013.1155

28. Brodlie KW, Carpenter L., Earnshaw R., et al. *Scientific Visualization: Techniques and Applications*. Springer Science & Business Media; 2012.

29. Kard ST, Mackinlay JD, Schneiderman B. *Readings in Information Visualization, Using Vision to Think*. San Francisco, Calif: Morgan Kaufmann; 1999.

CHAPTER 2. Impact of Data Visualization on Human Behavior: A Systematic Review

ABSTRACT

Introduction: Data visualization tools have the potential to support decisions for public health professionals. This review summarizes current science and evidence regarding data visualization and its impact on decision-making related behavior as informed by cognitive processes such as understanding, attitude, or perception and identifies key elements of appropriate study designs for testing this impact.

Methods: An electronic literature search was conducted using six databases including reference list reviews. Search terms were pre-defined based on research questions. Relevant studies were carefully selected and reviewed.

Results: Fourteen studies were included in the final analysis. Most data visualization interventions included in this review were found to have an impact on attitude, perception, and decision-making when compared to controls. These relationships between the interventions and the outcomes appear to be explained by mediating factors such as perceived trustworthiness and quality, domain-specific knowledge, basic beliefs shared by social groups, and political beliefs.

Conclusions: The evidence from this review generally suggests positive effects of data visualization depending on the control of confounding factors on attitude, perception, and decision-making. However, understanding about data visualization interventions specific to public health leaders' decision-making is still lacking, and there is little guidance for understanding a participant's characteristics and tasks. Future research to examine the impact of data visualization should be conducted with the appropriate methodology and should target public health leaders to improve the quality of evidence.

Key Words: Data visualization, decision-making, public health practice

INTRODUCTION

Data visualization refers to visual representations such as graphs and maps used to understand, communicate, and support collaborative engagement and decision-making with data and information. [1–4] Humans process visual representations, which are usually made with computers, through a cognitive system. [4] Data visualizations accelerate the interpretation of data, in contrast to numerical charts representing the same data. [4] Data visualizations help people to understand data, gain insight, answer specific questions, and discover underlying facts. [1,3–5] According to a recent computer science study, visualizing information was also effective in changing attitudes on certain topics among general audiences when compared with tabular displays of the same information. [6]

Data visualization tools have the potential to support public health practice by answering questions that previously were unanswerable, such as what makes certain areas in which a population lives or certain demographic groups within a population, healthier than others. Public health leaders in state and local health departments are expected to make the best possible decisions for resource allocation and public health planning. They are to do this based on their understanding of what factors influence which population and what intervention works to address health issues in a certain population and based on available scientific evidence through systematic uses of data and information systems. [7(p177),8,9] Making the best decisions on prevention programs and policies is crucial when considering their broad impact on a range of populations. [10] Since even before the 2008 U.S. economic recession, local health departments (LHDs) have faced severe cuts to budgets, programs, and staff, [11] and desire a means to be more strategic in allocating limited funding to effective programs. [10]

In recent years, numerous studies from various fields, especially computer science and information science, have developed data visualization tools and described the utility and usability of the tools in their field. However, few studies have examined the impacts of data visualization tools in public health practice. Also, only a few studies have used formal experimental designs (e.g., randomized controlled trials) to examine the impact of data visualization on human understanding, perception, or behavior related to decision-making.

I conducted a systematic review of the available literature on the use of data visualization to assess the status of current science and evidence regarding its impact on decision-making related behavior as informed by cognitive processes such as understanding, attitude, or perception and identified key elements of appropriate study designs for testing this impact. The objective of this review was to identify 1) theories used to explain how data visualization influences behavior, 2) methods used for testing data visualization, 3) interventions used for testing data visualization's impacts, 4) primary and secondary outcomes of the effect of data visualization, and 5) mediating and moderating factors known to interfere with the impact of data visualization.

METHODS

This systematic review was conducted under the guidance of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and the Centre for Reviews and Dissemination (CRD)'s book "Systematic Reviews." [12,13] With the help of librarians that specialize in health science and data science, I conducted an electronic literature search in November 2018 using PubMed, CINAHL Plus with full text, Web of Science, IEEE Xplore, PAIS international, and Academic Search Complete. I selected the databases that are most frequently used in the health sciences. [14] Then I expanded the search to other disciplinary databases relevant to computer science and information science where data sciences are actively studied[15]. To make the search more comprehensive, I also reviewed reference lists from searched articles that could be related to my research questions.

The search strategy was developed by defining two main concepts from my research questions.[12] The two search concepts included: 1) "data visualization" as an intervention and 2) "decision-making related cognitive and behavior change" as an outcome. In addition to the two concepts, a third concept, "community and public health systems" as an intervention setting was used specifically for the PubMed and Web of Science search. For searches other than PubMed and Web of Science, including all three concepts in a database search did not generate any results because the search terms were too restrictive.

I limited my search to articles published between 2008 and 2018 (2013-2018 for the Web of Science) that were peer-reviewed, written in English, and involved studies with humans. I did not have any restrictions on factors such as education, age, or geographic location in order to maximize the chance of finding relevant articles focused on the impacts of data visualization. More details about the search strategies are provided in the Appendix A.

Inclusion & Exclusion Criteria

I included only randomized and quasi-randomized controlled trials because randomized or quasi-randomized controlled trials are the most appropriate for identifying causal effects. Articles were included if 1) they described a behavioral, social, or organizational intervention using data visualization, 2) their main aims were assessing the use of data visualization, 3) they described a method of measuring the impacts of data visualization, and 4) they clearly described the study's objectives, methods, and findings. In order to fully capture studies related to my focus, the impacts of data visualization, I also included studies that did not specifically target public health professionals from state and local health departments and instead opened the search to the general population.

Articles were excluded if 1) they related to animals or plants, 2) they were books, book chapters, magazine review articles, systematic reviews, or commentaries, 3) the full text of the articles could not be retrieved, or 4) they were unpublished research reports. There were many articles on data visualization from the fields of biochemistry, genomics, immunology, oncology, pathology, and nutrition that were specific to microscopic biomedical research. I did not include these articles because my study focused on how visual displays that used individual or organizational level data affect human decision-making. I also did not include articles that described the process of developing a visualization tool or that used visualization to describe a methodology used for statistical modeling, because the focus of my review was experimental studies that measured the impact of visualization on cognitive and behavioral changes.

For the purposes of this review, data visualization was defined as any visual display (e.g., bars, pie charts, line graphs, pictures, 3-dimensional displays, etc.) used for presenting data, information, and evidence.

Quality Assessment and Analysis

Retrieved articles were assessed for their quality and relevance. Data from each article were abstracted into an Excel spreadsheet. The data abstracted included study aims, sample sizes, study populations, theories, study designs and settings, interventions, and findings.

I assessed the articles according to the following criteria: 1) if the study designs were quasi-experimental or experiments used for evaluating the impact of data visualization tools as interventions related to cognitive and behavioral changes, 2) if the methods used to measure the impact of visualization were quantitatively reliable, 3) if the confounding factors were measured and controlled for their possible effect in the analysis, and 4) if the sampling and randomization methods were clearly stated. Narrative synthesis with the tabular presentation was used as the approach for conducting the evidence synthesis because the quantitative outcomes of selected studies were not in comparable formats. [16(p135)]

RESULTS

A total of 786 articles were retrieved from electronic database searches, in addition to three additional articles identified by reviewing reference lists of articles. Combining the results of all searches and removing duplicates (n= 19) yielded 770 articles. Next, after a review of the title and abstract of the articles based on eligibility criteria, 60 articles were retained for full-text review. Of these, 46 additional articles were excluded because 1) they focused only on the process of developing data visualization tools (e.g., appropriateness of color and size of elements of visualization), 2) they were review or commentary articles, or 3) they did not use data visualization as an intervention. A flow diagram (Figure 2.1) shows the results of the screening process and reasons for exclusion based on eligibility assessment. The remaining 14 articles were used in the final qualitative synthesis. Nine studies' settings were located in the U.S.; two studies were from Taiwan[17,18] and others were from the Netherlands, [19] Spain, [20] and Sweden. [21] Results were organized below by 1) theories used to explain how data visualization influences behavior, 2) study designs and intervention used for testing data visualization, and 3) outcome measures, including mediating and moderating factors used for their analysis of the impact of data visualization.

Theory used for Data Visualization Experiments

Most of the papers did not explicitly describe theories specifically related to data visualization in their experiments but described a specific theory related to their topic rather than about data visualization. For example, a study of Pandey and his colleagues adopted a theoretical model of persuasion, the Elaboration Likelihood Model, and used this theory to structure their experiment to measure the persuasive power of data visualization on controversial social issues. Regarding hypotheses, many studies started from the assumption that visual displays help increase understanding of the problem by integrating large amounts of information from multiple sources while reducing cognitive burden, resulting in effective and efficient information-based decision-making involving complex real-world problems. [22–25]

Study designs, Participants and Interventions used for Testing Data Visualization

Design. Descriptive information about designs and interventions of included studies is reported in Table 2.1. Most studies used a quasi-experimental or experimental design except one study that surveyed study participants regarding the usefulness of a visualization tool after it was introduced to assist them in their decision-making. [26] Comparing two groups assigned as the treatment group and the control group for treatment effects was the most common design. Three studies compared pre- and post-effects of treatment. [6,19,23] and a study followed up one week after the treatment. [19] Among experimental studies, five studies reported that they used randomization in assigning their participants, but only one study reported the method of randomization. [18] Four other studies did not report their method of assigning participants to treatment and control groups. [24,25,27,28] Only four studies collected the participants' demographic information to compare homogeneity between treatment and control groups. [17,19,21,22] The ten other studies did not report the participants' demographic information and did not include this in their final analysis.

Participants. Sample sizes varied, ranging from 24 to 720 participants. Five studies recruited college students from universities where the experiments were conducted. [17,23–25,27] Two studies recruited participants through Amazon Mechanical Turk, a crowdsourcing Internet marketplace that enables anonymous individuals to participate in and perform tasks. [6,28] The most common participants were volunteers from universities or institutions (e.g., a hospital, mental health center) where the experiments were taking place. [17,23–25,27,29] In three studies, the volunteer participants were the decision-makers whose work was closely related to the visualization tools that were tested (e.g., tracking diseased patients' location in a hospital unit). [21,30,31] In two other studies, study participants were recruited from among high school students, either in the same grade or the same classrooms who were taking the same class or from university students selected to take a class in the same degree program.[18,20] This sampling method appeared as an effort to reduce heterogeneity between experimental and control groups. Most studies did not report how they calculated the appropriate sample size. Only one study clearly stated the result of power analysis to calculate an appropriate sample size based on the study design.[32] Four studies compensated participants for completing tasks or study activities, ranging from 50 cents to $26 depending on the choice of their activities or how close they were to the optimal task result. [6,24,25,27,28]

Interventions. Interventions using data visualization varied widely across studies. Three studies compared the effects of the visually displayed information on decision-making or attitude change with the effects of textual information. [6,17,19] Three studies compared the effects of visual displays against tabular information on decision-making. [21,24,25] Two studies used an interactive visual tool that automatically sorts values on a spreadsheet as an intervention and compared how the tool influenced decision quality in comparison to a typical sort of tool. [27,28] Of the two, the study by Kim added real-life scenarios (e.g., making a decision about finding an apartment while considering its price, size, maintenance, etc.) to participants' decision-making process and asked them to consider the scenario when engaged in the task. One study also compared the effects of a decision-making learning module that contained a visualization tool, collaborative learning, and metacognitive guidance on decision-making by comparing pre- and post-test scores. [18] Four studies held pre-treatments including quiz sections or training activities for prerequisite skills before the experiments. [6,18,25,27] In Samek et al.'s study, [27] the quiz was put in place to determine participants' basic understanding of how to use the visualization tool. If the participants did not pass the quiz, they were excluded from the study. Control groups also varied across studies and included the use of textual information only, baseline treatment with tabular information, or 2-dimensional images on PowerPoint slides.

Outcome Measures

Outcome measures, as well as mediating and moderating variables from the review, are summarized in Table 2.2. Eight studies reported decision-making as a primary outcome. [18,19,21,22,24,25,27,28] Two studies reported attitude changes as a primary outcome. [6,17] The primary outcomes of the final three studies involved changes in perception of water resource problems, [33] understanding of and engagement in emergency scenarios, [30] and motivation in computer programming tasks. [20]

Decision-making. Among the studies that measured decision-making as a primary outcome, three studies asked participants to invest, allocate, or sort money using visualization aids to maximize returns. [24,25,27] The results of these studies were compared with control groups. Three other studies used either adapted scales such as the decisional conflict scale (DCS*)* to measure the decision-making process, [19,22] or developed scoring rubrics such as decision-making tests and worksheets that score decision-making subskills from 0 to 2 points. [18] One study [21] asked decision makers at administrative and clinical levels to compare hospital performance and suggest actions for hospital personnel by looking at either league tables (a widely used format of presentation) or funnel plots. All findings in this review, except one study, [22] showed statistically significant effects on decision-making in the treatment groups compared to the control groups, [19,28] baseline information, [27] tabular information only, [6,24] and pre-and post-tests. [18–20191919]

Attitude*.* Two studies reported a measure of attitude change using multi-item or single-item questionnaires as their primary outcome. [6,17] One study identified a significant difference in attitude towards healthcare services between a group that watched online healthcare advertising using a celebrity performing as a doctor showing both text and picture displays and a group that watched advertising using a celebrity performing as a patient showing text display only, mediated with perceived quality and trustworthiness of the advertisement. [17] The other study found mixed results for attitude change between a treatment and control group. [6] When the participants' initial attitudes were not strongly polarized on socially controversial topics (e.g., the relationship between corporate income tax and unemployment rates in the US), visual displays with charts had a stronger effect on the likelihood of attitude change and persuasion than displays that only used tabular information. When the participants' initial attitudes were polarized, the reverse trend was evident, that is participants shown the tabular display had a higher percentage of attitude change than participants shown the chart display.

Perception. One study reported perception change regarding water resource problems, the causes of water shortage, and solutions for a water shortage using a multi-item questionnaire with a 5-point Likert-scale. [23] This study found both unchanged and changed perceptions depending on topics related to water management in response to visual information. The study also suggested that the changes in understanding or perception could be different depending on the types of information. For example, they found more highly significant changes on the perceived severity of water problems in the 2-D group (shown 2-D PowerPoint presentation in a standard classroom) compared to the 3-D group (shown 3-D visualization).

Mediating and moderating factors. Four studies collected demographic information about the study participants. [17,19,21,22] The demographic information included gender, age, racial/ethnic background, education level, location, disease type and severity, health literacy, numeracy, professional background, and organization position. However, most studies did not include the demographic information in their analysis as a moderating factor; instead, they used moderating factors to check homogeneity between experiment and control groups. Only one study used gender as a moderating variable that strengthened the relationship of perceived quality and trustworthiness about visual advertisement content on attitude towards healthcare services. [17] Two studies found that initial attitude or prior knowledge of given topics significantly affected the strength of the relationship between the visual displays and the change in attitude or perception. [6,23] For example, in Larson and Edsall's study, prior knowledge about the water shortage problem, basic beliefs shared by social groups, and political beliefs were considered as exploratory factors, and it was found that greater knowledge of water issues was significantly associated with increased perceived water risk. [23] Only one study clearly articulated mediating factors. The study by Wang and Doong hypothesized that a treatment group that watches online healthcare advertising with both textual display and pictures would experience higher perceived quality and trustworthiness than a group that was only shown textual information. Their findings supported this hypothesis indicating the mediating effects of perceived quality and trustworthiness on attitude towards the healthcare services. [17]

DISCUSSION

In this review, I assessed theories, methodological details, and results of 14 quasi- and experimental studies that aimed to examine how data visualization impacts attitude, motivation, perception, or decision-making.

Although there were mixed results regarding the influence of data visualization on changes in cognition and behavior, systematic synthesis of the included studies also showed that data visualization interventions have a positive impact on cognitive and behavioral change such as decision-making, attitude, motivation or perception when compared to controls. [6,19,20,24,25,27,34] Visualization appears to bring advantages by increasing the amount of information delivered and decreasing the cognitive and intellectual burden of interpreting information for decision-making. [25] Negative effects were also found, but these relationships between the interventions and the outcomes appeared to be affected by moderating and mediating factors such as perceived trustworthiness and quality, domain-specific knowledge, basic beliefs shared by social groups, and political beliefs. [6,17,23] Issues associated with trust in displayed data were also discussed in a qualitative study that evaluated the design of visualization techniques for integrated health information. [35] In this study, understanding where the data came from influenced how participants interpreted visual information due to limited trust toward the visualization.

However, overall evidence from this systematic review of the impact of data visualization was limited theoretically and methodologically. Most of these experimental and quasi-experimental studies did not present theories to explain how and why visualization tools impact cognitive processes and influence decision-making behavior. A theory is “a set of interrelated concepts, definitions, and propositions that present a systematic view of events or situations by specifying relations among variables to explain and predict the events or situations.” [36,37(p9)] In social and behavioral sciences, a theory is an essential part of guiding the whole process of an intervention study by providing “insight” on how to shape research questions, methods, and intervention programs. [36,38(p27)]

Due to the diversity of the academic disciplines and topics of the included studies and the heterogeneous nature of the data visualization tools, it was difficult to find consistency in methods and outcome measures from the findings. Unlike health science behavioral research, studies that test the impact of data visualization did not have either uniform protocols or guidelines for conducting behavioral studies. The heterogeneity of the studies may be because visualization researchers have only recently realized the need for learning scientific methodology, unlike researchers from cognitive psychology or social sciences who have been trained in the scientific method. [39,40] Most studies also did not appear to report the essential elements of an experimental study, which could jeopardize the perceived validity and reliability of the studies. For example, most studies did not report how they determined sample size or how they randomized participants into experimental and control groups. This issue was raised by Few in his critique as well. [40] Few criticized a study that tested the memorability of a presented visualization by pointing out missing evidence regarding the determination of reliable sample size, resulting a critical statistical flaw. Only a few studies considered potential confounding factors including mediating and moderating factors for their analysis. Incorporating individual differences such as intelligence, cognitive ability, and technical skill is important in designing studies because considering these individual characteristics in the early stages of studies enable research to be more generalizable. [41] This systematic review suggests that gender, age, racial/ethnic background, educational level, professional background, health literacy, numeracy, and place of living, as well as initial personal attitude, value judgments, and knowledge regarding data visualization seem to be possible mediating or moderating factors that should be considered in the assessment of the impact of data visualization. [6,17–19,21–23]

Differences in academic disciplines, target populations, aims, and methods across the study settings made this review challenging. Also, there were no relevant articles found to review the impact of data visualization on the audience of interest—public health leaders who make organizational decisions for programs and services in their agencies for and with their communities. However, the evidence from this review demonstrates the multiple benefits of data visualization and shows promise for its application to public health.

Building on the current literature, future studies should be designed with consideration of relevant theories and methodologies. Theories could be used to explain how visual displays help increase understanding of complex real-world problems using visual data and information from multiple sources resulting in effective and efficient decision-making. For example, in a study that describes the needs of a user to develop a geo-visualization (GeoVis) framework, researchers emphasized incorporating cognitive fit theory into their visual geospatial displays combined with principles of public health and a human-centered approach. [42] The use of such theory, in this case, cognitive fit theory, explained how different graphical displays are related to task performance by users and guided the researchers in the design of a system that utilized the right information based on the needs of the users and the tasks they performed. [42]

Demographic characteristics specific to public health leaders should be considered for future studies. Koenig et al.[43] reported that public health practitioners who were novice users of geographic information struggled with box plots and required assistance to complete tasks. This may have been due to a lack of numerical training in public health curricula content. [43] Besides the range of demographic characteristics of individual public health leaders, researchers should understand in what contexts they use and understand data, and how that informs the designing of visualization tools. Complete descriptions of tasks regarding how public health leaders make decisions for programs and services using data, information, and data visualizations should precede the development of visualization tools. As discussed in our reviewed articles, [28,33] decisions that were aided with data visualization were made differently depending on the context of the users. Due to the nature of public health systems, public health professionals' decisions are influenced by many environmental and organizational factors such as policies, mandates, funding, organizational culture, and leadership. [7,44–48] Researchers could further investigate specific tasks that public health professionals perform and in what context their decision-making lies for assuring best public health practice when they design and test data visualizations for public health professionals.

Strengths and Limitations

This is the first identified systematic review of the impact of data visualization on human behavior. Database searches were not limited to the discipline of nursing and public health but expanded to the related disciplines such and computer science, information science, and public policy and included comprehensive searches of six relevant databases. However, only a very few experimental studies about data visualization could be found, and none were found that specifically targeted public health professionals in state and local health departments. Restrictions in the search strategies (e.g., English-language) may have led to the exclusion of relevant studies.

Even though the search focused on experimental studies using data visualization, qualitative studies or usability studies may also be important to understand the use of data visualization for public health leaders and others. Also, the assessment of the quality of study articles was performed by one reviewer and did not employ a process for validating with other reviewers and was not performed using a formal checklist or numerical scale[16(p29)] due to the heterogeneity and scarcity of eligible articles. The lack of a formal checklist imposes potential biases on the interpretation and synthesis of the review. Finally, a meta-analysis of the primary studies could not be performed due to the heterogeneity and small number of studies.

CONCLUSION

In this review, I assessed theories, methodical details, and results of 14 quasi- and experimental studies that aimed to examine the impact of interventions incorporating data visualization. Even with increasing interest in visualization evaluation studies in recent years, [49] there is still relatively little research on how visualization impacts decision-making including cognitive and behavioral processes, especially in the discipline of public health.

The evidence from this review generally suggests positive effects of data visualization depending on the control of confounding factors on behavioral changes such as attitude, perception, and decision-making. However, there is a lack of understanding regarding data visualization interventions specific to public health leaders' decision-making; in addition to which, there is little guidance for understanding participants' characteristics and tasks. A visualization study that involves public health leaders from the beginning of the design process could aid in the development of effective visualization tools and interventions that accurately depict real-world problems and support the needs of public health leaders. Future research to examine the impact of data visualization should be conducted with the appropriate methodology, specifically targeting public health leaders to improve the quality of evidence.

REFERENCES

1. Telea A. *Data Visualization: Principles and Practice*. Second edition. Boca Raton: CRC Press, Taylor & Francis Group; 2015.

2. Few S. *Now You See It : Simple Visualization Techniques for Quantitative Analysis*. 1st edition. Oakland, California: Analytics Press; 2009. http://www.tableau.com/blog/stephen-few-data-visualization. Accessed December 16, 2015.

3. Few S. What Is Data Visualization? Visual Business Intelligence. https://www.perceptualedge.com/blog/?p=2636. Published May 4, 2017. Accessed January 23, 2019.

4. Ware C. *Information Visualization Perception for Design*. 3rd ed. Amsterdam: Amsterdam; 2013.

5. Tufte E. *The Visual Display of Quantitative Information*. second edition. Cheshire, Connecticut: Graphics Press; 2001. http://www.edwardtufte.com/tufte/books_vdqi. Accessed January 20, 2016.

6. Pandey AV, Manivannan A, Nov O, Satterthwaite M, Bertini E. The Persuasive Power of Data Visualization. *IEEE Trans Vis Comput Graph*. 2014;20(12):2211-2220. doi:10.1109/TVCG.2014.2346419

7. Brownson RC, Fielding JE, Maylahn CM. Evidence-Based Public Health: A Fundamental Concept for Public Health Practice. *Annu Rev Public Health*. 2009;30(1):175-201. doi:10.1146/annurev.publhealth.031308.100134

8. Brownson RC, Gurney JG, Land GH. Evidence-based decision making in public health. *J Public Health Manag Pract*. 1999;5(5):86–97.

9. Jenicek M. Epidemiology, evidenced-based medicine, and evidence-based public health. *J Epidemiol*. 1997;7(4):187–197.

10. Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. *Evidence-Based Clinical and Public Health: Generating and Applying the Evidence*. Office of Disease Prevention and Health Promotion; 2010. https://www.healthypeople.gov/sites/default/files/EvidenceBasedClinicalPH2010.pdf.

11. National Association of County & City Health Officials. *2013 National Profile of Local Health Departments*.; 2013. http: //www.naccho.org/topics/infrastructure/profile/upload/ 2013-National-Profile-of-Local-Health-Departments-report. pdf. Accessed December 8, 2015.

12. Centre for Reviews and Dissemination, ed. *Systematic Review: CRD's Guidance for Undertaking Reviews in Health Care*. 3. ed. York: York Publ. Services; 2009.

13. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–269.

14. University of Washington University Libraries. HSL - Database Searching - Library Guides at University of Washington Libraries. Library guides. http://guides.lib.uw.edu/hsl/database-searching. Accessed January 24, 2019.

15. University of Washington Engineering Library. Top engineering databases. Engineering Library. http://www.lib.washington.edu/engineering. Accessed January 24, 2019.

16. Booth A, Papaioannou D, Sutton A. *Systematic Approaches to a Successful Literature Review*. 1st ed. SAGE Publications Ltd; 2012. https://books.google.com/books/about/Systematic_Approaches_to_a_Successful_Li.html?id=MyI5uqU_y1QC. Accessed February 24, 2016.

17. Wang H-C, Doong H-S. Effects of Online Advertising Strategy on Attitude towards Healthcare Service. In: IEEE; 2014:2725-2732. doi:10.1109/HICSS.2014.342

18. Hsu Y-S, Lin S-S. Prompting students to make socioscientific decisions: embedding metacognitive guidance in an e-learning environment. *Int J Sci Educ*. 2017;39(7):964-979. doi:10.1080/09500693.2017.1312036

19. Westermann GMA, Verheij F, Winkens B, Verhulst FC, Van Oort FVA. Structured shared decision-making using dialogue and visualization: a randomized controlled trial. *Patient Educ Couns*. 2013;90(1):74-81. doi:10.1016/j.pec.2012.09.014

20. Velázquez-Iturbide Á, Hernan-Losada I, Paredes-Velasco M. Evaluating the Effect of Program Visualization on Student Motivation. *IEEE Trans Educ*. 2017;60(3):238-245. doi:10.1109/TE.2017.2648781

21. Anell A, Hagberg O, Liedberg F, Ryden S. A randomized comparison between league tables and funnel plots to inform health care decision-making. *Int J Qual Health Care*. October 2016. doi:10.1093/intqhc/mzw125

22. Dolan JG, Veazie PJ, Russ AJ. Development and initial evaluation of a treatment decision dashboard. *BMC Med Inform Decis Mak*. 2013;13:51. doi:10.1186/1472-6947-13-51

23. Larson KL, Edsall RM. The impact of visual information on perceptions of water resource problems and management alternatives. *J Environ Plan Manag*. 2010;53(3):335-352. doi:10.1080/09640561003613021

24. Rudolph S, Savikhin A, Ebert DS. FinVis: Applied visual analytics for personal financial planning. In: *IEEE Symposium on Visual Analytics Science and Technology*. IEEE; 2009:195–202. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5333920. Accessed December 20, 2015.

25. Savikhin A, Maciejewski R, Ebert DS. Applied visual analytics for economic decision-making. In: *The IEEE Symposium on Visual Analytics Science and Technology*. IEEE; 2008:107–114. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4677363. Accessed January 11, 2016.

26. Araz OM, John M, Lant T, Fowler JW. A new method of exercising pandemic preparedness through an interactive simulation and visualization. *J Med Syst*. 2012;36(3):1475-1483. doi:10.1007/s10916-010-9608-7

27. Samek A, Hur I, Kim S-H, Yi JS. An Experimental Study of Decision Process with Interactive Technology. *CESR-Schaeffer Work Pap*. September 2015. doi:http://dx.doi.org/10.2139/ssrn.2347698

28. Kim S-H. The Effectiveness of Interactive Visualizations for Multi-Attribute Decision Making with Contextual Data. *J Ergon Soc Korea*.:11.

29. Kettelhut VV, Vanschooneveld TC, McClay JC, Mercer DF, Fruhling A, Meza JL. Empirical Study on the Impact of a Tactical Biosurveillance Information Visualization on Users' Situational Awareness. *Mil Med*. 2017;182(S1):322-329. doi:10.7205/MILMED-D-16-00143

30. Araz OM, John M, Lant T, Fowler JW. A New Method of Exercising Pandemic Preparedness Through an Interactive Simulation and Visualization. *J Med Syst*. 2012;36(3):1475-1483. doi:10.1007/s10916-010-9608-7

31. Dolan JG, Veazie PJ, Russ AJ. Development and initial evaluation of a treatment decision dashboard. *BMC Med Inform Decis Mak*. 2013;13(1). doi:10.1186/1472-6947-13-51

32. Westermann GMA, Verheij F, Winkens B, Verhulst FC, Van Oort FVA. Structured shared decision-making using dialogue and visualization: A randomized controlled trial. *Patient Educ Couns*. 2013;90(1):74-81. doi:10.1016/j.pec.2012.09.014

33. Larson KL, Edsall RM. The impact of visual information on perceptions of water resource problems and management alternatives. *J Environ Plan Manag*. 2010;53(3):335-352. doi:10.1080/09640561003613021

34. Hui-Chih Wang, Her-Sen Doong. Effects of Online Advertising Strategy on Attitude towards Healthcare Service. In: *System Sciences (HICSS), 2014 47th Hawaii International Conference On*. ; 2014:2725-2732. doi:10.1109/HICSS.2014.342

35. Le T, Reeder B, Thompson H, Demiris G. Health providers' perceptions of novel approaches to visualizing integrated health information. *Methods Inf Med*. 2013;52(3):250-258. doi:10.3414/ME12-01-0073

36. Glanz K, Bishop DB. The Role of Behavioral Science Theory in Development and Implementation of Public Health Interventions. *Annu Rev Public Health*. 2010;31(1):399-418. doi:10.1146/annurev.publhealth.012809.103604

37. Kerlinger FN. *Foundations of Behavioral Research*. 3rd ed. New York: Holt, Rinehart, and Winston; 1986. https://books.google.com/books/about/Foundations_of_behavioral_research.html?id=PfpGAAAAMAAJ. Accessed February 28, 2016.

38. Glanz K, Rimer BK, Viswanath K, eds. *Health Behavior and Health Education: Theory, Research, and Practice*. 4th ed. San Francisco, CA: Jossey-Bass; 2008.

39. Kim S-H, Yi JS, Elmqvist N. Oopsy-daisy: failure stories in quantitative evaluation studies for visualizations. In: ACM Press; 2014:142-146. doi:10.1145/2669557.2669576

40. Few S. Information Visualization Research as Pseudo-Science. 2015:9.

41. Dillon A, Watson C. User analysis in HCI—the historical lessons from individual differences research. *Int J Hum-Comput Stud*. 1996;45(6):619–637.

42. Joshi A, Novaes M de A, Machiavelli J, et al. A Human-Centered GeoVisualization framework to facilitate visual exploration of telehealth data: A case study. *Technol Health Care*. 2012;20(6):487-501. doi:10.3233/THC-2012-0683

43. Koenig A, Samarasundera E, Cheng T. Interactive map communication: Pilot study of the visual perceptions and preferences of public health practitioners. *Public Health*. 2011;125(8):554-560. doi:10.1016/j.puhe.2011.02.011

44. Bekemeier B, Chen AL, Kawakyu N. Mandated activities and limited decision-making authority among local public health officials. *Front Public Health Serv Syst Res*. 2012;1(3):4.

45. Bekemeier B, Chen A, Kawakyu N, Yang Y. Local Public Health Resource Allocation. *Am J Prev Med*. 2013;45(6):769-775. doi:10.1016/j.amepre.2013.08.009

46. Fields RP, Stamatakis KA, Duggan K, Brownson RC. Importance of Scientific Resources Among Local Public Health Practitioners. *Am J Public Health*. 2015;105(S2): S288–S294.

47. LaPelle NR, Dahlen K, Gabella BA, Juhl AL, Martin, DA E. Overcoming Inertia: Increasing Public Health Departments' Access to Evidence-Based Information and Promoting Usage to Inform Practice. *Am J Public Health*. 2014;104(1):77-79. doi:10.2105/AJPH.2013.301404

48. Sosnowy CD, Weiss LJ, Maylahn CM, Pirani SJ, Katagiri NJ. Factors Affecting Evidence-Based Decision Making in Local Health Departments. *Am J Prev Med*. 2013;45(6):763-768. doi:10.1016/j.amepre.2013.08.004

49. Isenberg T, Isenberg P, Chen J, Sedlmair M, Moller T. A Systematic Review on the Practice of Evaluating Visualization. *IEEE Trans Vis Comput Graph*. 2013;19(12):2818-2827. doi:10.1109/TVCG.2013.126

Table 2.1. Study designs and intervention characteristics in a systematic review of data visualization intervention

Study	Designs	Intervention; dosage	Control Group
Araz et al.	survey design (n=32)	Arizona State University (ASU) Pandemic Influenza Tabletop Exercise, which included video clips with scenario information, geographical mapping, and the interactive computer simulation model. Asked to respond to a hypothetical pandemic influenza scenario and make iterative policy decisions in a group setting. The exercise was supplemented with an interactive simulation model for decision support. This model simulated the possible outcomes of decisions as they were being made based on different rates of disease transmission, case fatality, the timing of university evacuation decisions, and social distancing interventions. Four scenarios; each of 30 min. Compensation: not reported	No control group
Dolan et al.	feasibility pilot (n=25)	1. Brief introduction 2. Participants individually reviewed the dashboard running on a personal computer equipped with a touch screen monitor in a private setting. The patient decision aids were developed as a dashboard, consisting of five windows that summarize the relative performance of the treatment alternatives about each of the included drug information categories: effectiveness in relieving pain, risk of side effects, the possibility of drug-drug interactions, out-of-pocket cost, and how the drug is administered. time spent was measured as part of usability Compensation: not reported	No control group
Wang & Doong	experiment (n=104); Randomized as Tx group (n=51) Control group (n=53)	watch an online advertising video (online information display of text with picture about healthcare service) and fill out the questionnaire relevant to this advertising video afterward. 1 session: 10 min (video: 3min) compensation: not reported	online information with text only

Rudolph et al.	experiment (n=26); Tx group (n=13) Control (n=13) (no randomization)	Visual analytics tool (FinVis) that allows the non-expert casual user to interpret the return, risk, and correlation aspects of financial data and make personal finance decisions. This interactive exploratory tool helps the casual decision-maker quickly choose between various financial portfolio options and view possible outcomes. Seven sessions; duration of sessions was not reported compensation: $9-18	baseline treatment using a tabular display with the same information
Larson & Edsall	quasi-experiment (n=76); Randomized as Tx group 1 (n=35) Tx group 2 (n=41)	3D demonstration in Arizona State University's (ASU) Decision Theater (DT): a semi-immersive visualization environment featuring animated graphics that surround viewers on high-resolution large-screen displays. On seven walls of the 10-sided room are 10x8 foot screens capable of rendering computer displays in stereo, with three-dimensional viewing through polarized stereo '3D' glasses. The visual images in the DT primarily portrayed the physical depth of the underground aquifer. Images on a peripheral screen - the 'digital dashboard'- consisted of a series of small supplemental graphs and legends to guide the presenter. The same person narrated, live, both presentations from a standard script. 10 min presentations addressed excessive groundwater withdrawals in the region while illustrating water levels and flow to pinpoint cone of depression problems. compensation: not reported	2D PowerPoint presentation in a standard classroom: visual images in PowerPoint were designed to replicate the DT imagery as close as possible. Where two large images moved across the DT's multiple screens, the still images were captured and presented side-by-side in a PowerPoint slide to support a similar comparison. For PowerPoint presentation, small dashboard images were in the lower corner of the slides. Although the 2D presentations included 2.5D images, stereo glasses, and the large multiple screen environment of the DT were not used in the 2D setting. It is projected on a standard classroom screen, included smaller images in a narrower field of vision.

Westermann et al.	experiment (n=91); Randomized as Tx group (n=51) Control group (n=45)	Counseling in Dialogue (CD): Visualization aid was used for creating shared understandable and meaningful narrative of diagnostic formulation and treatment options for parents of a child of mental health problem. Both CD and CU; 1 session, 60min compensation; not reported	Care as a usual (CU); parents are counseled in an equally long counseling session as in CD. However, the therapist was free to select the topics to be discussed, the structure of the counseling, and the degree of shared decision-making.
Samek et al.	experiment (n=120) Tx group 1 Tx group 2 Control group (randomization not reported)	Written instructions before the experiment, short quiz on understanding Subjects were asked to make a decision and were incentivized to select an object (initial object) immediately, and continue improving on their selection (intermediate object) until they were satisfied with their selection (final object) using three different treatments; Baseline, Typical Sort, and Automated Sort. Typical Sort: subjects could sort objects by one attribute at a time by clicking on the column header Automated Sort; columns of each attribute were always sorted by value automatically in descending order and objects could be compared using visual highlighting in each attribute column One session; 140min (20 rounds, 180 seconds for each round including instruction time) compensation; based on the selection of the final object as instructed average earning $26	Baseline; limited interactive capability to select and mouse-over objects

Pandey et al.	experiment (n=720); Randomized as Topic 1; Tx group (n=54) Control group (n=45) Topic 2; Tx group (n=50) Control group (n=56) Topic 3; Tx group (n=55) Control group (n=63)	Nine stages that participants undergo during the experiment such as demographic information, pre-treatment questions, persuasive message (context + evidence + presentation with graphical evidence (charts), post-treatment questions, need for cognition scale) One treatment; 5-10 min compensation: $ 0.50	persuasive message by textual evidence (tables)
Savikhin et al.	experiment (n=56); Tx group 1 (n=23) Tx group 2 (n=14) Control group (n=19) (randomization not reported)	Pre-experiment quiz and training on how to read the information screen. Experiment: both the interactive visual analytics program and the simple visual representation use 2D graphics to show the profit the subject would have received at every possible company value at his/her bid amount after the outcome of the period was determined. allows exploring the dataset on their own time before making a decision compensation: $5-25	table display
Anell, A., Hagberg, O., Liedberg, F., & Ryden, S.	Web-based survey Randomized as Tx group 1 (n=113) Control group (n=108)	Data presentations in the form of league tables or funnel plots. Decision-makers compare hospital performance using either league tables or funnel plots for four different measures within the area of cancer care. For each measure, decision-makers were asked to suggest actions towards 12–16 hospitals (no action, ask for more information, intervene) and provide feedback related to whether the information provided had been useful. Compensation: not reported	data presentations in the form of League tables

Hsu, Y.-S., & Lin, S.-S.	experiment (n=74); pre- and post-tests and Randomized as Tx group 1 (n=42) Control group (n=32)	Decision-making learning module consisted of three activities as a small group discussion; 1) understanding the concept of evidence (reliability and validity of evidence), 2) Decision-making strategies, and 3) decision-making activities with/without embedded metacognitive guidance One learning module; 7 hours Compensation; not reported	no metacognitive guidance
Velazquez-Iturbide, J. A., Hernan-Losada, I., & Paredes-Velasco, M.	experiment (n=31); pre- and post-tests and Randomized as Tx group 1 (n=16) Control group (n=15)	A program visualization system, Srec that displays recursive processes coded in Java that is adopted to solve the class assignment. One session; 2 hours	any Java programming tools the students had mastered (typically, BlueJ or Eclipse)
Kettelhut, V. V., Vanschooneveld, T. C., McClay, J. C., Mercer, D. F., Fruhling, A., & Meza, J. L.	quasi-experimental survey (n=19) with pre- and post design	The VIZ that visualizes hospital spatial data linked to individual patients' clinical-laboratory data for tracking diseased patients' locations antibiotic administrations, contact data, and infection prevention intervention data One session; 30 minutes	current EHR-based data
Kim S-H	experiment (n=127) Tx group (n=67) Control group (n=60) (randomization not reported)	SimulSort for the interactive visualization interface. The participants had three trials with different data sets: apartment, laptop, and printer. For each trial, a scenario was given. compensation; average earning $2.90	Typical Sorting for the non-visual traditional interface

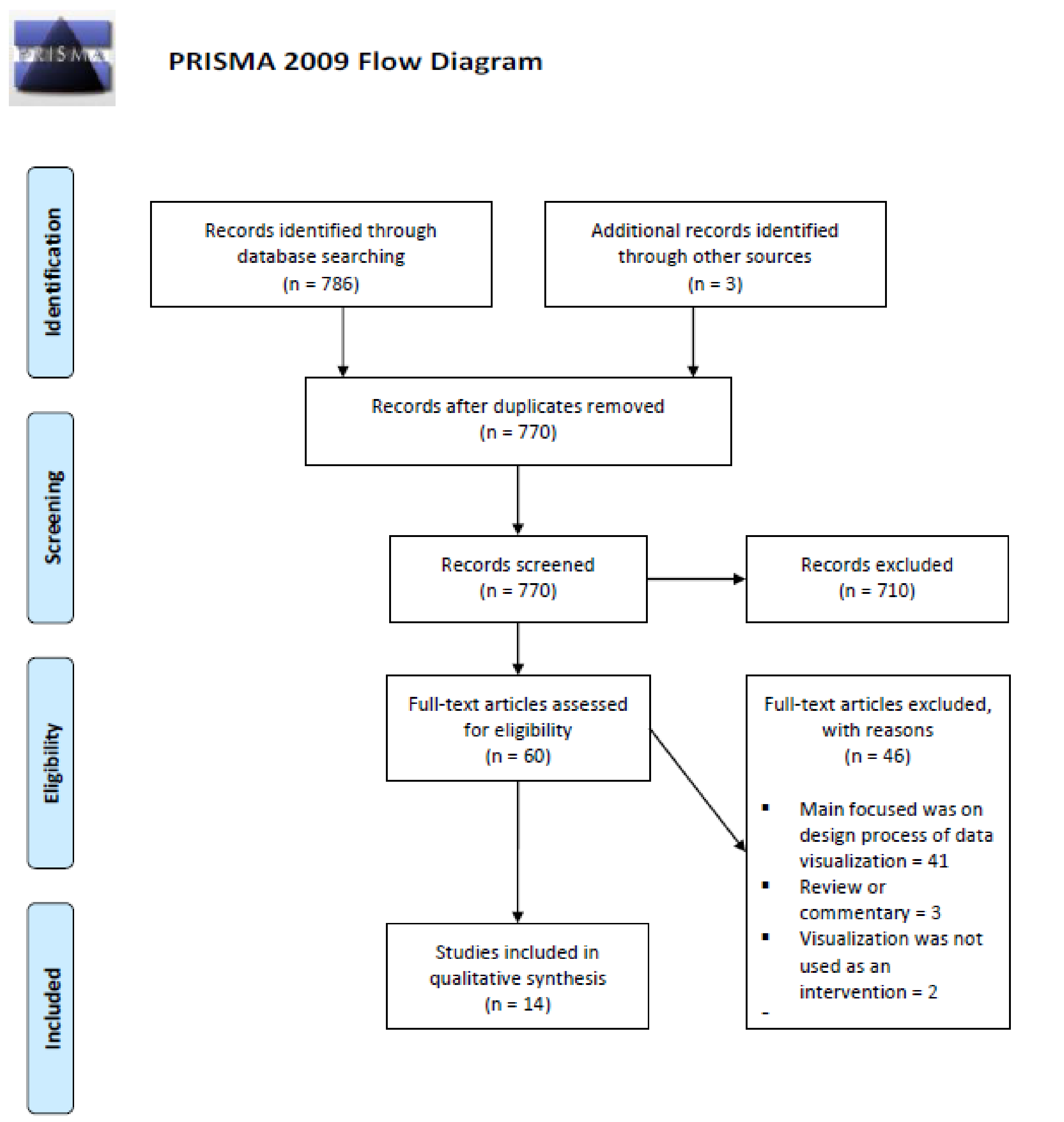
Table 2.2. Outcome measures, moderating and mediating variables

Study	Study Outcomes	Mediating Variables	Moderating Variables
Araz et al.	Help understanding and engaging in the emergency scenarios	Not reported	Not reported
Dolan et al.	Usability: Observation data (time, drug choice, use of optional features in the dashboard), ease of use and acceptability, mechanical ease of use, cognitive ease of use, and decision-aiding effectiveness Experiment: decision-making processes: the informed, values clarity, and uncertainty sub-scales (decisional conflict scale, nine items with 5-point scale) Follow-up interview: experiences of using dashboard and comparison with conventional information	Not reported	Gender, age, racial/ethnic background, highest educational level, health literacy, numeracy
Wang & Doong	Attitudes toward health care service (4 items with 4-point scale)	Perceived quality, Perceived Trustworthiness	Gender, age
Rudolph et al.	Decision-making (achieved a financial outcome) exploration and learning behavior (time spent) confidence and perception of usefulness (8-point scale)	Not reported	Not reported
Larson & Edsall	Perceptions of the magnitude of water resource problems and the causes of and solutions to water shortage (14 items with 5-point ordinal scales)	Prior knowledge (3 items with 5-point scale), Ecological worldviews; human control over nature, the fragility of the earth (6 items with 5-point scale), Political beliefs; free market economy and private property rights, high values indicating conservative beliefs (2 items with 5-point scale)	Gender, age, race/ethnicity, location

Westermann et al.	Primary outcome: Decisional conflicts; uncertainty, factors contributing to uncertainty, and perceived effective decision-making (Decisional conflict scale, 16 items with 5-point Likert scale) Secondary outcome: Decision made upon accurate information, satisfaction with shared decision-making, consensus diagnostic formulation, consensus recommended treatment, and acceptance recommended treatment (single item, 4-point scale)	Not reported	Gender, age, types of mental disease diagnosis, disease severity, education level of parents
Samek et al.	Decision-making process (achieved outcome value)	Not reported	Not reported
Pandey et al.	Persuasion likelihood (ratio of the number of persuaded participants and the total number of participants) Attitude change (single-item scales, changes in user's self-reported attitude)	Initial attitude before treatment (neutral or negatively polarized), degree of elaboration; involvement and need for cognition	Not reported
Savikhin et al.	Decision making achieved financial outcome value)	Not reported	Not reported
Anell, A., Hagberg, O., Liedberg, F., & Ryden, S.	Decision-making; Number of actions suggested by participants, the proportion of appropriate actions	Not reported	Age, gender, professional background, organizational position
Hsu, Y.-S., & Lin, S.-S.	Decision-making test and worksheets using scoring rubric (four decision-making subskills scored from 0 to 2 points each)	Not reported	Not reported

Velazquez-Iturbide, J. A., Hernan-Losada, I., & Paredes-Velasco, M.	Students' motivation; the EMSI questionnaire was used to measure students' four dimensions of motivation	Not reported	Knowledge level: checked by grades obtained in previous assignments; the participants were fourth-year computer science major students who elect to take a course on advanced algorithms.
Kettelhut, V. V., Vanschooneveld, T. C., McClay, J. C., Mercer, D. F., Fruhling, A., & Meza, J. L.	Subjects' perception (Level 1 SA) and comprehension (Level 2 SA) of the infection transmission risk factors with the use of the VIZ vs. the current EHR-based data.	Not reported	Unit-based staff vs. non-unit-based staff (clinical consultants and infection preventionists) The novice (1–5 years in health care) vs. the experienced (>5 years in health care).
Kim S-H	Decision quality (whether dominated or nondominated option) and confidence level of the decision (a 7-point Likert scale), time spent making a decision	Not reported	Not reported

Figure 2.1. Flow diagram of included studies in a systematic review of data visualization interventions



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

CHAPTER 3. Understanding data use and preference of data visualization for public health practitioners: A qualitative study

ABSTRACT

Background: Data visualization has great potential to improve access to and understanding of data for decision-making among public health professionals. However, little is known about their need for and use of data and information and their preferences regarding data visualization. This study aims to 1) assess how public health practitioners use data, information, and evidence, 2) understand perceptions of, and preferences regarding, data visualization to inform future design of a data visualization tool; and 3) examine how to further public health practitioners' engagement with the use of data for decision-making through data visualization.

Methods: Semi-structured interviews were conducted with 14 public health professionals from state and local health departments across five states. Interview transcripts were coded using a thematic analysis approach.

Results: Four themes were derived from the analysis: 1) collection of data, information, and evidence, 2) approaches to understanding data and information, 3) data for decision-making and communication, and 4) preference for data visualization and how it is being used. The analysis also showed associated challenges public health professional face in each step or combination of steps of data use processes.

Conclusion: Public health professionals believe in the great value of data, information and evidence for 1) understanding current public health issues, informing their decisions regarding their programs and services, 2) communicating with stakeholders for policy development and funding allocation and 3) educating the public. Yet there were profound challenges and a desire to use data much more and better. Data visualization tools can be used to improve understanding, communication, educating, and decision-making.

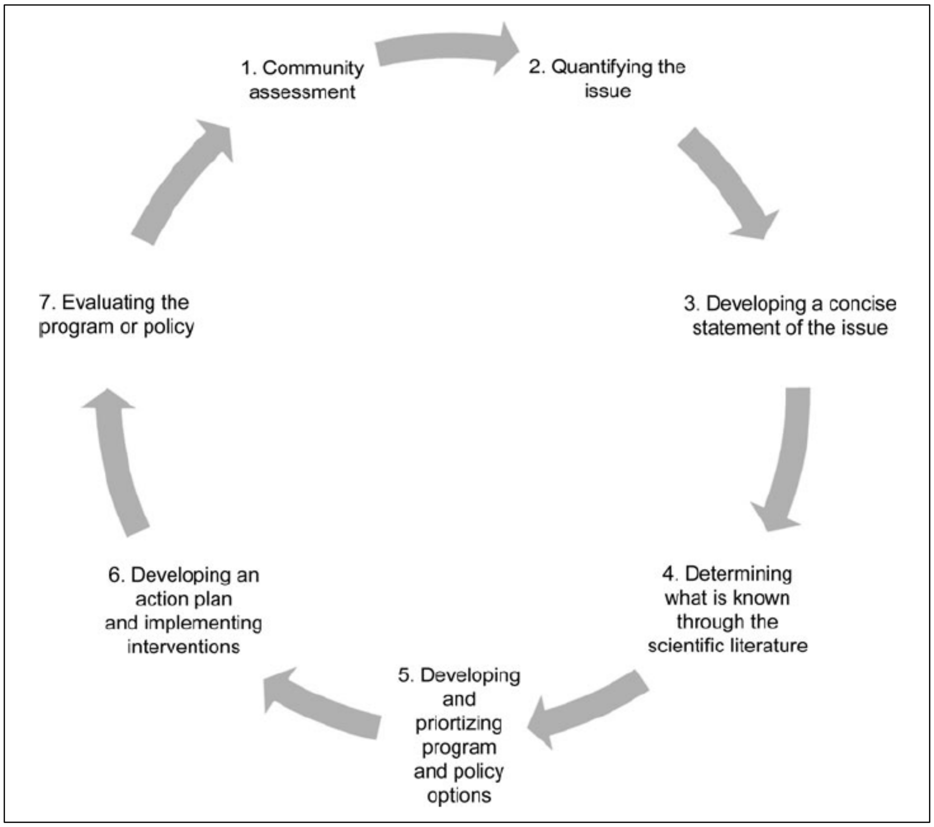
BACKGROUND

Public health leaders in state and local health departments are encouraged to make decisions for resource allocation and public health planning based on up-to-date and reliable data, information, and evidence.[1–3] Such Evidence-Based Public Health (EBPH) involves an "iterative" and "nonlinear" cycle of steps, including understanding public health issues based on data and literature and then developing, implementing, and evaluating programs and policies (Figure 3.1).[1] A specific set of competencies are required to implement EBPH such as "quantifying the issues" and "prioritizing health issues."[1]

However, there are barriers that often prevent public health leaders from using evidence as a means to drive decision-making for their public health planning. These include political and policy influence,[4] legal mandates and regulation,[5,6] limited and inflexible funding and rigid bureaucratic government systems and leadership that hamper change, as well as inadequate incentives and expertise within the workforce for effectively exploring and using data.[1,4–7] Moreover, tremendous funding cuts have undermined or led to the elimination of many local public health programs.[8] Such widespread funding cuts put public health leaders in a particularly tenuous position for systematically making decisions to effectively allocate their resources regarding services and programs.[8]

With the rise of data science, we have new opportunities for answering questions that were previously much harder to answer, such as why certain populations in some areas are healthier than those in other areas, or what socioeconomic factors impact health. A study by Ryan et al. presented a data-driven decision-making tool that was used to aid decisions on how to allocate funds for HIV/AIDS programs at a local health department in the US.[11] The study found that the data-driven decision-making tool revealed better intervention strategies and could be helpful in defending their programs in case of political restraint.[11]

Figure 3.1. Evidence-based public health framework by Brownson et al.[1]



A powerful means of rendering the results of such data science more useful is data visualization tools, which can facilitate these processes by supporting access to and improving our understanding of data. Data visualization refers to visual representations used to support understanding, communication, decision-making with data and information[12–15] as well as "amplifying cognition."[16] By looking at underlying patterns of complex public health problems that are informed by insights gained from data and information that are presented using visualization tools, researchers are better able to generate evidence with meaningful data and public health leaders can more effectively select programs and services that are well-matched to a community's needs, preferences, and organizational capacity[12] .

There are numerous studies regarding data visualization from multiple disciplines such as mathematics, statistics, computer science, information science, psychology, cognitive science, and engineering.[11] However, there are very few studies of data visualization used in public health, related to how the tools help develop a "deeper level of understanding" of data and information,[11,12] especially among public health leaders (S. Park's systematic review paper). Moreover, little is known about what the needs and preferences are of public health practitioners for visual data regarding local public health issues.

I used qualitative data to 1) assess how public health practitioners use data, information, and evidence, 2) understand perceptions of and preferences regarding data visualization to inform future design of a data visualization tool and 3) examine how to further public health practitioners' engagement with and use of data for decision-making through data visualization. This study was conducted to fill an important gap in understanding how improvements can be made to public health practitioners' use of data, information, and evidence.

METHODS

Data Collection

Semi-structured interviews were conducted via telephone to assess the needs of public health practitioners for visual displays of data regarding local public health issues. The interviews were a part of the Robert Wood Johnson Foundation- (RWJF) funded Public Health Activities and Services Tracking (PHAST) Study (PI. B. Bekemeier). Participants were recruited by emailing personnel who participated in one of the PHAST Study team's recent projects

(RWJF Grant #73270) (Appendix B.) as well as by advertising through an online newsletter distributed by the Northwest Center for Public Health Practice at the University of Washington's School of Public Health. Inclusion criteria for participants were that they had to be public health leaders working at a state or local health department who were in a position involving data use for analysis or decision-making as part of community health assessment and program planning. A total of 14 public health professionals were interviewed individually from April to September 2016. Detailed characteristics of the interview participants are presented in Table 3.1.

The interview questions used are specified in Appendix C. Participants were asked to describe their work as it involves decision-making, what kind of data and information they use, how they use it, and their challenges related to using data, information, and evidence throughout the decision-making process. Also, the interviewers asked about participants' specific preferences regarding data visualization for their work and their current use of data visualization.

Table 3.1. Characteristics of the interview participants

Characteristics	Category	Participants (n=14)
States (n)	Washington	9
	Georgia	1
	Ohio	1
	Minnesota	1
	New York	1
	Wisconsin	1
Level of Jurisdictions (n)	State level public health department	3
	County-level public health department	11
Positions (n)	Director/assistant director/administrator/health officer/health commissioner	4
	Epidemiologist	3
	Community assessment manager/coordinator	2
	Researcher	2
	Program director	2

The duration of the interviews was between 30-45 minutes. Interviews were recorded, and handwritten notes were taken for all interviews. Interviews were conducted with a note-taker and at least one interviewer. The lead author participated in each interview and asked supplementary follow-up questions or probes as needed. Institutional review board approval was granted for this study by the Human Subjects Division, University of Washington (IRB Number: 51569).

Data Analysis

I conducted a thematic analysis of the transcripts and identified themes. Thematic analysis is an appropriate approach to understand participants' experiences and contexts without being confined to a theoretical framework.[13(p9)] The recorded interviews were transcribed verbatim and imported into Atlas.ti. Software program version 7. The lead author read the interview transcripts several times to become familiar with the interview content and conducted coding based on identified recurring themes. To check the inter-rater reliability of our analysis, a data analyst who was present during the interviews acted as an additional coder for a sample of the interviews. Additional sample transcript coding was also performed by a Ph.D. student in Information Science. Inter-rater reliability was checked indicating a moderately satisfactory level of reliability (Kappa: 0.8) using the Coding Analysis Toolkit (CAT). [19] After determining inter-rater reliability, coding for the rest of the transcripts was finalized by this study's lead investigator (S. Park).

RESULTS

I identified four themes from the analysis: 1) collection of data, information, and evidence, 2) approaches to understanding data and information, 3) data for decision-making and communication, and 4) preference for data visualization and how it is being used (Table 3.2). The analysis also showed associated challenges they face in each step, or combination of steps, of data use processes (Table 3.2).

Collection of Data, Information, and Evidence

Participants described how they collect data, information, or evidence for their work. Even though I specifically asked about "data," the responses were not exclusively limited to data; participants used the terms data, information, and evidence inter-changeably or in combination. As a result, what the participants call data collection I have identified as "collection of data, information, and evidence." Broadly speaking, participants described two ways of collecting data, information, and evidence: informal and formal sources.

Before or in addition to looking at specific figures in data, administrative information, and evidence, they often seek out information informally from peers or other local government organizations such as law enforcement and environmental departments. For example, one participant expressed how sometimes informal communication can be helpful in their decision-making. The participant noted, "We started looking at death data from an (opioid) overdose. We saw a slight uptick but not great. But when we started talking with our police force, they were having more and more Narcan, naloxone, and we really used that and that helped drive our decision to start a Naloxone program for family members to come here and get it at no cost, be trained."

Table 3.2. Data Use Processes for Decision-Making among Public Health Practice

<table>
<tr><th>Theme</th><th>Domains</th><th>Challenges</th></tr>
<tr><td>Collection of data, information, and evidence</td><td>• Informal sources
▪ Hearing from peers
• Formal sources
▪ Lab reports
▪ Surveys, e.g. satisfaction survey from patients and customers
▪ Hospitals, clinics
▪ Schools, communities, e.g., data sharing contract with schools or communities
▪ State and federal government data portals
▪ Peer-reviewed journals</td><td>• Data quality issues
• Scattered data sources
• Limited access to data
• Data related to emergent issues (opioids, tattoo, e-cigarette)
• Lack of epidemiologist or IT experts
• Old and inconsistent information systems</td></tr>
<tr><td>Approaches to understanding data and information</td><td>• Making comparisons with other counties, their state, the nation, or themselves over time by combining with ancillary datasets (e.g., population size, poverty level, demographic information, etc.)
• Data are usually managed by data experts or epidemiologists where there are appropriate and enough staff in their jurisdictions</td><td>• Difficulty in combining multiple data sources due to the heterogeneity of data collection and measurement system
• Staff desire for having access to granular (sub-county) data
• Lack of epidemiologist or IT experts</td></tr>
<tr><td rowspan="3">Data for Decision-Making and Communication</td><td>Audience: Agency staff desire data access for
• Understanding the needs of community, e.g. community health assessment
• Quality and customer satisfaction for program improvement
• Setting priorities for program planning and program evaluation</td><td rowspan="3">• Need effective ways to communicate
• Lack of epidemiologist or IT experts</td></tr>
<tr><td>Audience: Staff use data with stakeholders and policymakers for
• Building a story using data
• Justifying funding
• Policy development
• Educating about public health issues</td></tr>
<tr><td>Audience: Staff use data with the Public for
• Communicating and educating the public about health issues and prevention</td></tr>
</table>

Most participants, however, described more formal ways to collect data, information, and evidence depending on the data sources and methods they use. The sources of data were varied and included lab reports, hospital data, surveys, state and governmental data, and peer-reviewed journals. Data were collected in many different ways. Primary data were collected directly from patients or customers who were served by their agencies. Secondary data were gained from hospitals, clinics, or schools or were found in publicly shared governmental data portals. They also noted that they search peer-reviewed literature in combination with local data.

Participants reported that they frequently used peer-reviewed literature to inform their decision makers (e.g., Board of Health) because evidence from literature already produced by experts saves time and effort compared with conducting original research. However, they also emphasized that using primary data that they analyzed was extremely informative for decision-makers especially when it involved trendy health issues such as "e-cigarettes."

Profound challenges were also noted related to the process of collecting data, information, and evidence. Several described that data sources are scattered, and even after the data were obtained it was not easy to use. One participant noted, "You can just get lost because you are going to so many different places and you might be getting conflicting numbers from different places." Also, even after they have gathered the data they need, they described facing difficulties in linking and analyzing datasets due to limited staff who have the relevant skills or knowledge to do so. Another challenge was that their data reporting system or data registry was often old and inconsistent across jurisdictions and over time.

In many cases, the registry was not mandatory resulting in incomplete datasets that caused them to question whether the numbers they saw reflected reality, or if they were due to reporting failures. Other challenges addressed included limited access to data, data that were old or not timely, and a simple lack of desired data. Finally, participants described their data reporting systems as not keeping up with newly emerging issues that need to be addressed by decision-makers (e.g., opioids and e-cigarettes).

Approaches to Understanding Data and Information

Most participants saw data, when available, as a valuable means for comparing themselves with other counties, their state, the nation, or themselves over time. Most practitioners expressed a desire to see what was going on in their county compared to others as well as what other community groups were working on so that they could better consider if they could or should adopt a similar strategy.

However, participants noted that comparisons should be used with caution because the performance of local public health systems often depends on their funding, size of staff and many other factors (e.g., the demography of population they serve). Practitioners also preferred to see rates or ratios rather than raw or total numbers of cases. If possible, participants wanted to see data combined with ancillary datasets to compare areas of similar geographical size, population size, poverty level, or by rural versus urban areas. For these kinds of comparisons, maps were expressed as a preference by participants because they were useful for understanding data. One participant noted "I love it when they have an interactive map. …And you can click on your state, and then your county…. you can zoom in and zoom out, and …get info by hovering your mouse over." Comparisons were also done or desired in which one examined the effectiveness of interventions. One practitioner noted, "We might work with school districts to collect BMI (Body Mass Index) data, you know, direct height and weight of kids because we have given them a mini-grant. … So, we use that kind of pre- and post to look at how their interventions are with how they might affect decreased BMIs." Participants representing local public health departments particularly stressed the need for more county-level data and even sub-county level data as being more important than regional or state level data. These local participants were more interested in understanding public health issues in their neighborhoods because the context of health issues is specific to the areas and neighborhood in which they are working.

However, county level or sub-county level data are often not available, compared to state or national-level data. Even when county-level data such as Pregnancy Risk Assessment Monitoring System (PRAMS) or Behavioral Risk Factor Surveillance System (BRFSS) are available, sample sizes of those survey datasets in small counties are too small to see generalizable trends and patterns. Other challenges were identified related to inconsistent data measurement and reporting systems over counties and states, or over time. Participants described, for example, that some counties were developing their scales to measure their public health performance for program planning where a standardized public health measurement system was absent across their jurisdictions. Moreover, they found it challenging to combine multiple data sources with exploring health outcomes of populations from different socioeconomic backgrounds because data are scattered in different sources and data collection methods, measurements and data structures are not consistent across jurisdictions and surveys.

Related to the problem of the sheer lack of desired data in public health, many participants expressed difficulty with understanding public health issues in the context of advocating for policies. A public health practitioner from a state health department, who had struggled to testify on a public health issue during a legislative session, noted, "We only had 132 counties using the system . . . So, we didn’t have a proper count statewide out of 159, to answer that [the legislative committee’s questions]. We did not have answers for those particular counties.”

Several participants noted that data and information are usually managed by data experts or epidemiologists, not by administrators or managers themselves when there are appropriate and enough staff in their jurisdictions. For example, a local health department that can hire a geographic information system (GIS) analyst may be available to develop maps for checking data quality and analyzing data as needed. One respondent from a jurisdiction serving a mid-sized population indicated, “We can see if there are things that are widely out of range…and we need to look at it [to see] if there is something wrong with that particular data field. It takes visualization. It is the easiest way to do that when you have hundreds of variables in a dataset.”

On the other hand, some participants reported that not having enough staff, particularly staff trained in data analysis and visualization (e.g., epidemiologists or informaticians), made it difficult for local health departments to utilize data adequately. This was especially true in small or rural local health departments. Administrators from small-sized local health departments reported that they sometimes ask other nearby counties to share data and information that might reflect their counties.

Data for Decision-Making and Communication

Public health professionals collected and understood data, information and evidence, then used them for decision-making and communication for three main audiences: 1) local or state health department staff themselves, 2) policymakers and stakeholders, and 3) the general public.

First, local or state health department staff saw themselves as an audience that uses data and information for various purposes, such as understanding the needs of their communities, so that they can set priorities for programs and services and allocate resources. They also wanted to assess client satisfaction with the programs and services they provide for the purposes of quality improvement. Often, they described using data and information to publish annual reports and research papers. During these processes, data are usually shared organizationally.

A second main audience is policymakers and stakeholders. Data and information are used to justify how their agency funding has been used and to advocate for more funding. For these audiences, participants noted, just presenting data is usually not enough to advocate for their programs and funding. Stakeholders and policy-makers, they indicated, are more likely to listen to stories and well-crafted narratives built upon data to more easily persuade them in their decision-making. Participants also reported that one-page factsheets or infographics are useful for first contacts with staff in legislators' offices, because those formats are easy to digest, can catch their attention visually, and can potentially initiate a dialogue in which additional information can be provided. As one participant noted, "We have a one-pager that's really an infographic, and we hand that out all the time at community meetings to legislators and presentations to our city council. We have it on our website." Data and information also can be used to advocate for developing policies. For example, one participant, a state director of Assessment, Policy Development & Evaluation, used data related to swimming pool inspections to communicate current needs regarding pool safety and to inform policymakers during the bill development process.

A third main audience was described as the general public and concerned the public's desire to access public health information. Participants noted that one-pagers or infographics are again very useful for educating or communicating to people broad concepts related to public health prevention issues, specifically when one wants to motivate behavior change. As one public health leader noted, visual displays of data are particularly more useful as a communication tool, with policymakers and the public, compared to their use for decision-making internally within their agencies. When visualization is used for communication, the participant saw his/her role as an interpreter of data for the audiences by producing and presenting data and information most convincingly based on their interpretation of it.

Preference for Data Visualization and how it is being used

This study revealed current use of data visualization and preferences for data visualization features among public health professionals, aiming to inform the design of future data visualization tools that may be useful for decision-making and communication among public health professionals.

Most participants showed high interest in applying visualization skills or tools in their public health practice. One respondent noted the value of visualizations by saying “Pictures always speak a million words.” Several points were made by participants regarding how public health practitioners are using visualization and their preference for it. Generally, most participants indicated liking simple and easily understandable visualizations even for lay audiences, as well as appealing data visualizations for communicating with stakeholders. By using visualization tools, they want to enhance the way they use data—comparing their county or state with and across counties, states, and the nation to know where their agencies or jurisdictions stand relative to others—more effectively.

Participants described specific preferences for design features. These included:

1) Overview first, zoom and filter, then details-on-demand[20]

   Participants expressed that they like visualized information, one-pagers, and executive summaries first, but then also wanted to dive down into detailed information by clicking on something. Detailed information includes data sources and methods of data collection, so one can discern for oneself whether the data are reliable and valid to use and share.

2) Well-organized information (e.g., standardization of categories)

   Participants expressed difficulty with finding data in visualization tools because data are usually categorized differently depending on the data source. "If you are looking at child and maternal health data versus environmental health and then that's broken down into the water, air and then chronic disease, you may have found this data under one heading in this website, but in another website, it is under a different heading."

3) Maps

   They consider maps an essential tool. Participants expressed that a lot of health issues are location-based. "Are we seeing respiratory issues more for people living along an interstate? Because we have an interstate that goes through our county …. that's a very specific ribbon of pollution that happens right around the highway, so are we seeing an increase in respiratory issues or things of that nature?" A respondent noted that by looking at the problem area, it reinforces the understanding that problems are not equally distributed, and it helps them to focus their limited resources to make a bigger impact in that area. "When you can map it out and show that truly these three zip codes are the areas of most need and that's why we are only going to do this program in these three zip codes. It makes a very compelling argument." This spatial information from maps can also serve to communicate with policy-makers because maps provide a snapshot of how the problems are occurring geographically.

4) Tables vs. Graphs

   Participants also expressed liking graphs and charts but not in all cases. Tables, they indicated, are sometimes preferred over charts, when graphs and charts become too complicated. "If there is too much information, people try to put too much into a chart with many lines or many bars." They noted that it takes some time or help to first understand visualizations, but once you are self-taught or taught by others, they become usable.

DISCUSSION

While the process of data use was complex and varied depending on the participant, I identified three major themes related to how public health professionals use data for their public health practice as well as challenges they faced. Although these themes were categorized separately for analytic purposes, they were non-linear in many situations and were described as not mutually exclusive. Public health professionals in this study expressed that they place great value on data, information, and evidence for 1) understanding current public health issues, informing their decisions regarding their programs and services, 2) communicating with stakeholders for policy development and funding allocation and 3) educating the public. I found profound challenges, however, in their use of public health data, information, and evidence. The challenges found in this study were consistent with other studies in the field, including old and inconsistent information systems and scattered data sources,[14–16] insufficient data,[7,15] inadequate sub-county data,[14,15] data that were old or not timely,[7] inadequate funding for information systems[4] and inadequate staff with the relevant skills and knowledge.[2,7,14]

Our results suggest how data visualization tools could be used to address these challenges, based on practitioner reports of how they use data and data visualization (Table 3.3). Several points can be discussed with regards to how data visualization can improve data use in each data use process. As reported by a participant, checking data quality using maps was helpful in checking if there are any missing data points or outliers. Visualizing data and information can help increase understanding of complex real-world problems by integrating large quantities of information from multiple sources and reducing cognitive burden, resulting in effective and efficient information-based decision-making.[12,24–26] For example, participants in our study expressed wanting to have an interactive visualization tool that combines multiple data sources representing social determinants of health, so that they can understand what factors contribute to health disparities in certain areas by looking at maps and graphs of their choice. Interactive data visualization further provides additional information that users can gain access to and manipulate using filters, colors, technical documentation, and links to detailed information.

Public health professionals in my study expressed repetitively that visualized data and information can be shared and presented more easily between agencies, in-house, with stakeholders, policymakers, or with the public at large. By using visualization skills, they would enable them to make more effective communication tools with policymakers for advocating their funding, programs, and services.[27] A study by Pandey et al. on an effect of visualization on persuasion found that chart displays have statistically significant effects on attitude change for socially controversial topics (e.g., the relationship between strict imprisonment and decreases in crime), depending on the participants' initial opinion on the topics.[28]

Findings from this study regarding specific preferences for types of data visualization have also confirmed a previous study.[29] That study similarly found that health policy experts liked to have simple data visualization with interactivity, colors, and tooltips as well as a map of health indicators. This study also confirms a special preference for using maps.[29] An interactive map can be a powerful tool for understanding and effectively communicating public health issues with multiple layered data sets on a geospatial presentation.[30] Kard and colleagues described in his book[16] that scientific data tends to be based on physical data (e.g., the human body, the earth). Public health data and indicators are not inheritably based on physical space (e.g., budget, number of staffs, number of outbreaks of communicable disease), however, those data are combined with data with geocoding that present geographical space. Maybe that is why public health professionals in this study and others prefer having a map that enables the public health data to be presented geographically for improved understanding.

Table 3.3. Addressing gaps through visualization

Theme	Visualization relevance to improving data use
Collection of data, information, and evidence	• Can improve data clarification and identification of missing or wrong data quickly by giving a visual overview
Approaches to understanding data and information	• Can include multiple data in one place to improve understanding • Can represent trends well • Can improve data analytics by showing results visually and interactively (filtering, color, map, customization) • Can support non-data experts in interacting with data
Data for Decision-Making and Communication	• Can support decision-making • Can support communication (infographics, map)

Although data visualization has the potential to bring benefits associated with understanding and use of data, this study suggests that the use of data visualization cannot be achieved without addressing systematic challenges such as old and inconsistent data collection and management systems and lack of funding for staff with relevant skills. Public health leaders in our study and other research stressed the need for improvements in data quality and in their use of data to better fulfill their needs.[21] Those challenges exist across themes in this study regarding public health data use; the collection of data, information, and evidence, understanding data and information, and data for decision-making and communication. For example, if public health leaders only have data that are incomplete and inadequate, the interpretations they make from these data have limitations for understanding public health needs and may not be reliable for decision-making or communication. On the other hand, when a health department does not have enough staff with relevant skills, their capacity to collect adequate data can be limited, or data use itself can be restricted even when reliable data exist to meet needs.

Limitations

There are potential limitations of this study. Although interview participants were varied regarding the states they are from, level of jurisdictions to which they belonged, and positions in which they were involved in using data, the themes presented here may not have captured the full range of perspectives. Our findings are based on a purposive sample of interview participants drawn from willing participants who are likely to be more data-savvy or interested in the use of data, thus, they may under-represent the public health practice challenges to data access and use. Also, the majority of our participants were from Washington State (64.2%) and local health departments (78.6%). This could limit our participants' perspectives, as the status of data use in public health could differ depending on which state one is in. Perspectives between a state health department and local health department leaders may also be very different. Finally, our research focused on public health professionals in positions that use data for their decision-making; therefore, our findings are not generalizable to all public health practitioners at every level.

CONCLUSION

In this study, I assessed how public health leaders use data, information, and evidence, and their current use of and preferences regarding data visualization in their practice. I found that public health leaders use data, information, and evidence from various resources daily for communication with co-workers, stakeholders, and the public, and for decision-making regarding their programs and services in various settings. They desire to use data much more and better. Many public health professionals also appear to already be adopting some form of data visualization such as graphs, charts, maps, and infographics to improve their practice. They face systematic challenges that impair adequate data, information, and evidence use.

Opportunities exist, however, for ways to support public health leaders in using data better through data visualization and by supporting those challenges at the agency level. Such support needs to be also accompanied by large system changes making data more reliable and accessible.

REFERENCES

1. Brownson RC, Fielding JE, Maylahn CM. Evidence-Based Public Health: A Fundamental Concept for Public Health Practice. *Annu Rev Public Health*. 2009;30(1):175-201. doi:10.1146/annurev.publhealth.031308.100134

2. Brownson RC, Gurney JG, Land GH. Evidence-based decision making in public health. *J Public Health Manag Pract*. 1999;5(5):86–97.

3. Jenicek M. Epidemiology, evidenced-based medicine, and evidence-based public health. *J Epidemiol*. 1997;7(4):187–197.

4. Leider JP, Resnick B, Kass N, et al. Budget- and Priority-Setting Criteria at State Health Agencies in Times of Austerity: A Mixed-Methods Study. *Am J Public Health*. 2014;104(6):1092-1099. doi:10.2105/AJPH.2013.301732

5. Bekemeier B, Chen A, Kawakyu N, Yang Y. Local Public Health Resource Allocation. *Am J Prev Med*. 2013;45(6):769-775. doi:10.1016/j.amepre.2013.08.009

6. Bekemeier B, Chen AL, Kawakyu N. Mandated activities and limited decision-making authority among local public health officials. *Front Public Health Serv Syst Res*. 2012;1(3):4.

7. Baum NM, DesRoches C, Campbell EG, Goold SD. Resource allocation in public health practice: a national survey of local public health officials. *J Public Health Manag Pract*. 2011;17(3):265–274.

8. Lovelace KA, Aronson RE, Rulison KL, Labban JD, Shah GH, Smith M. Laying the Groundwork for Evidence-Based Public Health: Why Some Local Health Departments Use More Evidence-Based Decision-Making Practices Than Others. *Am J Public Health*. 2015;105(S2): S189–S197.

9. Sosnowy CD, Weiss LJ, Maylahn CM, Pirani SJ, Katagiri NJ. Factors Affecting Evidence-Based Decision Making in Local Health Departments. *Am J Prev Med*. 2013;45(6):763-768. doi:10.1016/j.amepre.2013.08.004

10. National Association of County & City Health Officials. *2013 National Profile of Local Health Departments*.; 2013. http: //www.naccho.org/topics/infrastructure/profile/upload/ 2013-National-Profile-of-Local-Health-Departments-report. pdf. Accessed December 8, 2015.

11. Ryan GW, Bloom EW, Lowsky DJ, et al. Data-Driven Decision-Making Tools To Improve Public Resource Allocation For Care And Prevention Of HIV/AIDS. *Health Aff (Millwood)*. 2014;33(3):410-417. doi:10.1377/hlthaff.2013.1155

12. Telea A. *Data Visualization: Principles and Practice*. Second edition. Boca Raton: CRC Press, Taylor & Francis Group; 2015.

13. Few S. *Now You See It : Simple Visualization Techniques for Quantitative Analysis*. 1st edition. Oakland, California: Analytics Press; 2009. http://www.tableau.com/blog/stephen-few-data-visualization. Accessed December 16, 2015.

14. Few S. What Is Data Visualization? Visual Business Intelligence. https://www.perceptualedge.com/blog/?p=2636. Published May 4, 2017. Accessed January 23, 2019.

15. Ware C. *Information Visualization Perception for Design*. 3rd ed. Waltham, Mass: Morgan Kaufmann; 2013.

16. Kard ST, Mackinlay JD, Schneiderman B. *Readings in Information Visualization, Using Vision to Think*. San Francisco, Calif: Morgan Kaufmann; 1999.

17. Brodlie KW, Carpenter L., Earnshaw R., et al. *Scientific Visualization: Techniques and Applications*. Springer Science & Business Media; 2012.

18. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101. doi:10.1191/1478088706qp063oa

19. Texifter. Coding Analysis Toolkit. https://cat.texifter.com/. Accessed April 30, 2018.

20. Shneiderman B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: *The Craft of Information Visualization*. San Francisco: Morgan Kaufmann; 2003:364-371. doi:10.1016/B978-155860915-0/50046-9

21. Leider JP, Shah GH, Williams KS, Gupta A, Castrucci BC. Data, Staff, and Money: Leadership Reflections on the Future of Public Health Informatics. *J Public Health Manag Pract*. 2017;23(3):302-310. doi:10.1097/PHH.0000000000000580

22. Shah SN, Russo ET, Earl TR, Kuo T. Measuring and Monitoring Progress Toward Health Equity: Local Challenges for Public Health. *Prev Chronic Dis*. 2014;11. doi:10.5888/pcd11.130440

23. Lovelace K, Shah GH. Informatics as a Strategic Priority and Collaborative Processes to Build a Smarter, Forward-Looking Health Department: *J Public Health Manag Pract*. 2016;22: S83-S88. doi:10.1097/PHH.0000000000000452

24. Savikhin A, Maciejewski R, Ebert DS. Applied visual analytics for economic decision-making. In: *The IEEE Symposium on Visual Analytics Science and Technology*. IEEE; 2008:107–114. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4677363. Accessed January 11, 2016.

25. Dolan JG, Veazie PJ, Russ AJ. Development and initial evaluation of a treatment decision dashboard. *BMC Med Inform Decis Mak*. 2013;13(1). doi:10.1186/1472-6947-13-51

26. Rudolph S, Savikhin A, Ebert DS. FinVis: Applied visual analytics for personal financial planning. In: IEEE; 2009:195-202. doi:10.1109/VAST.2009.5333920

27. Speyer P, Pagels B, Park N. Communicating data for impact. *Commun Data Impact – Seattle Inst Health Metr Eval Forum One*. 2015.

28. Pandey AV, Manivannan A, Nov O, Satterthwaite M, Bertini E. The Persuasive Power of Data Visualization. *IEEE Trans Vis Comput Graph*. 2014;20(12):2211-2220. doi:10.1109/TVCG.2014.2346419

29. Zakkar M, Sedig K. Interactive visualization of public health indicators to support policymaking: An exploratory study. *Online J Public Health Inform*. 2017;9(2). doi:10.5210/ojphi.v9i2.8000

30. Sopan A, Noh AS-I, Karol S, Rosenfeld P, Lee G, Shneiderman B. Community Health Map: A geospatial and multivariate data visualization tool for public health datasets. *Gov Inf Q*. 2012;29(2):223-234. doi:10.1016/j.giq.2011.10.002

CHAPTER 4. Tables versus visualization for understanding data among public health practitioners

ABSTRACT

Introduction: Despite the potential value of data visualization for improving the understanding of data in a manner that results in better decision-making in health and healthcare, there is little understanding of its applications in public health practice, particularly for public health program planning and resource allocation. This study examined how data were understood differently by public health practitioners when displayed in table format versus more visualized formats. This study also examined practitioners' confidence in their understanding and perceived ease of use with table and visualization presentations.

Methods: An exploratory mixed-methods design using online surveys was used. The study developed an online dashboard that included a set, including a table and an interactive visualization that were used in the online survey. The online survey included scenarios and related questions that participants were asked to answer by exploring the dashboard.

Results: Twenty-two public health practitioners participated in this survey. Content analysis of the qualitative data collected by open-ended questions described relationships that participants found by looking at data in a table and a visualization. Public health practitioners found similar types of relationships through the comparisons they made with both the table and visualization format data, such as comparing to mean, median, min, or max, comparing by population size categories, or between subcategories. It was found that data visualization has benefits in advancing the understanding of data among public health practitioners. Data visualization seems to make it easy for participants to find information from data, reduce errors in assessing information, as well as process data, and help generate more meaningful knowledge.

Conclusions: The study findings suggest that data visualizations could be useful in advancing the understanding of data among local public health leaders, and in filling a gap in understanding how to improve their use of data and data visualizations for their practice and decision-making.

Keywords: data visualization, interactive data visualization, public health practice

INTRODUCTION

One of the eight domains of core competencies for public health professionals is analytical and assessment skills that include collecting, understanding, analyzing, and interpreting valid and reliable data, and making evidence-based decisions to address community health needs. [1] Public health practitioners are expected to use "data and information, and evidence systematically" when making decisions for programs for disease prevention and health promotion to incorporate the concept of Evidence-Based Public Health (EBPH). [2–4, 5 (p177)] In a study assessing how urban US city health departments use data to inform their work, local public health leaders responded that they believed using local data was good for informing decisions regarding public health programs and policies and educating decision-makers about certain community health problems. For example, the study cites how practitioners using data to cite notably higher smoking rates in their district compared to ones in neighboring districts drew more attention from decision-makers to the problem. [2]

Public health professionals, however, underutilize data, information, and evidence for assessment and analysis of public health issues as guides in decision-making. The literature consistently indicates the need for public health practitioners to have accessible, "timely, easy to digest, and up-to-date information that is filtered, summarized, and synthesized from the authoritative content source." [3(p412)] A qualitative study that conducted interviews with New York Local Health Department (LHD) leaders and upper-level staffs found that not all decisions made were based on evidence; some were instead based on "gut feeling", even though the LHD leaders were aware of the importance of evidence and desired to use it more. [4] In a nationwide survey study with LHD practitioners, researchers found that LHD practitioners ranked funding guidance from legislative authorities or state health departments as more important for their decision-making (e.g., program planning, policy development, or funding) than scientific resources such as systematic reviews of scientific literature (e.g., *Community Guide*), scientific reports, and general literature review articles. [5] In a more recent NAM report, the importance of information and information systems for monitoring community health needs were emphasized, but the report expressed concern about "inadequate access to information systems and communication tools" for the public health infrastructure. [6]

Visualizing data and information can help increase understanding of complex real-world problems by integrating large amounts of information from multiple sources, making data more accessible, and reducing cognitive burden, resulting in effective and efficient information-based decision-making. [7–10] Visualization broadly refers to any visual representation of data and information such as charts, graphs, bars, maps, and images that aid in the understanding and communication of data. [11] Good visualizations help explore unknown problems with data and answer questions about real-world problems. [12]

Previous experimental studies have found that data and information aided with visualization tools are more likely to impact human understanding, perception, motivation, attitude, and decision-making. [7–9,13–15] In a study examining how visualization tools help emergency unit staff understand situations better by looking at visualized map versus a table, researchers found that staff with a visualized map had a significant increase in their situation awareness and their confidence regarding their understanding. A study by Hsu and Lin reported that a decision-making learning module that included visualization tools, cognitive guide, and group discussion, improved high school student's decision-making related to socio-scientific issues (e.g., building a water reservoir while considering the surrounding environment). [16] Another study by Kim revealed that a SimulSort, a visual aid for multi-attribute decision-making, enhanced the probability of making better decisions (e.g., selecting the best apartment considering its location, price, size, etc.). [15] As found in my previous qualitative study (S, Park, dissertation paper), public health practitioners expressed that, based on their experience, data visualization was useful, because of its potential value for understanding and communicating data.

Despite the potential value of data visualization for improved understanding of data and therefore better decision-making in health and healthcare,[17,18] there is little understanding of its applications in public health practice, particularly for public health program planning and resource allocation. To begin to build this body of public health research to support the use of data for decision-making using data visualization, this study examined how data were understood differently by public health practitioners when displayed in table format versus more visualized formats. While tables are a more traditional way of collecting and displaying data in many governmental organizations, including public health, this study also sought to examine whether visual displays of data and information were easier for public health practitioners to understand compared with tables of the same data and information.

METHODS

Participants

I used convenience sampling to recruit eligible participants for the study described here, drawing a sample from among the study participants of an ongoing research project (RWJF #73187, P.I Betty Bekemeier) by the Public Health Activities and Services Tracking (PHAST) research team. The participants in the PHAST study had provided de-identified aggregated secondary public health administrative data that were used for creating a data visualization tool for a larger study. Participants in the larger PHAST study had already demonstrated an interest in improving their access to and understanding of existing data and, thus, were expected to be more likely to participate in the study described here. In order to recruit from among the PHAST study participants, I sent out recruiting email invitations in September 2017 and contacted representative public health leaders from Missouri, Nebraska, and New Jersey for whom we had contact information through PHAST's previous research. The recruiting email asked recipients to share information about this study with colleagues in their agencies to identify more potential study volunteers. After receiving reply emails from volunteers, I screened them for eligibility. Inclusion criteria for this study were: 1) working at a local or state health department, and 2) serving in a position of director, administrator, program lead, or director of finance, etc. that involves program planning, finance, or decision-making in a local or state health department.

Participants, who fully completed the survey, were invited to voluntarily provide their email address for a random drawing, in a link separate from the survey. The compensation for the two randomly selected participants was $600 each towards registration for a conference of their choice. This study was approved by the University of Washington Human Subjects Division Institutional Review Board.

Design, Procedures, and Analysis

An exploratory mixed-methods design using online surveys was employed to capture how data were understood through two different formats of data presentation – a table and visualization. I also examined participants confidence in their understanding as well as their perceived ease of use with the table and visualization presentations.

Recruited participants were sent an email with a link to an online survey that included scenarios and related questions. Survey participants were first asked to read a scenario in which they imagined that they were working for a fictitious LHD and doing budget planning or budget reporting about a communicable disease control program within that LHD. The scenario asked participants to develop an understanding of budget-related data from their LHDs and then compare them with data from other LHDs within their state or in other states. Then, they were asked to consider how much money should be allocated within their own (fictitious) LHD to fund different communicable disease control programs in their agency. Finally, they followed a link to a budget-related dataset that could be explored in two formats: a table and a visualization. To prevent bias, participants were randomly allocated to two groups; one group looked at the table format data first and then looked at the visualization format data, and the other group started with the visualization format first. Participants were then asked to 1) compare “their” (fictitious) agency with other (also fictitious) agencies in “their” state with regard to communicable disease control programs and to summarize what they saw in the comparison, 2) compare “their” agency with others in “other” states in terms of communicable disease control programs and to summarize what they saw in the comparison, and 3) compare “their” agency with “others of a similar population size” in terms of communicable disease control programs and to summarize what they saw in the comparison. Participants were asked to provide these three written summaries in order to understand the insights that participants gained from the comparisons. The purpose of asking them to summarize the data they saw was not to check if they were “correct” in their responses or to check how quickly they understood the data; rather it was to capture their description of how they understood the data as they might in the context of a typical work task.[19] After exploring both types of information (table and visualized formats), subjective levels of confidence about their comparison summaries were assessed using Likert scales (e.g.1- very confident, 2- somewhat confident, 3- neutral, 6- not confident at all). I also asked which formats of data display were easier to use for making comparisons, using a multiple-choice question (e.g., Table was much easier than visualization, Visualization was somewhat easier than the table).

Finally, they were asked to provide general demographic information such as gender, age, years of public health experience, background fields (e.g., epidemiology, economics, public health), education level, graph literacy, and numeracy level. Numeracy (quantitative literacy), the ability to understand and use numerical information for decision-making, [20] was assessed using a modified five item-numeracy measure based on Lipkus, Samsa, & Rimer’s 11-item numeracy scale. [21] The objective scale used included a free-response format that assessed the ability to convert between proportion and percentage. Graph literacy, “the ability to understand graphically presented information” was assessed using a 5-item subjective graph literacy (SGL) scale. [22,23]

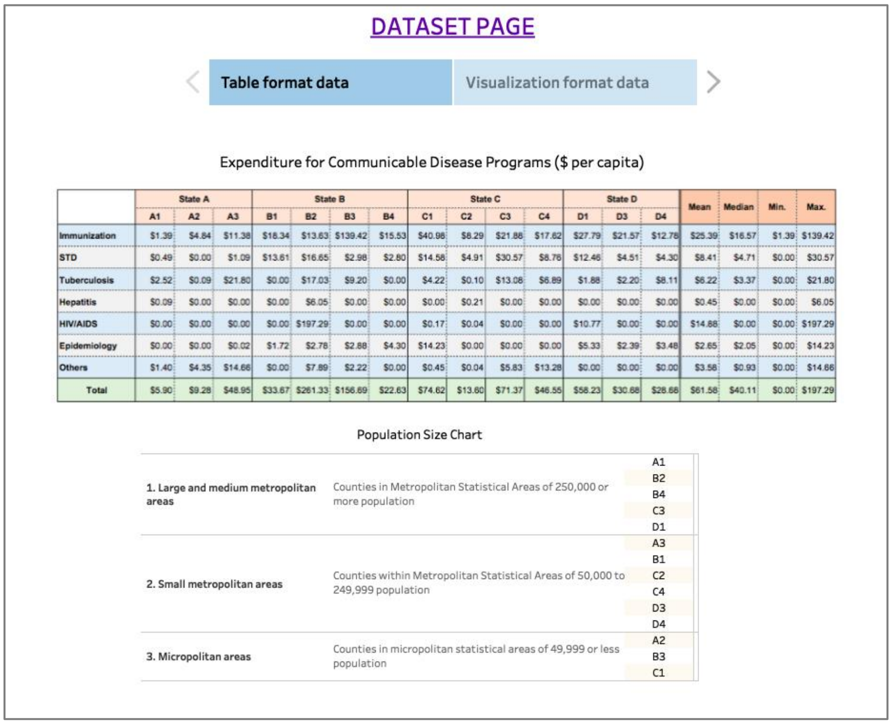
Table versus Visualization

The online dashboard used in the survey instrument included a set of tables and a visualization. The same dataset was used for both the table and the visualization. The dataset represented per capita expenditures of communicable disease control programs and activities in 14 LHDs from four states and was used by permission from a previous PHAST study project. The LHD data for the participants’ fictitious LHD were unlabeled in terms of an LHD's actual name to maintain data confidentiality and to focus the participant on the data and not its source.

The table used in this study (Figure 4.1) displayed per capita expenditures using multiple categories (the subcategory of communicable disease program on the row intersected with columns of agencies [e.g., A1, A2, A3] and states [e.g., State A]). The table allowed participants to compare per capita expenditures between subcategories of communicable disease programs (e.g., immunizations, tuberculosis) as well as between agencies within and across the four states. At the end of each row, a descriptive summary of the per capita expenditures such as mean, median, minimum, and maximum were added to aid their exploration.

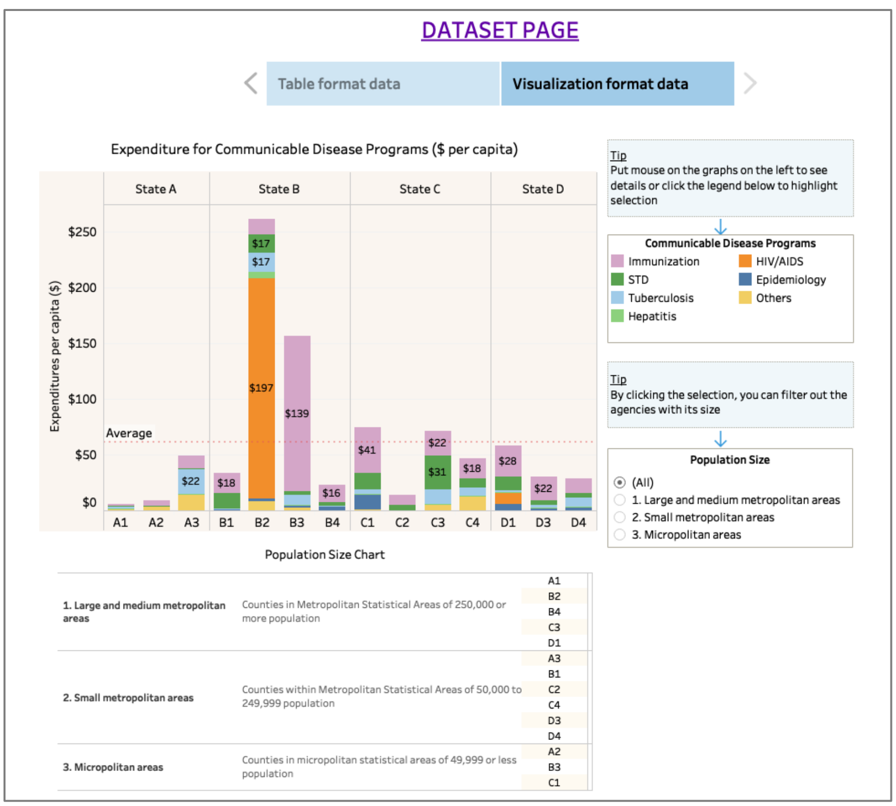
The visualization used was developed using Tableau 10.3.13, for which I used stacked bar charts to show the distribution of per capita expenditures of LHDs separated by states (Figure 4.2). To support users' exploration of the visualization, I included several features such as filtering, hovering, selecting, and color coding with the addition of short notes to guide participants' use of the visualization. The participants could, for example, hover over the graph in Figure 4.2 to see details or click the legend to highlight the selection of sub-programs under the communicable disease programs. They could also select large and medium metropolitan areas, small metropolitan areas, or micropolitan areas (Figure 4.2). To help the participants know what population size category "their" LHD belonged to, a population size chart was included on the bottom of the bar chart. This population size category was intended to help them compare "their" agency with similar-sized agencies, which had been a desire frequently expressed by public health leaders in previous PHAST studies. [24]

Figure 4.1. Table for a set of per capita expenditure of communicable disease programs



	State A			State B				State C				State D			Mean	Median	Min.	Max.
	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	D1	D3	D4				
Immunization	$1.39	$4.84	$11.38	$18.34	$13.63	$139.42	$15.53	$40.98	$8.29	$21.88	$17.62	$27.79	$21.57	$12.78	$25.39	$16.57	$1.39	$139.42
STD	$0.49	$0.00	$1.09	$13.61	$16.65	$2.98	$2.80	$14.58	$4.91	$30.57	$8.76	$12.46	$4.51	$4.30	$8.41	$4.71	$0.00	$30.57
Tuberculosis	$2.52	$0.09	$21.80	$0.00	$17.03	$9.20	$0.00	$4.22	$0.10	$13.08	$6.89	$1.88	$2.20	$8.11	$6.22	$3.37	$0.00	$21.80
Hepatitis	$0.09	$0.00	$0.00	$0.00	$6.05	$0.00	$0.00	$0.00	$0.21	$0.00	$0.00	$0.00	$0.00	$0.00	$0.45	$0.00	$0.00	$6.05
HIV/AIDS	$0.00	$0.00	$0.00	$0.00	$197.29	$0.00	$0.00	$0.17	$0.04	$0.00	$0.00	$10.77	$0.00	$0.00	$14.88	$0.00	$0.00	$197.29
Epidemiology	$0.00	$0.00	$0.02	$1.72	$2.78	$2.88	$4.30	$14.23	$0.00	$0.00	$0.00	$5.33	$2.39	$3.48	$2.65	$2.05	$0.00	$14.23
Others	$1.40	$4.35	$14.66	$0.00	$7.89	$2.22	$0.00	$0.45	$0.04	$5.83	$13.28	$0.00	$0.00	$0.00	$3.58	$0.93	$0.00	$14.66
Total	$5.90	$9.28	$48.95	$33.67	$261.33	$156.69	$22.63	$74.62	$13.60	$71.37	$46.55	$58.23	$30.68	$28.68	$61.58	$40.11	$0.00	$197.29

Figure 4.2. Stacked bar chart for a set of per capita expenditure of communicable disease programs



Analysis

I used a mixed method approach for the analysis in coding qualitative data and analyzing quantitative survey questions.

Qualitative coding of open-ended questions. Qualitative content analysis was used to code themes in the open-ended questions regarding a participant's summary of what they understood from the table and the visualization to describe differences in decoding the data between table and graph. [25] The study author (Park, S) read the responses and performed open coding. Codes were assigned to emerging themes and iteratively refined and collated into broader concepts as more of the transcripts were coded. After the coding was completed and to ensure rigor, I used a "peer review" approach, discussing my coding and emerging themes with a PHAST team member and checking for agreement. [26–28] This discussion reached almost total coding consensus. The coded themes were tabulated for their frequencies by comparing table vs. visualization.

Quantitative descriptive analysis. Descriptive statistics including mean, variation, and frequencies were used to describe the study participants, perceived ease of use between a table and visualized format data and scaled confidence for an understanding of table and visualized formats of data. Also, perceived ease of use was examined relative to age, years of working experience, graph literacy, and numeracy.

RESULTS

Participants' Characteristics

A total of 22 public health practitioners participated in this online survey. Among participants, 63.6% were female, the average age was 48 (SD: 12.6) years, and 45.4% had graduate-level degrees (Table 4.1). Most participants (77.3%) held positions as directors, administrators, health officers, or program managers, suggesting that they were in decision-making positions in their agencies.

Table 4.1. Demographic characteristics of the study participants

Characteristics	Category	Participants (N=22)
Age in years (n, %)	Less than 40	5 (22.7%)
	40-49	5 (22.7%)
	50-59	8 (36.4%)
	60-69	4 (18.2%)
Gender (n, %)	Female	14 (63.6 %)
	Male	8 (36.4 %)
Highest Education (n, %)	Some college	1 (4.5 %)
	Bachelor’s Degree	11 (50.0 %)
	Master's Degree	9 (40.9 %)
	Doctoral Degree	1 (4.5%)
Training Field (multiple answers)	Public Health	8
	Business	4
	Nursing	3
	Administration	3
	Epidemiology	2
	Accounting	2
	Others (Geography, Biology, etc.)	3
Job Title	Director/administrator/Health Officer	10
	Program Manager	7
	Fiscal Director, Accountant	3
	Health Educator	1
	Researcher	1
Years of Working in Public Health (Mean, SD)		8 (7.8)
Numeracy Score out of 5 (Mean, SD)		4.62 (0.59)
Graph Literacy Score out of 25 (Mean, SD)		19.68 (3.56)

Comparisons of Tables versus Visualization

Content analysis of the qualitative data produced 18 themes from the set of three open-ended questions for the table and visualization. Ten themes were mentioned and discussed by participants that related to comparison relationships that participants summarized when looking at the table and the visualization (Table 4.2). Also, eight themes were identified that were about insights separate from comparison relationships (Table 4.3). Most themes, regarding the comparisons (Table 4.2) and insights separate from comparisons (Table 4.3) were identical for the table and the visualization, but some themes occurred only with the table or with the visualization. The themes are presented below and in both tables in order of their frequency of occurrence.

Comparison Relationship. The most frequent types of comparison the participants made between both the table and visualized data were 'comparing to mean, median, min, or max; although it should be noted that for visualization data it was only possible to compare to mean.' The next most frequent was followed by 'comparison by population size categories' and 'comparison by subcategories under the communicable disease' (Table 4.2). Then, for the table, participants tended to compare by ranking LHDs in terms of per capita expenditure or listing exact values of per capita expenditures of each LHD. As noted by a participant, "[Our agency] is spending almost the lowest amount on STD [Sexually Transmitted Diseases], 2nd only to D4 agency…" Some comparisons were only found in the visualization, such as, when participants referred to comparisons using the visualization, they tended to describe what they compared in terms of an overall picture of the information. For example, a participant reported, "Though most agencies of my size spend the bulk of their communicable disease funds on immunizations, A3 [agency] spends the bulk on Tuberculosis services." Another participant noted, "With the exception of B2 [agency], all seem to dedicate the majority of resources to immunizations." Others also compared the data values by gauging the size of the bar graph, as noted by a participant, "...the amount spent on Immunization in C3 [agency] was half [of] the amounts spent in C1 [agency], whereas the amount spent on TB [Tuberculosis] was double that of C1." In addition, participants made comparisons using filters that the visualization tool provided to sort agencies by population size. Such a filtering option was not included or possible in table format data.

Table 4.2. Differences between table and visualization responses regarding comparisons

	Table (n)			Visualization (n)			Table	Visualization
	Within state	Across state	Within similar population size	Within state	Across state	Within similar population size	Total (n)	Total (n)
Compare to mean/median/min/max	8	16	7	7	7	5	31	19
Compare by population size	5	3	13	3	2	13	21	18
Compare by each category under the communicable disease	8	4	3	8	5	1	15	14
Compare by ranking	5	2	4	2	1	4	11	7
Exact/individual values	4	1	5	4	1	2	10	7
Variation	2	3	2	3	3	3	7	9
Part to Whole (percentage)	2	3	2	3	1	4	7	8
Comparison: overall pictures	0	0	0	0	7	1	0	8
Compare by size of bar graph	0	0	0	4	0	1	0	5
Using filter function	0	0	0	3	1	1	0	5

Table 4.3. Differences between table and visualization insights unrelated to comparisons

	Table (n)			Visualization (n)			Table	Visualization
	Within state	Across state	Within similar population size	Within state	Across state	Within similar population size	Total (n)	Total (n)
Using adverbs of degree	5	3	2	2	7	2	10	11
Ah-ha moments	4	2	0	4	4	3	6	11
Lazy answer/takes time	0	0	6	0	1	0	6	1
Making error	1	2	3	1	0	0	6	1
Answers not clear	1	3	2	0	0	0	6	0
Finding Outliers (zero value)	1	3	1	2	2	0	5	4
Express difficulty in finding information	0	2	3	0	1	0	5	1
Express easiness in finding information	0	0	0	1	1	1	0	3

General Insights. Aside from the relationships that participants found for comparison, other insights were reported (Table 4.3). Participants frequently used adverbs of degree such as ‘significantly more' or 'extremely less' to describe comparisons they found from the data both from the table and visualization. Participants also reported outliers that stood out in both the table and visualization.

There also were differences reported between table and visualization. Participants looking at the visualization, more frequently expressed questions or speculations regarding what other factors might be related to the pattern they saw from the data, or their insights from the data—expressing, thereby, an 'ah-ha' moment or a type of musing. One participant, for example, stated "I would be interested in looking at these areas geographically to see what other factors could be involved. Are they closely located to a small metropolitan area or micropolitan areas? Are they in a state, which helps to fund local public health efforts? What are the demographics?" Another participant stated, "Surprisingly when compared to similar-sized counties, I now feel that my spending per capita on Immunization is much higher." Incorrect assessments about the data that were reported by participants were more common in responses related to looking at the table. For example, one participant compared per capita expenditure on each of their agency's communicable diseases programs with two other agencies in their state but mistakenly concluded that their agency spent more on epidemiology than the other two, when this was not the case.

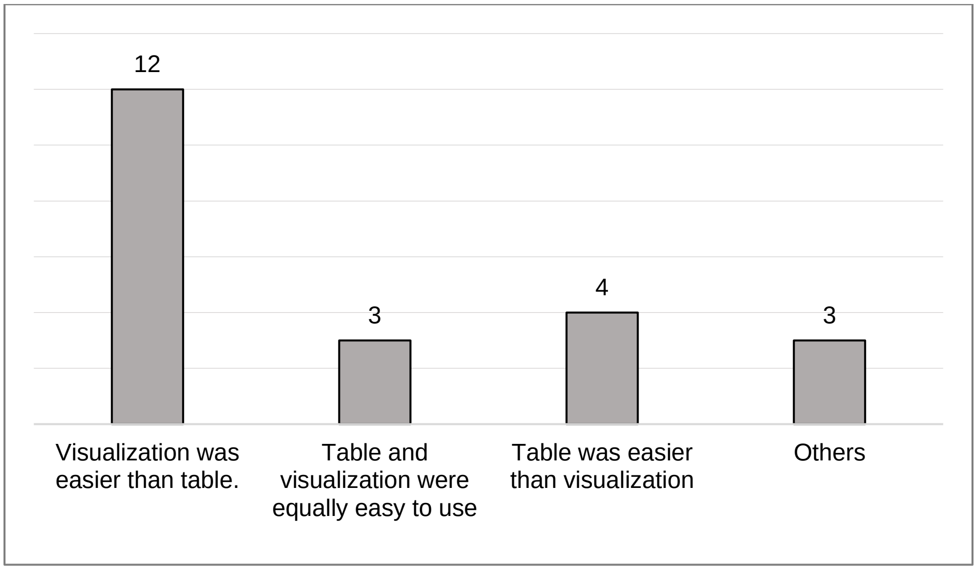
Also, participants more frequently expressed difficulty in finding information when they looked at the table, especially when they were asked to compare the data across states (comparing across multiple columns that were likely more distant columns as well) and within agencies serving similar population sizes. In contrast, participants expressed, in their open-ended responses, that it was easy to find information when they were looking at the visualization format data. For example, as noted by a participant, "…It is very easy to use the filter for each of the disease programs to determine how our agency compares…." Another participant also reported, "Once I determined that our agency is a small metropolitan area, it was very easy to filter on the population size and see only the small metro areas.…"

Confidence in their Comparison. After answering the three open-ended questions about the table and another three about the visualization, the survey asked how confident participants were in each of their answers (total of six questions). Participants answered that they were either “very confident (n=38)”, or “confident (n=87)” in addition to “not sure (n=6)” about the comparisons they summarized for the three open-ended questions for the table and visualization. No participant answered that they were “not confident” about their answers. Among the participants who answered that they were “very confident (n=38)”, 21 participants answered “very confident” when looking at the visualization compared to 17 participants answering this way when looking at the table.

Ease of use regarding data display--table versus visualization

After completing comparisons using both the tables and visualization, participants were asked about which display was easier to use to find and understand the information they needed to make comparisons, far more participants indicated that the visualization was easier (n = 12, 54.5%)) than the table (n = 4, 18.1%) (Figure 4.3). However, some participants stated that the table and visualization formats had both advantages and disadvantages. For example, participants reported that the visualization feature that enabled sorting the LHDs by similar-sized areas was very useful and that it made it easier to compare them with LHDs across states, whereas the table was easier to use to determine total spending or comparing expenditure values of a specific program across all states. There were no apparent differences in the perceived ease of use for data between the visualization and table displays among public health agency participants relative to their age, gender, race, education, experience, graph literacy, or numeracy.

Figure 4.3. Ease of use regarding data display--table versus graph (N=22)



DISCUSSION

In this study, I explored how data presented via visualization was understood differently compared to data presented in a more traditional table format. Qualitative analysis of open-ended questions described comparison relationships that LHD participants found by looking at data in a table and a visualization format. In this study, public health practitioners found similar types of relationships through the comparisons they made with both the table and visualization format data, such as comparing to mean, median, min, or max, comparing by population size categories, or between subcategories. Also, participants answered that they were "very confident" or "confident" in the comparisons they arrived at from both the table and visualization formats and no one answered that they were "not confident." Participants found that the table was helpful for understanding data, especially for getting exact numbers of per capita expenditures and ranking values from top to bottom. These findings correspond with previous research that has found that tables are particularly valuable for making it easy for users to look up individual values.[29]

This study, however, suggests particular benefits of using visualization in improving the understanding of public health data among public health practitioners. Public health participants more frequently expressed ease in finding information when looking at the visualization, especially when they were asked to compare data across states (comparing across multiple columns) and within agencies of similar-sized populations. This benefit of data visualization confirms a previous study by Zakkar and Sedig[30] that found their study participants were particularly satisfied with the interactive capabilities of visualizations such as filtering and adjusting displayed information. Also, participants reviewing the visualization tended to make fewer errors in assessing information than they did when reviewing the table. These apparent benefits of visualization are supported by literature describing the advantages of visualization in reducing the cognitive burden of data processing.[12,31,32] While public health practitioners typically receive data in traditional table formats in their actual work settings, more than half of our study participants (13 out of 22) tended to express greater ease of understanding with visualization of data. This suggests that improving how data are displayed and made available to public health professionals may improve their ability to make use of data for decision-making.

Interestingly, our qualitative analysis also indicated that, when looking at a data visualization display, participants seemed more likely to proceed in trying to gather 'knowledge' from the data displayed. They tended to look more deeply into the visualization data for further understanding than they did with the table, posing questions and speculating about other factors at play. On the data-information-knowledge-wisdom hierarchy continuum, knowledge is generally built on data and information that is processed through the individual with experience, opinions, beliefs, and skills allowing for better decision-making. [33] Our respondents reviewing the visualization made more mention of the factors they felt explained differences in the patterns they saw. Visualization, therefore, might be a means of improving the processing of data or information and could help to generate more meaningful knowledge that is useful to public health practitioners.

Implications for research and practice

By analyzing and communicating data through visualization, [34] public health practitioners may be better able to discover local public health needs and issues, resulting in more informed decision-making and more effective communication with decision-makers. Future research should examine how data visualization might influence or facilitate better decision-making and communication with more refined and tailored data visualization tools. Designing visualizations involves a multi-step process, including planning, user-task analysis, conceptual design, prototyping, and expert guideline-based evaluation. [35,36] This means that involving public health professionals from the beginning of the design process is crucial. Given challenges in data use (S.Park's qualitative paper) and in evidence-based decision making in public health practice such as inconsistent information systems, limited financial and organization resources, and political inflexibility, [4,37–39] understanding the specific needs of public health professionals and considering other factors that might affect their use of data and visualization in decision making should be considered. [30]

A survey study regarding public health workforce interests and needs by Sellers, et al.[40] found that 'gathering reliable information to answer questions' and 'communicating ideas and information in a way that different audiences can understand' or 'communicating in a way that persuades others to act' are perceived as important workforce competencies. Yet, it was also found that public health employees from state health agencies were unable to perform these skills or were only able to do so at the level of a beginner.[40] Training is, thus, needed for public health leaders in how to create and use data visualization and data assessment tools in their daily work, including generating additional knowledge about their workflow and current data needs and data uses.

Limitations

This study had limitations. The sample size of this study was small (n=22). As such it was difficult to find statistically significant differences between the table and visualization formats in terms of how public health practitioners understood the data. In designing the visualization for this study, the visualization tool used was not rigorously evaluated through usability testing by actual users, since the focus of this study was not evaluating the tool itself. I did, however, incorporate visualization design principles from previous literature,[29,41,42] and the findings correspond with other visualization research, suggesting that such tools can increase ease of understanding and reduce the number of user errors.[43]

CONCLUSIONS

This study highlights how public health practitioners understand data differently when data are presented in a visualization display compared to table format data. Even though participants in this study showed similar types of comparisons when they used table and visualization formats, the findings suggest that data visualization has benefits for using data. These include ease in using data visualization to find information from data, reduction of errors in assessing information, and improvement in participants' processing of data such that they generate more meaningful knowledge. Data visualizations could be particularly useful in improving the understanding of data among local public health leaders, and in filling an important gap in understanding how to improve their use of data and data visualization in their practice and decision-making.

REFERENCES

1. Core Competencies for Public Health Professionals. http://www.phf.org/resourcestools/pages/core_public_health_competencies.aspx. Accessed March 29, 2016.

2. Castrucci BC, Rhoades EK, Leider JP, Hearne S. What Gets Measured Gets Done: An Assessment of Local Data Uses and Needs in Large Urban Health Departments. *Journal of Public Health Management and Practice*. 2015;21: S38-S48. doi:10.1097/PHH.0000000000000169

3. Revere D, Turner AM, Madhavan A, et al. Understanding the information needs of public health practitioners: A literature review to inform design of an interactive digital knowledge management system. *Journal of Biomedical Informatics*. 2007;40(4):410-421. doi:10.1016/j.jbi.2006.12.008

4. Sosnowy CD, Weiss LJ, Maylahn CM, Pirani SJ, Katagiri NJ. Factors Affecting Evidence-Based Decision Making in Local Health Departments. *American Journal of Preventive Medicine*. 2013;45(6):763-768. doi:10.1016/j.amepre.2013.08.004

5. Fields RP, Stamatakis KA, Duggan K, Brownson RC. Importance of Scientific Resources Among Local Public Health Practitioners. *American journal of public health*. 2015;105(S2): S288–S294. http://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2014.302323. Accessed November 23, 2015.

6. Committee on Assuring the Health of the Public in the 21st Century. *The Future of the Public's Health in the 21st Century*. Washington, D.C.: National Academies Press; 2003. http://www.nap.edu/catalog/10548. Accessed December 8, 2015.

7. Larson KL, Edsall RM. The impact of visual information on perceptions of water resource problems and management alternatives. *Journal of Environmental Planning and Management*. 2010;53(3):335-352. doi:10.1080/09640561003613021

8. Rudolph S, Savikhin A, Ebert DS. FinVis: Applied visual analytics for personal financial planning. In: *IEEE Symposium on Visual Analytics Science and Technology*. ; 2009:195-202. doi:10.1109/VAST.2009.5333920

9. Savikhin A, Maciejewski R, Ebert DS. Applied visual analytics for economic decision-making. In: *The IEEE Symposium on Visual Analytics Science and Technology*. IEEE; 2008:107–114. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4677363. Accessed January 11, 2016.

10. Dolan JG, Veazie PJ, Russ AJ. Development and initial evaluation of a treatment decision dashboard. *BMC Medical Informatics and Decision Making*. 2013;13(1). doi:10.1186/1472-6947-13-51

11. Williams J, Sochats K, Morse E. Visualization. *Annual Review of Information Science and Technology*. 1997;30:161-207. http://eric.ed.gov/?id=EJ518328. Accessed January 20, 2016.

12. Telea A. *Data Visualization: Principles and Practice*. Second edition. Boca Raton: CRC Press, Taylor & Francis Group; 2015.

13. Thai L, Blaine R, Jane C, Hilaire T, George D. Design of smart home sensor visualizations for older adults. *Technology and Health Care*. 2014;(4):657–666. doi:10.3233/THC-140839

14. Samek A, Hur I, Kim S-H, Yi JS. An Experimental Study of Decision Process with Interactive Technology. *CESR-Schaeffer Working Paper*. September 2015. doi:http://dx.doi.org/10.2139/ssrn.2347698

15. Kim S-H. The Effectiveness of Interactive Visualizations for Multi-Attribute Decision Making with Contextual Data. *Journal of the Ergonomics Society of Korea*.:11.

16. Hsu Y-S, Lin S-S. Prompting students to make socioscientific decisions: embedding metacognitive guidance in an e-learning environment. *International Journal of Science Education*. 2017;39(7):964-979. doi:10.1080/09500693.2017.1312036

17. Gotz D, Borland D. Data-Driven Healthcare: Challenges and Opportunities for Interactive Visualization. *IEEE Computer Graphics and Applications*. 2016;36(3):90-96. doi:10.1109/MCG.2016.59

18. Shneiderman B, Plaisant C, Hesse BW. Improving Healthcare with Interactive Visualization. *Computer*. 2013;46(5):58-66. doi:10.1109/MC.2013.38

19. Cleveland WS. *The Elements of Graphing Data*. Rev. ed. Murray Hill, N.J.: Murray Hill, N.J. : AT & T Bell Laboratories; 1994.

20. Nelson W, Reyna VF, Fagerlin A, Lipkus I, Peters E. Clinical Implications of Numeracy: Theory and Practice. *Annals of Behavioral Medicine*. 2008;35(3):261-274. doi:10.1007/s12160-008-9037-8

21. Lipkus I, Samsa G, Rimer B. General performance on a numeracy scale among highly educated samples. *Medical decision making*. 2001;21(1):37–44. http://mdm.sagepub.com/content/21/1/37.short. Accessed August 9, 2016.

22. Garcia-Retamero R, Cokely ET, Ghazal S, Joeris A. Measuring Graph Literacy without a Test: A Brief Subjective Assessment. *Medical Decision Making*. 2016;36(7):854-867. doi:10.1177/0272989X16655334

23. Garcia-Retamero R, Galesic M. Chapter 4 Graph Literacy for Health. In: *Transparent Communication of Health Risks: Overcoming Cultural Differences*. New York, NY: Springer New York; 2013:53-65. doi:10.1007/978-1-4614-4358-2_4

24. Bekemeier B, Park S. Development of the PHAST model: generating standard public health services data and evidence for decision-making. *Journal of the American Medical Informatics Association*. 2018;25(4):428-434. doi:10.1093/jamia/ocx126

25. Colorafi KJ, Evans B. Qualitative Descriptive Methods in Health Science Research. *Health Environments Research & Design Journal*. 2016;9(4):16-25. doi:10.1177/1937586715614171

26. Morse JM. Critical Analysis of Strategies for Determining Rigor in Qualitative Inquiry. *Qualitative Health Research*. 2015;25(9):1212-1222. doi:10.1177/1049732315588501

27. Elo S, Kääriäinen M, Kanste O, Pölkki T, Utriainen K, Kyngäs H. Qualitative Content Analysis: A Focus on Trustworthiness. *SAGE Open*. 2014;4(1):215824401452263. doi:10.1177/2158244014522633

28. Thomas E, Magilvy JK. Qualitative Rigor or Research Validity in Qualitative Research: Scientific Inquiry. *Journal for Specialists in Pediatric Nursing*. 2011;16(2):151-155. doi:10.1111/j.1744-6155.2011.00283.x

29. Few S. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. 2nd ed. Burlingame, Calif.: Analytics Press; 2012.

30. Zakkar M, Sedig K. Interactive visualization of public health indicators to support policymaking: An exploratory study. *Online Journal of Public Health Informatics*. 2017;9(2). doi:10.5210/ojphi.v9i2.8000

31. Ware C. *Information Visualization Perception for Design*. 3rd ed. Waltham, Mass: Morgan Kaufmann; 2013.

32. Whitney H. *Data Insights: New Ways to Visualize and Make Sense of Data*. 1st ed. Amsterdam: Elsevier, Morgan Kaufman; 2013. http://www.barnesandnoble.com/w/data-insights-hunter-whitney/1112252098. Accessed February 29, 2016.

33. Rowley J. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*. 2007;33(2):163-180. doi:10.1177/0165551506070706

34. Few S. What Is Data Visualization? Visual Business Intelligence. https://www.perceptualedge.com/blog/?p=2636. Published May 4, 2017. Accessed January 23, 2019.

35. Fuhrmann S, Pike W. Chapter 31 - User-centered Design of Collaborative Geovisualization Tools. In: *Exploring Geovisualization*. Oxford: Elsevier; 2005:591-609. http://www.sciencedirect.com/science/article/pii/B9780080445311504498.

36. Torres R. *Practitioner's Handbook for User Interface Design and Development*. 1st ed. Upper Saddle River, New Jersey: Prentice Hall; 2001.

37. Baum NM, DesRoches C, Campbell EG, Goold SD. Resource allocation in public health practice: a national survey of local public health officials. *Journal of Public Health Management and Practice*. 2011;17(3):265–274. http://journals.lww.com/jphmp/Abstract/2011/05000/Resource_Allocation_in_Public_Health_Practice___A.10.aspx. Accessed December 7, 2015.

38. Bekemeier B, Chen A, Kawakyu N, Yang Y. Local Public Health Resource Allocation. *American Journal of Preventive Medicine*. 2013;45(6):769-775. doi:10.1016/j.amepre.2013.08.009

39. Lovelace KA, Aronson RE, Rulison KL, Labban JD, Shah GH, Smith M. Laying the Groundwork for Evidence-Based Public Health: Why Some Local Health Departments Use More Evidence-Based Decision-Making Practices Than Others. *American journal of public health*. 2015;105(S2): S189–S197. http://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2014.302306. Accessed November 24, 2015.

40. Sellers K, Leider JP, Harper E, et al. The Public Health Workforce Interests and Needs Survey: The First National Survey of State Health Agency Employees. *Journal of Public Health Management and Practice*. 2015;21: S13-S27. doi:10.1097/PHH.0000000000000331

41. Tufte E. *The Visual Display of Quantitative Information*. 2nd ed. Cheshire, Connecticut: Graphics Press; 2001. http://www.edwardtufte.com/tufte/books_vdqi. Accessed January 20, 2016.

42. Nielsen J. 10 Usability Heuristics for User Interface Design. https://www.nngroup.com/articles/ten-usability-heuristics/. Published in 1995. Accessed January 19, 2016.

43. Benefits of User-Centered Design. Usability.gov. /what-and-why/benefits-of-ucd.html. Published December 1, 2017. Accessed September 26, 2018.

CHAPTER 5. CONCLUSION

SUMMARY

Throughout my Ph.D. dissertation research, I explored what can be learned from the literature through systematic review and from interviews with public health professionals, as a means to informing improvements in decision-making among public health professionals by developing an understanding and use of data through data visualization.

In my first paper (Chapter 2), I provided a systematic review exploring the literature that examines how visualization can impact decision-making for general populations. Fourteen papers were included in the final review. Even though the evidence is limited due to a deficit of theoretical and methodological strength, the studies suggested that interventions that include data visualization have a positive impact on cognitive and behavior change such as decision-making, attitude, motivation or perception. The evidence from this review demonstrates the benefits of data visualization and shows promise for its application to public health.

In my second paper (Chapter 3), I explored how public health professionals use data, information, and evidence for their practice, especially for their decision-making, and their current use of and preferences for data visualization. I found that public health leaders use data, information, and evidence from various resources daily for communication with co-workers, stakeholders, and the public and for decision-making regarding their programs and services in various settings. I also found that the public health professionals interviewed have a high level of interest in applying visualization skills or tools in their public health practice as well as preference regarding the features or types of data visualization that they want to have. For example, they preferred simple, easily understandable and appealing data visualizations that help communication with stakeholders and policymakers. However, they face systematic challenges that discourage adequate data, information, and evidence use.

In my third paper (Chapter 4), I examined how data presented to public health professionals via visualization was understood differently compared to data presented in a more traditional table format. Qualitative analysis of open-ended questions described comparison relationships that LHD participants found by looking at data in a table and a visualization format. Even with the small sample participating in this study, I found that there are apparent benefits to using data visualization such as making finding information from data easier and reducing errors in tracking data. This paper focused on how data visualization improves the “making sense of data” step – a step followed by decision-making processes, analyzing and summarizing data, translating data into information, and then synthesizing the information into knowledge.

Implications for Future Practice and Research

This dissertation suggests that public health professionals can gain various benefits from using data visualization. Prior literature suggests that data visualization improves decision-making by improving understanding, analysis, and interpretation of data.[19,20,29] Similarly, this dissertation described how public health professionals may be better able to discover and understand local public health needs and issues, resulting in more informed decision-making. Data visualization can also be used to communicate better between agencies, within an agency, and with stakeholders, policy-makers, or with the public at large. Through the use of visualization, public health professionals may be able to make more effective communication tools for use with policymakers when advocating for funding, programs, and services. This potential benefit of data visualization could help public health leaders overcome political and systematic challenges that can hamper evidence-based decision-making. These findings also support the PHAST Model and related research upon which this dissertation was, in part, built. In the PHAST Model, efforts to make data more accessible such as data visualization can facilitate the use of data for public health practitioners by helping them to discover public health needs.[18]

There are, however, several existing challenges and barriers that undermine public health professionals' use of data and data visualization in their decision-making, even though public health professionals have a desire to use more data and are aware of the importance of data use and the particular benefits of data visualization. Major challenges include political and policy-related influences on decision-making, inadequate funding for information systems, inadequate staff with relevant skills and knowledge, and insufficient or old data. Given these challenges, understanding the specific needs of public health professionals and considering other factors that might affect their use of data and visualization in decision-making should be considered. Such considerations could include the engagement of public health professionals in advancing public health information systems, developing data visualization tools from the beginning of the design process, or applying new technology (e.g., Informatics, Big data).

Future research should also examine how data visualization influences or facilitates better decision-making and communication using more refined and tailored data visualization tools. Involving public health professionals from the beginning of the design process is crucial because data visualization can be designed in the context of how the tool is used—in this case, public health task performance. Finally, as described in the PHAST Model, training is needed for public health leaders in how to create and use data visualization and data assessment tools in their daily work, including generating additional knowledge of their workflow and current data needs and data uses.

In conclusion, this dissertation presents the current evidence regarding data visualization and its applications for public health professionals. This work suggests that data visualization could be an effective approach for improving decision-making and communication among public health professionals.

APPENDIX

Appendix A. Literature Search Strategy

Literature searches started by breaking down the research question and conducting scoping searches on the PubMed database to identify search terms according to PICOS (Population, Interventions, Comparisons, Outcomes, and Study Design).

- Research question: Data visualization as a behavioral intervention for policy-related behavior changes in state and local public health
- Population(s)/Patient(s): public health leaders
- Intervention(s)/Treatment(s): data visualization
- Comparator(s): table form of result
- Outcome(s): policy related behavior change
- Study Design: Any study design
- Setting: community and public health system, or state and local health department

Next, using the developed search strategies, searches were conducted on PubMed, CINAHL Plus with full text, Web of Science, IEEE Xplore, PAIS international, and Academic Search Complete and retrieved primary studies. The publication period ranged from 2008 to 2018. Foreign language papers were excluded. Two hundred fifty-five articles from PubMed, 48 articles from CINAHL Plus with full text, 131 articles from Web of Science, 244 articles from IEEE Xplore, 35 articles from PAIS international, and 73 articles from Academic Search Complete were found. Three additional articles were identified by reviewing reference lists of highly relevant articles. Overall a total of 786 articles were found (19 including duplicates). Concept tables were constructed as I proceeded with the search on the databases.

Concept Table

PUBMED	Concept: VDQI/data visualization	Concept: policy-related behavior	Concept: public health

Mesh Terms	Visual display of quantitative information
	Data visualization
	Information visualization
		Health planning
		Health Planning Organizations
		Decision-making
		Decision Making, Organizational
		Policy Making
		Decision Support Techniques
		Behavior
		    Information Seeking Behavior/subheading
		Attitude
		    Attitude to Computers/subheading
		Comprehension
			Public health
			Community health
			Public Policy
			    Health policy/subheading
			Administrative personnel

CINAHL Plus	Concept: VDQI/data visualization	Concept: policy-related behavior	
CINAHL headings	Visual display of quantitative information Data visualization Information visualization	Health and Welfare Planning Health Services Needs and Demand Marketing National Health Programs State Health Plans Strategic Planning Decision making, organizational Explode Decision Support Techniques Explode Problem Solving Health Service Administration Information management Explode Knowledge Management Nursing Administration Organizational Policies Planning Techniques Explode Resource Allocation Shared Governance/subheading Policy Making Public Policy Behavior Attitude understanding	

Web of Science	Concept: VDQI/data visualization	Concept: policy-related behavior	Concept: public health
	Visual display of quantitative information Data visualization Information visualization	Health planning Decision-making Policy Making Decision Support Behavior Attitude Comprehension	Public health Community health Public Policy Administrative

PAIS international	Concept: VDQI/data visualization	Concept: policy-related behavior	
	Visual display of quantitative information Data visualization Information visualization	Health planning Decision making Policy Making Decision Support Behavior Attitude Comprehension Understanding	

Academic Search Complete	Concept: VDQI/data visualization	Concept: policy-related behavior	
	Visual display of quantitative information Data visualization Information visualization	Health planning Decision making Policy Making Decision Support Behavior Attitude Comprehension Understanding	

IEEE Xplore	Concept: VDQI/data visualization	Concept: policy-related behavior	
	Visual display of quantitative information Data visualization Information visualization	Health planning Decision making Policy Making Decision Support Behavior Attitude Comprehension Understanding	

Search strategies and retried articles

PubMed

No.	Search	Query	Record
15	filter	Publication dates: 10 years Species: Humans Languages: English	255
14	#11 and #12 and #13	(((((((((("Health Planning"[Mesh] OR "Health Planning Organizations"[Mesh]) OR "Decision Making"[Mesh:noexp]) OR "Decision Making, Organizational"[Mesh]) OR "Policy Making"[Mesh]) OR "Decision Support Techniques"[Mesh]) OR "Behavior"[Mesh:noexp]) OR "Information Seeking Behavior"[Mesh]) OR "Attitude"[Mesh:noexp]) OR "Attitude to Computers"[Mesh]) OR "Comprehension"[Mesh]) AND ((("Eurograph IEEE VGTC Symp Vis"[Journal] OR ("data"[All Fields] AND "visualization"[All Fields]) OR "data visualization"[All Fields]) OR ("Inf Vis"[Journal] OR ("information"[All Fields] AND "visualization"[All Fields]) OR "information visualization"[All Fields])) OR (visual[All Fields] AND display[All Fields] AND quantitative[All Fields] AND information[All Fields])) AND (((("Public Health"[Mesh] OR "Community Health Services"[Mesh]) OR "Public Policy"[Mesh]) OR "Health Policy"[Mesh]) OR "Administrative Personnel"[Mesh]) Search ((((((((((((("Health Planning"[Mesh]) OR "Health Planning Organizations"[Mesh]) OR "Decision Making"[Mesh:NoExp]) OR "Decision Making, Organizational"[Mesh]) OR "Policy Making"[Mesh]) OR "Decision Support Techniques"[Mesh]) OR "Behavior"[Mesh:NoExp]) OR "Information Seeking Behavior"[Mesh]) OR "Attitude"[Mesh:NoExp])) OR "Attitude to Computers"[Mesh]) OR "Comprehension"[Mesh])) AND (((data visualization) OR information visualization) OR visual display of quantitative information)) AND ((((("Public Health"[Mesh]) OR "Community Health Services"[Mesh]) OR "Public Policy"[Mesh]) OR "Health Policy"[Mesh]) OR "Administrative Personnel"[Mesh])	554
13	#13	Search (((("Public Health"[Mesh]) OR "Community Health Services"[Mesh]) OR "Public Policy"[Mesh]) OR "Health Policy"[Mesh]) OR "Administrative Personnel"[Mesh]	7476403
12	#12	Search ((data visualization) OR information visualization) OR visual display of quantitative information	25357
11	#1 OR #2 OR # 3 OR #4 OR # 5 OR #6 OR #7 OR #8 OR #9 OR #10	Search (((((((((("Health Planning"[Mesh]) OR "Health Planning Organizations"[Mesh]) OR "Decision Making"[Mesh:NoExp]) OR "Decision Making, Organizational"[Mesh]) OR "Policy Making"[Mesh]) OR "Decision Support Techniques"[Mesh]) OR "Behavior"[Mesh:NoExp]) OR "Information Seeking Behavior"[Mesh]) OR "Attitude"[Mesh:NoExp])) OR "Attitude to Computers"[Mesh]) OR "Comprehension"[Mesh]	587651
10	#10	Search "Health Planning Organizations"[Mesh]	4644
9	#9	Search "Comprehension"[Mesh]	12761
8	#8	Search ("Attitude"[Mesh:NoExp]) OR "Attitude to Computers"[Mesh] Schema: syn	49622
7	#7	Search "Information Seeking Behavior"[Mesh]	1904
6	#6	Search "Behavior"[Mesh:NoExp]	28651
5	#5	Search "Decision Support Techniques"[Mesh]	72675

4	#4	Search "Policy Making"[Mesh]	24120
3	#3	Search "Decision Making"[Mesh:NoExp]	86760
2	#2	Search "Decision Making, Organizational"[Mesh]	10989
1	#1	Search "Health Planning"[Mesh]	328428

CINAHL Plus

No.	Search	Query	Record
10	#10	filter added: English last 10 years	48
9	#7 AND #8	(#1 OR #2 OR #3 OR #4 OR #5 OR #6) AND (#7 AND #8)	50
8	#8	#1 OR #2 OR #3 OR #4 OR #5 OR #6	297,546
7	#7	data visualization OR information visualization OR visual display of quantitative information	224
6	#6	understanding	
5	#5	(MM "Behavior") OR (MM "Attitude")	
4	#4	(MM "Public Policy+") OR (MM "Policy Making")	148,330
3	#3	(MM "Organizational Policies") OR (MM "Information Management") OR (MM "Knowledge Management+") OR (MM "Nursing Administration+") OR (MM "Resource Allocation+") OR (MM "Shared Governance+") OR (MM "Planning Techniques+")	17,739
2	#2	(MM "Health Services Needs and Demand+") OR (MM "Marketing") OR (MM "National Health Programs") OR (MH "State Health Plans") OR (MM "Strategic Planning+") OR (MM "Health and Welfare Planning")	53,775
1	#1	(MM "Decision Making, Organizational") OR (MM "Decision Support Techniques+") OR (MM "Problem Solving+")	27,539

Web of Science

No.	Search	Query	Record
4	#1 AND #2 AND #3	#3 AND #2 AND #1 Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Last 5 years	131
3	#3	(TS=(health planning) OR TS=(decision making) OR TS=(policy making) OR TS=(decision support) OR TS=(behavior) OR TS=(attitude) OR TS=(comoprehension) OR TS= (policy related behavior)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Last 5 years	1020858
2	#2	(TS=(public health) OR TS= (community health) OR TS=(public policy) OR TS=(administrator)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Last 5 years	178631
1	#1	(TS=(data visualization) OR TS=(information visualization) OR TS=(Visual display of quantitative information)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Last 5 years	17576

PAIS international

No.	Search	Query	Record
4	#4	**Limited by:** year: 2008; 2009; 2010; 2011; 2012; 2013; 2014; 2015; 2016; 2017; 2018; Document type: Article	35
3	#3	#1 AND #2	41
2	#2	(Health planning) OR (Decision making) OR (Policy Making) OR (Decision Support) OR Behavior OR Attitude OR Comprehension OR Understanding	101580
1	#1	(data visualization) OR (information visualization) OR (visual display of quantitative information)	119

Academic Search Complete

No.	Search	Query	Record
3	#1 AND #2	(Health planning OR Decision making OR Policy Making OR Decision Support OR Behavior OR Attitude OR Comprehension OR Understanding) AND (#1 AND #2) Limiters - Published Date: 2005-2015, English	73
2	#2	Health planning OR Decision making OR Policy Making OR Decision Support OR Behavior OR Attitude OR Comprehension OR Understanding	2757392
1	#1	data visualization OR information visualization OR visual display of quantitative information	5266

IEEE Xplore

No.	Search	Query	Record
1	#1	Filter: 2008-2018 Index term: data visualization ((data visualization OR information visualization OR visual display of quantitative information) AND (Health planning OR Decision making OR Policy Making OR Decision Support OR Behavior OR Attitude OR Comprehension OR Understanding) AND (Public Health OR Health))	244

Appendix B.

Email inviting potential subjects to participate in this research study

Dear ________
The Public Health Activities and Services Tracking (PHAST) Project at the University of Washington would like to invite you to participate in a study examining data use and needs for public health practice. This study is part of a grant from the Robert Wood Johnson Foundation to facilitate state and local health systems' use of relevant, standardized data. These data allow comparisons at the local level that would not otherwise be possible, to identify and address variation in the delivery of public health services.

Your perspective and insights can help PHAST understand local public health data use and needs. If you choose to participate, PHAST study staff will arrange a phone interview during which you will be asked questions about the use of and need for data in your public health practice setting. The questions focus on how data fits into your workflow processes, how data are used for decision making, and what challenges your agency faces that could be solved with access to the relevant data. The interview will be between 30 and 60 minutes depending on your responses. An audio recording of the interview will be made to ensure no details from your responses are lost.

If you would like to participate or would like to learn more about the interviews and the work PHAST is doing, please contact XXX: XXXXX@uw.edu.

Appendix C.

State/Local Health Department Leader Interviews: Questions on Data Use

Key Informant Interviewee		**Date**	
Interviewer(s)		**Phone Number**	

Opening Script –
Welcome/Introduction:

My name is _________ and I am [role with the PHAST project]. Thank you again for responding to the email invitation and for taking the time to have this conversation. With these interviews, we are trying to understand how people use data, what they need, and how they hope to use data. We want to find out about the value of standardized data depicting public health services and activities at the local level. The aim of this research is to make data accessible and useful for practitioners based on their actual experience.

PHAST is developing a system for putting data into action. We are working with public health partners nationwide to learn about how data are used, and what those partners' needs are. As more states adopt the uniform set of measures to satisfy data needs, public health practitioners working at the local level will gain access to a greater wealth of relevant information to inform their work.

Today's interview should take at most 45 minutes. Is that still going to work for you? (Y/N) As a reminder, the interview will be recorded. This is only to make it easier for us to transcribe your responses and to ensure we heard you accurately. After transcription, the recording will be erased. Do you have your permission to record this interview? (Y/N) Okay, we'll go ahead and start recording.

Q#	QUESTION	RESPONSE
We're going to start with some questions about how you work and make decisions, and how you might use data to shape those decisions.		
1	Broadly tell us about your agency or your unit within a larger agency, and the kind of work you do that involves decision making. *Probe: For example, can you walk us through a typical example of decision making?* *Or, can you describe a recent situation in which you needed to make a decision about public health activities?*	
2	Do you use data or information (such as internal or external data, tables, charts, medical records, texts, reports, scientific journal, etc.) when you make decisions?	

Now we will ask you for some details to follow up on how you use data in decision making or why you don't use data in decision making.		
3a	**If yes**, please tell us more details. What type of data and information you use when you make decisions? How do you use data and information for your decision making? *Probe: Do you have data or information that you check yearly (or regularly) to see trends?* *If so, which data or information? How did you choose sources of those? What do you look for from the data? (e.g. trends or comparison?)* *Probe: How does knowing this data and information influence your decision making about programs and services?* *Probe: How relevant to your jurisdiction is the data that you have access to?*	
or 3b	**If not**, why don't you use data to influence your decisions? *Probe: How could you use data to influence your decisions?* *Probe: How relevant to your jurisdiction is the data that you have access to, but are not using?*	
Now we're going to ask some questions about data you'd like to have and about challenges data could solve for you.		
4	What types of data or information do you need that either you don't currently have, or you've found the data are incomplete? *Probe: Do you have an example of a situation in which you have data, but you need more details and information than that data currently offers?* *Probe: What are some challenges to getting the data you need? (e.g. lack of technical expertise, lack of staff to use it, hard to find right data, administrative, etc.) If you have workarounds to those challenges, what are they? Ideally how would accessing data look if you didn't have to do the workaround?*	
5	Can you tell us about a time when you were unable to access data or it was missing, which created a challenge when addressing a public health concern, or led to a missed opportunity for your agency?	
6	If you had information about what programs LHDs around the state or the country are working on, would that be useful to you, and why or why not?	

7	How could standardized data (which allows you to compare an LHD to similar LHDs across states) influence or support any efforts you are engaged in or planning toward accreditation, community health assessment, or quality improvement?	
8	Thinking only of visual elements, what makes for good and useful data visualization, to you in your work? *Probe: Can you describe a data visualization tool that has been useful and what about it was useful to you?* *Probe: Examples of data visualization elements include bar graphs, time trends, interactive displays, maps, infographics, tables, etc. It could also be how a site is designed or some other visual component.*	
9	Is there anything else you'd like to share about the work you do that will help us understand your use of data?	
10	Do you have any suggestions for how we could improve our questions or the interview process?	