MEDINFO 2013
C.U. Lehmann et al. (Eds.)

doi:10.3233/978-1-61499-289-9-697

Visualization of Infectious Disease Outbreaks in Routine Practice

David Karlsson[a], Joakim Ekberg[a,b], Armin Spreco[b], Henrik Eriksson[a], Toomas Timpka[a,b]

[a] *Department of Information and Computer Science, Linköping University, Sweden*
[b] *Department of Medical and Health Sciences, Linköping University, Sweden*

Abstract

Throughout the history of epidemiology, visualizations have been used as the interface between public-health professionals and epidemiological data. The aim of this study was to examine the impact of the level of abstraction when using visualizations on routine infectious disease control. We developed three interactive visualization prototypes at increasing levels of abstraction to communicate subsets of influenza outbreak surveillance information. The visualizations were assessed through workshops in an exploratory evaluation with infectious disease epidemiologists. The results show that despite the potential of processed, abstract, and information-dense representations, increased levels of abstraction decreased epidemiologists' understanding and confidence in visualizations. Highly abstract representations were deemed not applicable in routine practice without training. Infectious disease epidemiologists' work routines and decision-making need to be further studied in order to develop visualizations that meet both the quality requirements imposed by policy-makers and the contextual nature of work practice.

Keywords:

Public health, infectious disease outbreaks, epidemiology, data display, health information systems, visualization.

Introduction

Throughout the history of epidemiology, visualizations have been used as the interface between public-health professionals and epidemiological data. The two main components of modern visualizations are graphics and interaction. Graphics take advantage of the flexibility of human perception in identifying patterns coupled with adaptive decision-making mechanism [1]. Interactivity provides the coalescing layer of exploration that is required to support the dialogue between users and data. Over time, health professionals have learned to integrate increasingly complex visualization methods into the management of population health. There is a body of research describing how changes in technology, work, and organizational forms have led to and made visualization part of epidemiological practice. In infectious disease control, local disease maps have been used since Seaman used "spot maps" in 1798 to trace yellow fever cases in New York [2,3]. A few decades later, Farr introduced vital statistics, that is, abstract mathematical methods for detecting associations between mortality and social circumstances [4]. Since then, the appropriate level of abstraction to be used in analyses of infectious disease outbreaks has been debated. Such interplay between analogue and abstract representations of increasing complexity is displayed in modern public-health practice. The epidemiological evidence on concurrent health threats is understood to underdetermine the tested hypotheses, providing neither complete proof nor disproof [5]. To be able to judge the weight of epidemiological evidence, both scientific and extrascientific values must be included in the assessment [6,7], and the nature of this evidence and these values may shift between contexts. Modern visualization technologies provide extended possibilities for epidemiologists to explore and judge underdetermined evidence. Furthermore, integrated health information systems have opened new potentials for collection and management of community-level epidemiological data on infectious diseases [8,9]. In addition, advanced methods for detection of infectious disease outbreaks have been developed [10,11]. The aim of this study was to examine the impact of the level of abstraction when using visualizations in routine infectious disease control.

Materials and Methods

A two-step approach was used for the study. Three interactive visualization prototypes at increasing levels of abstraction were developed to communicate subsets of surveillance information about an influenza outbreak. These visualizations were then assessed through workshops in an exploratory evaluation [12] involving senior infectious disease epidemiologists.

Study setting

The data repository connected to the Electronic Health Record (EHR) systems at Östergötland County Council, Sweden [9], collects clinical data from all primary care centers (PHCs), hospitals, and clinical laboratory units in the county. The repository is also connected to the Swedish national telenursing system Healthcare Direct, an around the clock, nurse-led, telephone advice service. The system has strong similarities to the NHS Direct system in the UK and Health Direct in Western Australia. The specially trained nurses use a computerized decision-aid program and an EHR system for every call, and are, according to Swedish law, obliged to keep a record of each patient interaction. Upon answering a call, the nurse collects the patient's unique personal identification number and opens a new call record. Over the course of the conversation, the nurse consults one or more online guidelines; that is, electronic documents about conditions or symptoms, such as 'fever (adult)' or 'dizziness'. A chief complaint from a fixed-field terminology register is recorded after each call.

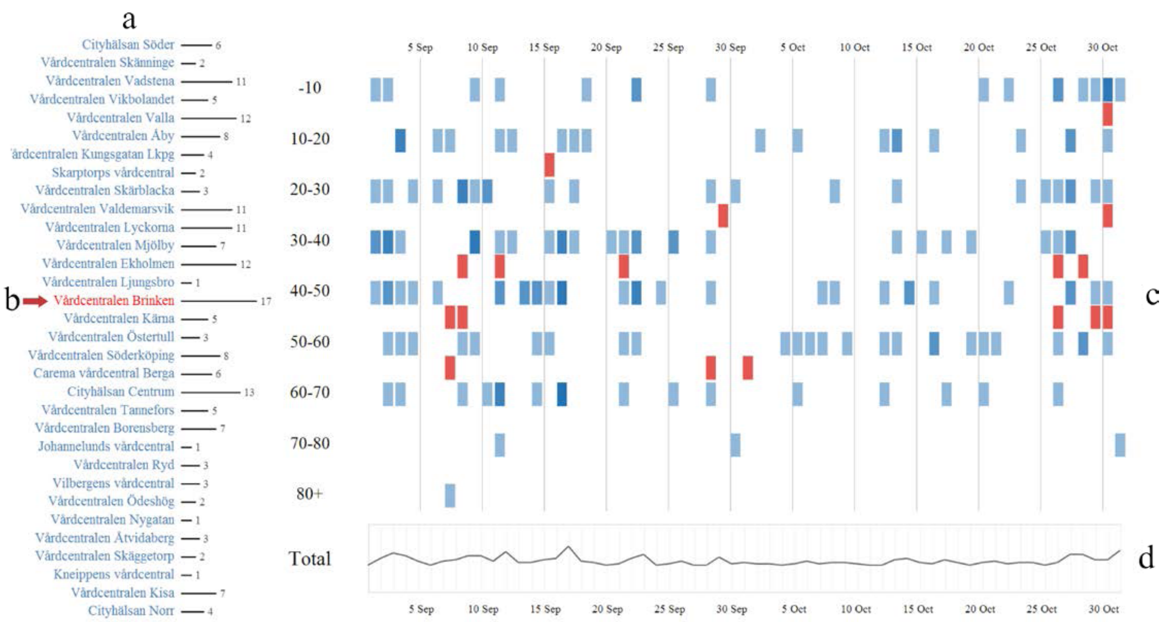

Figure 1- Disease map prototype (P1) based on a listing of PHCs (a). P1 provides the possibility to select a specific PHC (b), view the local case incidence per day and age group (c), and compare the local data to a time graph showing the incidence rate for the entire county (d).

Visualization data

Aggregated population data on sex, age, and residence were collected from Statistics Sweden. Data from Östergötland residents who had been clinically diagnosed with influenza or had contacted the telenursing service were identified from the data repository connected to the countywide data repository. Influenza cases were identified by the ICD-10 codes for influenza (J10.0, J10.1, J10.8, J11.0, J11.1, J11.8). For individuals having received a diagnosis at both primary and secondary levels of care, only the latter record was used for the analyses. Influenza-like-illness-related telenursing call cases were identified by the chief complaint codes associated with influenza symptoms (dyspnea, fever (child, adult), cough (child, adult), sore throat, lethargy, syncope, dizziness, and headache (child, adult)) from the fixed-field terminology register.

For this study, two data sets were prepared. An EHR data set was compiled, consisting of influenza-related diagnoses from the county-wide EHR. The data set also included information on age and the PHC the patient visited.

Second, a telenursing data set was gathered, consisting of information about the individuals that had called the telenursing service, categorized into chief complaints and age group (preschool, school, adult, and elderly).

Ethics Statement

The study was based on administrative public health databases established to systematically and continuously develop and secure the quality of service. According to Swedish legislation (SFS 2008:355), personal identification data were removed from the records. The Regional Research Ethics Board in Linköping, Sweden (dnr. 2012/104-31) approved the study design.

Design of prototypes

The prototypes were developed using Data-Driven Documents (D3) (http://d3js.org/, accessed December 8th, 2012), a JavaScript library for production of visualizations from digital data. D3 manipulates HyperText Markup Language (HTML), Scalable Vector Graphic (SVG), and Cascading Style Sheets (CSS), and eanbles interactive and dynamic visualizations displayable in modern web browsers.

Disease map prototype (P1)

Prototype 1 (P1) shows a visualization concept based on spatio-temporal information using the EHR data set. The basic rationale was to display the progress of the disease burden for specific PHCs. The visualization is based on a list of all PHCs in the county (Figure 1a). For the infectious disease epidemiologists, a listing of the centers with their geographical uptake areas was assumed to convey the most affected locations during an outbreak. The interactive component of P1 consists of PHC selection, providing the user the case incidence for each day and age group (Figure 1b) and cumulative incidence for the local geographical area (Figure 1c). P1 also provides a time graph representation of the daily incidence rate for the entire county (Figure 1d).

Descriptive abstract prototype (P2)

Prototype 2 (P2) uses a descriptive visualization concept to present the telenursing dataset using a matrix where the rate of telenursing calls per day in the county is displayed, separated by the chief complaints (symptoms) recorded (Figure 2a). P2 enables a detailed representation of symptoms, but also a more aggregated, abstract representation using color saturation to represent differences in rates. The interactive component of P2 consists of day selection (Figure 2b), updating both the total number of calls per chief complaint and the age distribution per chief complaint (Figure 2c).

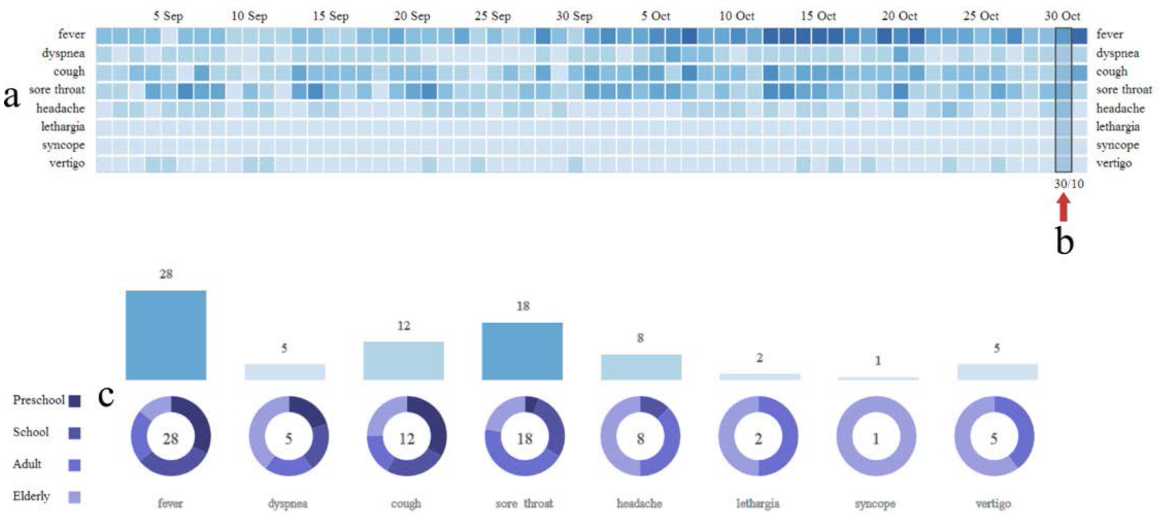

Figure 2-Descriptive abstract prototype (P2) displaying rates of telenursing calls divided by chief complaints (symptoms) (a). Any day can be selected (b) to obtain a breakdown of the calls by chief complaint and age group for that day (c).

Analytic abstract prototype (P3)

Prototype 3 (P3) displays an outbreak detection visualization concept based on the telenursing dataset. As in P2, the rates of chief complaints per day are displayed in a matrix (Figure 3a). P3 expands the abstract component included in P2 by utilizing an algorithm for prediction of call rates for the next 5 days (Figure 3b). In addition, highlights in red indicate elevated rates of chief complaints compared to a computed baseline. The interactive component of P3 consists of day selection, which shows the user recorded (or predicted) calls and their age distribution for the selected day, divided by chief complaint (Figure 3c).

Data collection

Two workshops using think-aloud techniques [13] were conducted with senior infectious disease epidemiologists experienced in epidemiological surveillance. Qualitative data were gathered from the workshops through audio-recordings and field-notes. The think-aloud sessions were complemented with questions about usefulness to the participants' work practices. Before introduction of the prototypes, the participants were instructed to construe them as possible tools for use during influenza outbreaks. Each prototype was shown to the participants, who were asked to think aloud when exploring and interacting with them. After every think-aloud session, open questions were then asked about the possible usefulness of the visualizations in their routine practice.

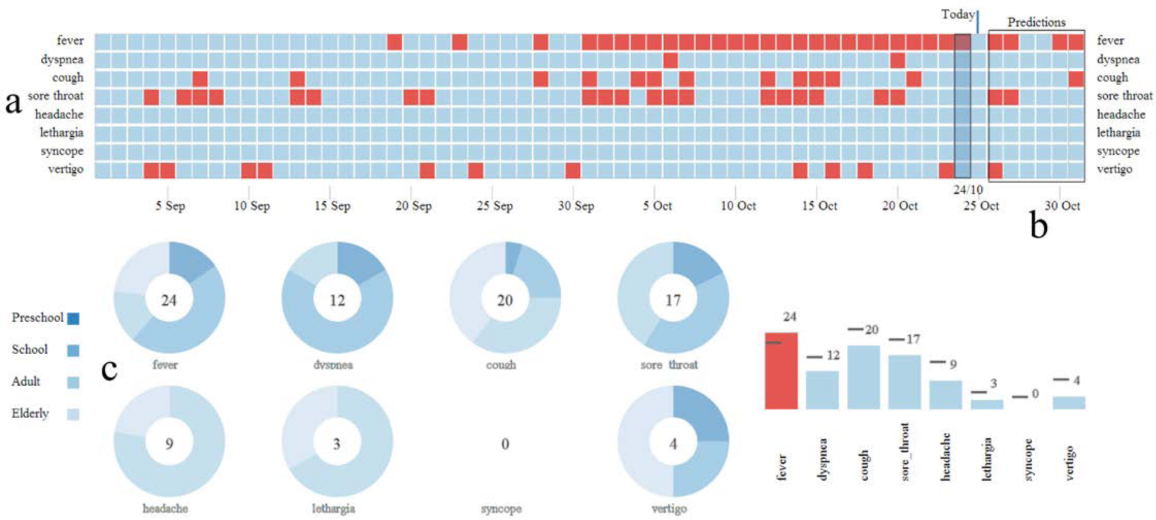

Figure 3-Analytic abstract prototype (P3). Prototype P2 is extended by prediction of call rates for the next 5 days (b,c).

If the usefulness of a specific part of the prototype was not covered in the answer then they were asked about it specifically.

Data analysis

The workshop data were transcribed and analyzed using content analysis [14]. For each comment, the feature in the visualization to which it applied was identified. The comments were categorized and further analyzed to infer what impact the participants thought the feature could have on their work practice.

Results

In general, participants had positive opinions and believed that most visualization features in the prototypes could, for different reasons, fill a function in their work practices.

Disease map prototype (P1)

The disease map prototype was considered a tool for both senior infectious disease epidemiologists and general practitioners. It was considered adequate because identifying the most affected PHCs was reported to give information about where the outbreak was most intensive and helped to visualize how the outbreak was evolving geographically.

One senior epidemiologist thought that the interactive possibility to click on a PHC and inspect the squares corresponding to incidence rates in the matrix was interesting. This functionality was thought to be useful for both infectious disease epidemiologists and general practitioners because it could show the distribution of cases for every PHC. P1 could thus display where the outbreak had the most impact at any given time and help to establish a shared situational awareness about the propagation of the disease. Another senior epidemiologist reported that by identifying the most affected PHCs, it was possible for the infectious-disease epidemiologist to warn not yet intensively affected PHCs to initiate vaccinations, if vaccines were available. For general practitioners, it was thought to be useful to be able to identify whereto the outbreak was propagating and the case rates for their own PHC. This feature was believed to create local outbreak awareness, which was considered beneficial for local preparedness and response. The participants reported that the line graph was useful because it showed changes over time in the rate of infected (infectious) individuals. They explained that influenza outbreaks develop in waves and that it is important to know when the peak is reached.

Descriptive abstract prototype (P2)

Access to age groups in the descriptive abstract prototype was considered useful because the composition of age groups varies between counties, and people of dissimilar age both differ with regard to immunity and act in different ways during outbreaks. The participants intended to display the data by age because it can support planning of interventions and communication with the community. One senior epidemiologist explained that if there are numerous small children with fever showing in the visualization, then kindergartens and schools could be warned, and accordingly take actions that might curb the outbreak. Another senior epidemiologist reported that seeing distribution of age groups communicates much about how the health-care service is going to be affected by the outbreak. For example, if mainly young people were affected, health services would probably not be burdened to the same extent. Conversely, if the main portion of the infected population consisted of seniors, not only PHCs, but also hospitals, would be burdened more quickly.

Moreover, it was expressed that bar graphs and pie charts were more useful than the matrix, because more detailed data is usually only interesting on the closest past days during influenza outbreaks. However, one senior epidemiologist explained that the ability to inspect data from previous days through interaction can be useful, because it is always valuable to make comparisons between days.

The possibility to move back in time through interaction was explained as providing two benefits; one is that it can be hard to remember all the different distributions, and that the interaction can make work easier as it supports memory. The other benefit was that senior infectious disease epidemiologists can learn much if they can go back and retrospectively review the outbreak when it has ended.

Analytic abstract prototype (P3)

For the analytic abstract prototype, the participants at both workshops were of the opinion that it would be useful if the baseline was correct and could be trusted. However, they were not sure if they fully understood what the prototype said about the outbreak, other than that the levels were higher than they were supposed to be. The participants were of the opinion that seeing predictions about the approaching days would be useful, but only if the foundation for the predictions could be inspected, understood, and trusted.

They also thought that these more complex visualizations could be useful if the users had received specific training in using and interpreting the conveyed information. If the predictions show that the symptoms fever and cough are increasing in the county, this could be an indicator for an influenza outbreak, and by predicting it, would allow vaccinations and other precautions to be implemented before the outbreak disseminates fully. However, making such decisions in routine practice was thought to require trust in the information that is provided, which, in turn, would require additional training and experience.

Discussion

We set out to examine the impact of the level of abstraction when using visualizations in routine infectious disease control. The results indicate that for visualizations to be useful in routine practice, the level of abstraction must be chosen to support the epidemiologist in their work practice. The participants described how (procedurally) and why (in terms of the quality of their decisions) the visualizations were associated to their decision making. Although the participants acknowledged the value of including symptoms and a summative abstraction of the progress of influenza spread, the geographical component and readily comprehensible setup of the disease map prototype (P1) was considered superior in terms of transparency and close ties to routine practice. Thereby, our results suggests that visualizations of data from infections disease outbreaks should support decisions in terms of why, how, and what to do in practice, similar to what is required from decision support systems [15]. The descriptive component was delivered in P1 by describing incidence rates and geographical location. In P2 and P3, the descriptive component was increasingly processed into a more abstract, summative representation, making participants less confident in their ability to use the visualization.

A wealth of benefits from using visualization in public health and infectious disease control has previously been reported [8,16]. The findings from this study add nuances to this picture. Participants in this study reported uncertainty about the meaning of highlighting and what the baseline meant in the outbreak predictions. They experienced insecurity in their interpretation of the visualization. These observations are in accordance with reports from studies of visualization use in clinical settings, such as in neurosurgery planning [17]. In the latter setting, it was found that abstract and information-dense visualizations, although potentially more powerful, failed to support the clinicians as intended. Therefore, introducing results of complex computations into visualizations, such as outbreak progress predictions, may be detrimental when the computations cannot be understood or do not comply with heuristics and other decision making practices. In addition to making the visualizations more transparent, it would be beneficial to integrate interpretation of more complex analyses in the training of infectious disease epidemiologists, thereby helping them explore and judge underdetermined evidence [6,7].

This study has several limitations that need to be considered when interpreting the results. The use of a qualitative approach based on workshops, where the purpose was to elicit rich accounts of the experience of infectious disease epidemiologists, may have led to important aspects and areas being omitted completely in the evaluation data. However, the similarity of the accounts collected from the two workshops suggests that the findings can be regarded as trustworthy. Furthermore, the prototypes do not represent all types of analogue and abstract representations of infectious disease data, and other principles not implemented may have benefited the participants. Thus, the conclusions drawn from this study are limited by the range of visualizations used in the prototypes.

Conclusion

Modern visualization technologies provide extended possibilities for epidemiologists to explore and judge underdetermined evidence. This study shows that despite the potential of processed, abstract, and information-dense representations, increased levels of abstraction decreased epidemiologists' understandings and confidence in visualizations. Highly abstract representations of infectious disease outbreaks were deemed not immediately applicable in routine practice without preceding training. Infectious disease epidemiologists' work routines, and especially their decision-making, need to be further studied in order to develop visualizations that meet both the quality requirements posed on infectious disease control by policymakers and the contextual nature of work practice.

Acknowledgments

This study was supported by grants from the Swedish Civil Contingencies Agency (2010-2788), VINNOVA (2011-03231), and the Swedish Science Council (2008-5252, 2009-6291). The funders had no role in study design, data collection and analysis, decision to publish, or manuscript preparation.

References

[1] Ware C. Information Visualization: Perception for design 3rd ed. Waltham: Morgan Kaufmann, 2012

[2] Stevenson L. Putting disease on the map: the early use of spot maps in the study of yellow fewer. J History Allied Sci 1965;20:226-61.

[3] Shannon G. Disease mapping and early theories of yellow fewer. Prof Geographer. 1981;33:221–7.

[4] Farr W. Vital statistics: memorial volume of selections from the reports and writings. Bull WHO 2000;78:88-96.

[5] Bradford-Hill, A. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine 1965;58: 295-300.

[6] Weed, DL. Underdetermination and Incommensurability in Contemporary Epidemiology. Kennedy Institute of Ethics Journal. 1997; 7(2):107-124.

[7] Dowie J. Analysing health outcomes. J Med Ethics. 2001;27(4):245-50.

[8] Cheng CK, Ip DK, Cowling BJ, et al. Digital dashboard design using multiple data streams for disease surveillance with influenza surveillance as an example. J Med Internet Res. 2011 Oct 14;13(4):e85.

[9] Timpka T, Eriksson H, Gursky EA, et al. Requirements and design of the PROSPER protocol for implementation of information infrastructures supporting pandemic response: a Nominal Group study. PLoS One. 2011 Mar 28;6(3):e17941

[10] Kass-Hout TA, Xu Z, McMurray P, et al. Application of change point analysis to daily influenza-like illness emergency department visits. J Am Med Inform Assoc. 2012 Nov 1;19(6):1075-81. doi: 10.1136/amiajnl-2011-000793.

[11] Hoen AG, Keller M, Verma AD, et al. Electronic event-based surveillance for monitoring dengue, Latin America. Emerg Infect Dis. 2012 Jul;18(7):1147-50. doi: 10.3201/eid1807.120055.

[12] Ellis G, Dix A. An Explorative Analysis of User Evaluation Studies in Information Visualisation. New York: BELIV, 2006

[13] Ericsson K, Simon HA. Protocol Analysis: Verbal reports as data. Rev ed. Cambridge.: The MIT Press, 1993

[14] Krippendorff K. Content analysis: An introduction to Its methodology. Thousand Oaks: Sage Publications, 2004.

[15] Abbott PA, Zytkowski M. Supporting Clinical Decision Making. In: Englebardt SP, Nelson R. Health Care Informatics. St. Louis, Missouri: Mosby, 2002

[16] Bures V, Otcenáskova T, Antos K. A proposal for a computer-based framework of support for public health in the management of biological incidents: the Czech Republic experience. Perspect Public Health 2012;132(6):292-8

[17] Dinka D, Nyce JN, Timpka T, Holmberg K. Adding value with 3D visualization and haptic forces to radiosurgery-a small theory-based, quasi-experimental study. J Med Syst. 2006 Aug;30(4):293-301

Address for correspondence

David Karlsson
Department of Information and Computer Science
Linköping University
SE-581 83 Linköping, Sweden
+46 (0)70 2212948
davka442@gmail.com