BRIEF REPORT

Use of Digital Pens for Rapid Epidemiologic Data Collection During a Foodborne Outbreak Investigation

Abigail A. Mathewson, DVM, MPH; Elizabeth R. Daly, MPH; Steffany J. Cavallo, MPH; Adnela Alic

ABSTRACT

Objective: Public health investigations require rapid assessment, response, and initiation of control measures. In 2012, the New Hampshire Department of Health and Human Services used digital pens to rapidly acquire epidemiologic data during a gastrointestinal illness outbreak.

Methods: Menus were obtained and a standard questionnaire was administered to exposed persons using digital pens. Questionnaire data were downloaded into an electronic file for analysis.

Results: Sixty-nine (74%) of 93 exposed persons completed a questionnaire. Of 6389 data entries made on digital paper, 218 (3%) required correction; of these, 201 (92%) involved a free-form variable and 17 (8%) involved a check-box variable. Digital pens saved an estimated 5 to 6 hours of data-entry time.

Conclusions: This outbreak provided an opportunity to assess the value of digital pens for decreasing data-entry burden and allowing more timely data analysis in an emergent setting. Depending on the size of the outbreak and complexity of the survey, there is likely a threshold when use of digital pens would provide a clear benefit to outbreak response. As new technology becomes available for use in emergency preparedness settings, public health agencies must continuously review and update response plans and evaluate investigation tools to ensure timely disease control and response activities. (*Disaster Med Public Health Preparedness*. 2015;9:349-353)

Key Words: outbreak response, digital pen, rapid survey, foodborne outbreak, investigation methods, public health preparedness

Public health investigations require rapid assessment and response in order to institute control measures quickly and prevent additional illness. In the United States, state and local health departments are required to establish and maintain public health emergency preparedness and response capabilities[1] in order to receive specific federal funds for such purposes. Funded jurisdictions are regularly assessed and ranked on their ability and readiness to respond to large-scale emergencies and natural disasters.[2] As such, public health agencies must continuously review and update response plans and evaluate investigation tools to ensure timely disease control and response activities.

Collection of epidemiologic information is a key component of a public health investigation and often involves surveying people to determine their health status and identify exposures that may pose a threat to others. Historically, the most common approaches to conducting public health surveys were by mail, in person, or via telephone, and responses were documented using pen and paper, followed by data entry and then analysis. Over the past decade, as new technologies have been developed and accepted into public health practice, electronic methods for rapid data acquisition have emerged. Some electronic methods are now frequently used, such as online surveys,[3] but others are less commonly used, such as computer-assisted telephone interviewing.[4] More recently, the use of tablets or personal data assistants have been explored for data entry in the field.[5] Digital pens are another technology available to minimize or eliminate data-entry burden by electronically recording data handwritten on specially printed paper and uploading questionnaire responses into an electronic file for data analysis. Use of digital pens has been reported in clinical settings,[6-8] but use during public health responses has been infrequently described.[9]

Some of the most common types of public health investigations are foodborne illness outbreaks; approximately 1200 occur in the United States each year.[10] Because of their frequency, state and local health departments are provided with an opportunity

to exercise preparedness and response plans during routine investigations of foodborne illness outbreaks. In February 2012, the New Hampshire Department of Health and Human Services (DHHS) investigated gastrointestinal illnesses among wedding attendees to assess extent and possible sources of illness. During the investigation, the DHHS piloted the use of digital pens for data collection. The goal of the pilot was to evaluate the digital pens for ease of use in an outbreak setting, establish a model to increase investigation efficiency and more timely availability of epidemiologic data, and provide insight into best practices for use of digital pens for public health response in the future.

METHODS

After the outbreak was identified, menus for all wedding-related events were obtained, and a standard questionnaire was created using Capturx digital pen software (Adapx, Inc, Seattle, Washington) and printed for investigator use. The questionnaire asked attendees about 72 different food exposures, symptoms of gastrointestinal illness, visits to health care providers, and specimen submission for pathogen testing. The questionnaire included 81 check-box and 25 free-form variables in total. Contact information for wedding attendees was obtained, and the questionnaire was deployed by telephone from February 8 through February 18, 2012.

Eleven staff members were trained to use the digital pens to conduct phone interviews with exposed persons. Training included instruction on use of the paper survey instrument, operation of the digital pens, and data security and confidentiality. The digital pens were used in a similar manner to a standard pen with the added functionality of using a unique dot matrix printed on each paper interview to record each respondent's answers. After interview completion, the pens were placed into a docking station and questionnaire data were uploaded into an electronic Microsoft Excel (Microsoft Corporation, Redmond, Washington) file using the Capturx digital pen software. Once data from the pens were made available in the electronic file, corrections were required to remove duplicate records from prior data transmissions and correct critical data-capture errors before analysis of data. The number of corrections required was documented, and data were analyzed to assess overall accuracy and completeness of data. Data analyses were performed using EpiInfo 7 (Centers for Disease Control and Prevention, Atlanta, Georgia) and Microsoft Excel.

RESULTS

Of 93 exposed wedding attendees, 69 (74%) completed a standard questionnaire. Of those interviewed, 29 cases were identified as having gastrointestinal illness for an attack rate of 43%. Norovirus was identified as the etiologic agent, and illness was associated with consumption of vegetable dip served at the wedding welcome reception.

After content of the questionnaire was determined, the survey was created using digital pen software, a process that took approximately 3 hours. After the electronic form was developed, questionnaire administration was initiated, which required a specialized printing process on a standard office printer and training of staff members. The digital pens saved approximately 5 to 6 hours of data-entry time based on estimates of 69 surveys at 4 to 5 minutes each.

To determine the number of corrections needed, both the electronic and hard-copy records were examined for each patient interview. Of 6389 data entries made using digital pens, 5510 (86%) were check-box variable entries and 879 (14%) were free-form variable entries. Overall, 218 (3%) entries required correction; 201 (92%) corrections involved a free-form variable and 17 (8%) involved a check-box variable. Among free-from variables, 24 (96%) of 25 variables required at least 1 correction (range = 1-27 corrections per variable), and corrections were required in 23% of all free-form variable entries recorded electronically. Among check-box variables, data corrections were required in 13 (16%) of 81 variables (range = 1-4 corrections per variable). Corrections were required in <1% of all check-box variable entries recorded.

DISCUSSION

Digital pens were successfully used to rapidly collect epidemiologic information during a foodborne outbreak investigation in New Hampshire. Upon notification of the outbreak, the DHHS was able to pilot the use of digital pens in an outbreak setting to determine their efficacy in potentially reducing the burden of data entry, thereby allowing for more timely data analysis and identification of the source of an outbreak.

A key aspect of assessing the use of digital pens was how well the handwritten information transferred into an electronic format. Each paper survey instrument was reviewed and compared with the database for accuracy and completeness, and as expected, the majority of necessary corrections involved free-form variables. All records required at least 1 correction, but the number of corrections were dependent on the method of data collection on the paper form. Among the most common reasons for corrections were style and format of information written on the surveys. Data transferred from the pens encountered the most difficulty with special characters (eg, w/, &, <, >, fractions), letters that were partially written outside the confines of the allotted space, and places where a second line of text was begun above the initial line in a free-form field box (eg, starting text at the bottom of the box and then trying to write above the initial entry when space became limited). Additional problems occurred as a result of the formatting of the spreadsheet into which the data were downloaded. Cells that were specifically formatted for an anticipated format of response that was not subsequently

https://doi.org/10.1017/dmp.2015.43 Published online by Cambridge University Press

FIGURE 1

Comparison of Theoretical Standard Data Entry Flow to Digital Pen Data Flow

Timeline comparing standard and digital pen epidemiologic investigation tools using 50 survey example (standard tools preferred)

Standard Epidemiologic Investigation	Develop Survey Content		Create Database for Entry of Survey Data			Patient Interviews			Data Entry of Completed Surveys			Clean Data	Analyze Data											
Timeline (hours)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Digital Pen Epidemiologic Investigation	Develop Survey Content		Create Digital Form			Patient Interviews			Dock Pens Upload Data		Clean Data				Analyze Data									

Timeline comparing standard and digital pen epidemiologic investigation tools using 100 survey example (tools perform similarly)

Standard Epidemiologic Investigation	Develop Survey Content		Create Database for Entry of Survey Data			Patient Interviews			Data Entry of Completed Surveys						Clean Data	Analyze Data								
Timeline (hours)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Digital Pen Epidemiologic Investigation	Develop Survey Content		Create Digital Form			Patient Interviews			Dock Pens Upload Data		Clean Data					Analyze Data								

Timeline comparing standard and digital pen epidemiologic investigation tools using 200 survey example (digital pens preferred)

Standard Epidemiologic Investigation	Develop Survey Content		Create Database for Entry of Survey Data			Patient Interviews			Data Entry of Completed Surveys												Clean Data		Analyze Data	
Timeline (hours)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Digital Pen Epidemiologic Investigation	Develop Survey Content		Create Digital Form			Patient Interviews			Dock Pens Upload Data		Clean Data						Analyze Data							

provided by the investigator resulted in a nonsensical or incorrect entry in the data field, such as date or number fields.

Suggestions for improving data quality when using digital pens were identified by comparing the necessary corrections in the electronic database to what was recorded on the paper survey instrument. These included use of check-box variables when possible, creation of larger free-form boxes to accommodate longer answers, and use of a general text format for all free-form variables to prevent population of variables with nonsensical data when downloaded into the spreadsheet. More in-depth training for staff conducting interviews on the proper method for using free-form variables may also improve data quality and reduce the number of errors observed. Training should include instruction to avoid the use of special characters and to begin writing at the top left corner of a free-form box and then write below the first line if additional lines are needed.

A limitation of this report is that the digital pen method and the standard method (standard pen and paper followed by data entry) were not conducted in parallel, and time savings could only be estimated based on other epidemiologic data-acquisition experiences. Data capture and upload from the digital pens is estimated to have saved approximately 5 to 6 hours of data-entry time. This estimate is based on the time it took to manually enter a survey into a prepared database versus downloading the data from the pens. Time required to dock the pens and import the data was less than 10 minutes per pen. Not included in this estimate of time savings is the amount of time it took to manually correct the survey records before epidemiologic analysis as that parameter was not recorded in real time, nor was the length of time recorded for training purposes. It is expected that as investigators become more familiar with this technology and adept at using it, both the time required to correct uploaded data and perform training will be negligible. Additionally, in the standard approach, the survey can be quickly generated and made available for the initiation of interviews, allowing for development of the data-entry tool while interviewers are collecting the data. With digital pens, the printed survey is generated directly from the data-entry tool so interviewing cannot commence until the data-entry tool is created, which, in this investigation, delayed interviewing by approximately 3 hours. Though digital pens saved data-entry time, time needed to set up the system, train staff, and correct errors may

https://doi.org/10.1017/dmp.2015.43 Published online by Cambridge University Press

TABLE 1

Consideration Factors in Deploying Digital Pens

Factor and Response	Manual Data Entry	Digital Pens
Number of expected respondents: Determine threshold for the number of surveys that would result in a data-entry time savings		
Fewer surveys	√	
More surveys		√
Length and complexity of survey: Quality of digital pen survey data is improved if most variables are presented as a check box versus free-form.		
Simple survey with many check-box fields		√
Complex survey with many free-form fields	√	
Location of interview staff: Staff who will conduct interviews must be centrally located in order to access the digital pen and printed surveys.		
Centralized		√
Decentralized	√	
Staff preference for handwriting during survey administration: Staff who are uncomfortable with computer interfaces may prefer to take handwritten notes.		
Handwriting preferred		√
Staff uses computer interfaces with ease[a]	√	

[a]Consider direct entry of interview data into a web-based survey, local database, or tablet in real time during the interview.

have made up the difference in their initial deployment. There is likely a threshold for the number of anticipated surveys that would result in data-entry time savings from using the digital pens (Figure 1). This threshold would be affected by staffing resources available for response, staff level of expertise with the technology, and the consideration that manipulating the digital pen database may require more data analysis expertise, while traditional survey methods require greater data-entry staff availability. Correspondingly, factors to consider when choosing to use digital pens in an event include number of respondents, length and complexity of survey, whether interviewers are centrally located or in various response locations, and staff preferences (Table 1).

Similar to other electronic data acquisition methods (eg, tablet and web surveys), digital pens offer a particular benefit over standard methods for field deployments, such as going to a community to do door-to-door interviewing or when interviewing at the site of an outbreak or event. The near real-time availability of data would also provide a clear benefit in a mass vaccination, point of dispensing, or mass shelter incident. Rapid screening of an incoming clinic or shelter population would allow responders to more quickly respond to the needs of attendees (eg, additional medical screeners for complicated patients or specific food, shelter, or medication needs). This is a clear benefit over the more traditional method of collecting data through the use of paper survey instruments with manual entry after the event. As patient data are collected, they can be imported into a master database as frequently as desired by the agencies coordinating the event for evaluation and analysis rather than consuming the time of additional staff with data entry as time allows.

CONCLUSIONS

As new technology becomes available for use in the public health setting, public health agencies must continuously review and update response plans and evaluate investigation tools to ensure timely disease control and response activities. This outbreak provided an opportunity to assess the value of digital pens, which could reduce data-entry burden during public health investigations and provide more timely data for analysis and, ultimately, institution of disease-control measures. Although the use of digital pens was not without problems, there is great potential for the use of these devices for rapid data acquisition in outbreak investigations and other public health events.

About the Authors

New Hampshire Department of Health and Human Services (Dr Matthewson, Ms Daly, Ms Cavallo, and Ms Alic), Concord, New Hampshire

Correspondence and reprint requests to Abigail A. Mathewson, DVM, MPH, Infectious Disease Surveillance Section, New Hampshire Department of Health and Human Services, 29 Hazen Dr, Concord, NH 03301-6504 (e-mail: abigail.mathewson@dhhs.state.nh.us).

Acknowledgments

We wish to thank the following persons who assisted with interviewing survey participants or otherwise helped collect the data used as the basis for this report: Christine Adamski, Kim Budde, Rick Cricenti, Katrina Hansen, Sarah Krycki, Maureen MacDonald, Erin Metcalf, Darlene Morse, and Colleen Smith of the New Hampshire Department of Health and Human Services.

Funding and Disclaimer

This work was supported, in part, by the Centers for Disease Control and Prevention, Epidemiology and Laboratory Capacity for Infectious Diseases (Grant numbers 3U50CI000874-02S2 and 1U50CK000256-01) and

https://doi.org/10.1017/dmp.2015.43 Published online by Cambridge University Press

Hospital Preparedness Program and Public Health Emergency Preparedness (Grant number 5U90TP000535-01) Cooperative Agreements. The findings and conclusions in this report are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or Assistant Secretary for Preparedness and Response. The use of trade names and names of commercial sources is for identification only and does not imply endorsement by these agencies.

Published online: May 5, 2015.

REFERENCES

1. Centers for Disease Control and Prevention (CDC), Office of Public Health Preparedness and Response. Public Health Preparedness Capabilities: National Standards for State and Local Planning. http://www.cdc.gov/phpr/capabilities/index.htm. Published March 15, 2012. Accessed November 10, 2014.
2. Trust for America's Health. Ready or Not? Protecting the Public from Diseases, Disasters, and Bioterrorism. http://healthyamericans.org/report/101/. Published December 2012. Accessed November 10, 2014.
3. Yasmin S, Pogreba-Brown K, Stewart J, et al. Use of an online survey during an outbreak of *Clostridium perfringens* in a retirement community-Arizona, 2012. *J Public Health Manag Pract*. 2014;20:205-209.
4. Greene SK, Daly ER, Talbot EA, et al. Recurrent multistate outbreak of *Salmonella* Newport associated with tomatoes from contaminated fields, 2005. *Epidemiol Infect*. 2008;136:157-165.
5. Ricchetti-Masterson K, Horney J. Social factors as modifiers of Hurricane Irene evacuation behavior in Beaufort County, NC. *PLOS Curr*. 2013; Jun 5:5.
6. Cooper S. More time for patients: Digital pens for clinicians. *Br J Health Care Manag*. 2012;18:249-250.
7. Helm MA, Hauke JA, Schlechtriemen TB, et al. Digital pen and paper—introducing a new technology for prehospital data recording in German Helicopter Emergency Medical Service. *Eur J Emerg Med*. 2011;18:363-364.
8. Vowden KA, Vowden PB. Digital pen and paper technology: a tool for data capture and patient management. *Wounds UK*. 2010;6(2):74-78.
9. Anoto Group AB. Public Safety Improves with Wireless Technology. Healthcare Technology Online. http://www.healthcaretechnologyonline.com/doc/public-safety-improves-with-wireless-0001. Published September 3, 2008. Accessed November 10, 2014.
10. Centers for Disease Control and Prevention. Surveillance for foodborne disease outbreaks—United States, 1998–2008. *MMWR Morb Mortal Wkly Rep*. 2013;62:1-34.

https://doi.org/10.1017/dmp.2015.43 Published online by Cambridge University Press