Epidemiologic Study of Health Effects Associated With Low Pressure Events in Drinking Water Distribution Systems

Request for an Extension of Existing Collection of Information

September 1, 2015

Contact:
Katie Fullerton, MPH
Epidemiologist, Waterborne Disease Prevention Branch
National Center for Emerging and Zoonotic Infectious Diseases
Centers for Disease Control and Prevention
Mailstop C-09
1600 Clifton Rd. NE
Atlanta, GA 30333
Tel: 404-718-4714
Fax: 404-639-2205
E-mail: kgf9@cdc.gov

Table of Contents: Part B

List of Reference Attachments

    A. Authorizing Legislation
    B. 60 day Federal Register Notice
    C. Pilot Evaluation Logical Framework
    D. Pilot Summary and Evaluation Report
    E.1. Utility Press Release
    E.2. Utility Radio PSA
    E.3. Utility Newsletter Paragraph
    E.4. Utility Social Media Posts
    E.5. Utility Promotional Postcard
    E.6. Utility Customer Fact Sheet
    E.7. Utility Water Sample Fact Sheet
    F. Advance Letter
    G. Cover Letter – Paper
    H. Cover Letter – E-mail
    I. Consent Brochure
    J. Household Survey – Paper
    K. Household Survey – Web – Screen Shots
    L. Thank you/Reminder Letter
    M. Replacement Survey Cover Letter – Paper
    N. Replacement Survey Cover Letter – E-mail
    O. Reminder Phone Script
    P. Final Appeal Letter
    Q. IRB Approval Continuation
    R. Low Pressure Event Form
    S. Utility Customer Information

Epidemiologic Study of Health Effects Associated With Low Pressure Events in Drinking Water Distribution Systems

iii. Extension

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Sampling Methods

We plan to conduct a prospective cohort study among households that receive water from five water utilities. For each low pressure event, the water utility will provide the study team with contact information for utility customers from an exposed and an unexposed area. From the utility company lists, CDC will randomly select a number of exposed and unexposed households to be sent the survey. For each low pressure event (LPE), unexposed and exposed households will be recruited in a 2:1 ratio. Sample size calculations are based on a matched (at the LPE level) design, thus calculations depend on the number of LPEs and the number of exposed and unexposed households per LPE. Based on calculations outlined below, we aim to include 65 LPEs in the multi-site study, to obtain information from 4,050 households (1,320 exposed and 2,700 unexposed). Based on an anticipated 60% response rate, we will send surveys to approximately 6,750 households (4,050/0.60) during the time period of the study. We will not be contacting persons younger than 18 years of age. The household respondent will provide information on exposures and illnesses among all household members, and we assume there will be two individuals per household, on average. Thus, we should have data on 8,100 persons from the multi-site study.

The potential universe of respondents for this study would include all households that receive water from the five participating water utilities. The participating water utilities, which are located in the southeast, northeast, northern midwest, southern midwest, and pacific northwest, regions of the U.S., are volunteering to work with CDC on this research study and therefore make up a convenience sample of all utilities in the U.S. The participating utilities were selected on the basis that they had the capacity and management support to participate in this research study, that they typically experience about one low pressure event per month, that they use chlorine or chloramine as a secondary disinfectant, and that they are located in different regions of the U.S. Potential survey respondents will have two options for responding to the survey: paper surveys filled out by study respondents and returned via postal mail to CDC study staff; and web-based surveys filled out by respondents who fill in the study survey responses via a password-protected website housed at CDC. Links to the web survey will be sent to some utility customers by email, when email address information can be provided by the utility.

We anticipate a 60% response rate. Two recent studies related to drinking water conducted in the U.S. have had lower response rates (33% and 48%; see Ailes et al. and Roy in Table B.1.1). In order to increase the anticipated response rate for this study, we incorporated aspects of the Dillman Tailored Design Method. Special attention was given to the presentation and wording of the survey and consent materials. To encourage

participation by fostering a social exchange relationship, we will include a token gift of a refrigerator magnet calendar that highlights the two week period of interest for the survey. To further foster a social exchange relationship, we will include a handwritten post-it note in the second survey mailing. Further, based on results from our pilot study, we estimate that 60% of individuals who respond will use the postal version of the survey, and 40% will respond via the web based version.

Table B.1.1 – Survey Response Rates

First author (country)	Year of study	Survey methods	Overall Response Rate	Proportion Responding via: Web	Mail	Telephone
Ailes (U.S.)	2009	Mail	33%	--	100%	--
Roy (U.S.)	2008	Web, mail	48%	64%	36%	--
Griffith (U.S.)	2008	Web	53%	100%	--	--
Ghosh (U.S.)	2006	Web, telephone	≥78[*] (telephone), ≥74% (web)	--	--	--
Jones (U.S.)	2004	Telephone	33%	--	--	100%
Smith (U.S.)	2001-2003	Web, mail	--[†]	55%	45%	--

[*] There were multiple survey groups per survey method
[†] Longitudinal cohort study follow-up; survey respondents had already agreed to participate in the study but in this round of the study, they were given an option to complete questionnaires via the internet.

We have performed power calculations to assess whether this study will have enough power to provide statistically useful information. In addition to analyzing data from all participants overall, we plan to stratify our analyses by type of water treatment used (i.e., chlorine versus monochloramine as a secondary disinfectant). Therefore, both of these study arms will individually require sufficient statistical power for analysis. For these power calculations, the following assumptions were made:

- The population incidence of AGI among unexposed persons is estimated to be 5% (based on prevalence of AGI during a month-long period in U.S. survey data (Jones, McMillian et al. 2007).
- The event sizes will vary. Of the 65 events for the multi-site study, we aim to include15 small events (affecting at least 10 households), 35 medium events (affecting at least 37 households), and 15 large events (affecting at least 53 households). This distribution of expected event sizes was informed by our pilot utility.
- In an unstratified analysis, with a 60% response rate and Type I error = 0.05, the study will have over 90% power to detect an odds ratio of 1.6, 67% power to detect an odds ratio of 1.4, and 25% power to detect an odds ratio of 1.2.
  In each of the two arms of a stratified analysis (e.g., stratifying on disinfectant type), the study will have over 85% power to detect an odds ratio for AGI of 1.8 or larger.

Under these assumptions and goals, we estimate that we will have enough respondents to meet the goal of detecting a statistically significant odds ratio of 1.6 or greater in an unstratified analysis (an increased risk similar to the Nygard study), and further that we would be able to detect a statistically-significant odds ratio of 1.8 or greater for AGI

within each disinfectant type subgroup. Finding non-significantly elevated risks would suggest that additional research with additional precision and statistical power is needed.

Table B.1.2 – Power Calculations for Overall Study*

Power	AGI Unexposed	AGI Exposed	Detectable Odds Ratio	Exposed Households	Unexposed Households	Unstratified Sample Size
25%	5%	6%	1.2	1,340	2,680	4,020
67%	5%	7%	1.4	1,340	2,680	4,020
93%	5%	8%	1.6	1,340	2,680	4,020
99%	5%	9%	1.8	1,340	2,680	4,020
100%	5%	10%	2.0	1,340	2,680	4,020

*Assumes an event size distribution of 15 small events (at least 6 affected households respond), 35 medium events (at least 22 affected households respond), and 15 large events (at least 32 affected households respond) for the multi-site study; Assumes 2:1 sampling of unexposed to exposed households, and a similar proportion of responding households in the unexposed areas.

If we recruit five utilities for the multi-site study, we would need, on average, approximately one low pressure events per month per utility. However, many low pressure events affect smaller numbers of households, thus the number of households invited to participate will vary based on the number of utility connections affected (Table B.1.3). Using the estimated event size distribution, which was informed by our pilot utility, we estimate an approximate sample size of 4,050 households in the multi-site study (Table B.1.3). Smaller utilities may have less than one event per month; however, we should still be able to achieve a minimum sample size of 4,020 households because we anticipate that larger utilities will report more than one event per month (Table B.1.4).

Table B.1.3 – Distribution of low-pressure event size and number of expected enrolled households.

Event type	Number of events	Exposed households/ event	Unexposed households/ event	Total exposed households	Total unexposed households	Total households
Multi-site study:						
Small	15	6	12	90	180	270
Medium	35	22	44	770	1,540	2,310
Large	15	32	64	480	960	1,440
Multi-site total:	**65**	**60**	**120**	**1,340**	**2,680**	**4,020**

Table B.1.4 – Distribution of low-pressure event size and number of events by utility, assuming one event per utility per month.

Utility	Number of small events	Number of medium events	Number of large events	Total Events	Events / month
A	3	7	3	13	1-2
B	3	7	3	13	1-2
C	3	7	3	13	1-2
D	3	7	3	13	1-2
E	3	7	3	13	1-2
Total:	**15**	**35**	**15**	**65**	

2. Procedures for the Collection of Information

For a detailed discussion of sample selection, sample size, and statistical power for this study see Section B.1.

Identification of exposed and unexposed areas associated with each low pressure event will be undertaken collaboratively between the utility and the CDC study team, including individuals with distribution system engineering expertise. Water utilities will provide information on up to two LPEs per month. Exposed households will be identified as addresses that experience a loss or lack of water pressure due to an LPE in the water distribution system. The utility will identify the households affected by the LPE with the assistance of their water system information systems, and convey this information to CDC, using street names and a map if available. We will not ask residents to participate in the study more than once, therefore events that occur in areas that were previously sampled as either exposed or unexposed areas will not be eligible.

Using the address of the event and map of affected households, the CDC study team will determine the census block(s) affected, as well as the census block group and census tract. A map showing these areas will be provided to the utility to narrow down the area in which to seek a group of homes that are comparable to the homes affected by the LPE. Because census tracts are designed to be homogeneous with respect to population characteristics, identifying unaffected areas within the same block group (or census tract for large events) will help to minimize bias by assuring that affected and unaffected areas are reasonably well matched in terms of demographic characteristics. Using its own water system information systems, in collaboration with the CDC study team, the utility will then identify similarly-sized or larger areas within the same census block group (or census tract if necessary) that are unaffected by the LPE because they are in a different pressure zone, upstream of the event, involve a separate loop, or are in an area of the distribution system separate from the area of the LPE. From the potential areas identified, utilities will identify an area that matches the affected (event) area in terms of predominant housing type, pipe size and material, and origin of water, as detailed on the Low Pressure Event form. If multiple potential matching areas are found, CDC will randomly select the area for inclusion.

The utility will provide the billing addresses of the exposed and unexposed households or the mailing addresses from public assessor information to CDC research staff within one week of an LPE. From this address list, a random sample of exposed and unexposed households will be generated by CDC study staff (see section B.1 for details on

sampling). Households will only be surveyed once (either as a household exposed to an LPE or as an unexposed household).

Household respondents must be English-speaking and ≥18 years of age in order to be eligible to participate in the study. Children will not be targeted to be enrolled in the study per se, as we will not survey anyone <18 years of age. However, the parent or guardian of a child may respond to the questionnaire regarding recent illness and exposures of children in his/her household. Households that have been invited to participate in response to a previous event will not be included in the sampling frame for subsequent events.

Water connection addresses that are linked to businesses will be excluded. Residential facilities such as nursing homes or long-term care facilities will also be excluded because populations in these facilities generally have a higher rate of underlying medical conditions which predisposes them to a higher baseline incidence of diarrheal illness. They will be excluded to avoid bias. To exclude potentially ineligible residential properties (e.g., rental properties or seasonal homes), we will exclude addresses for which the mailing and premise addresses do not match.

All potential study participants will have advance notice of the survey. An advance letter with information about the study (Attachment F) will be mailed to each selected household approximately 1-1.5 weeks after the low pressure event. The letter will be personalized with the utility customer's name and home address using CDC stationary, and will be signed by the principal investigator. This letter will introduce the importance of the study and inform residents that they will receive a survey packet in a few days, with options to complete the paper survey and mail it in or to complete it online using a unique code. The advance letter will also inform study participants that they will receive a small gift (the refrigerator magnet) in their study packet as a token of our appreciation for their participation.

A survey packet, including an introductory letter, study brochure with consent information (Attachments G, H, and I) and the household survey (Attachments J and K), will be mailed from CDC to each randomly selected exposed and unexposed household approximately two weeks after the LPE (i.e., 0.5-1 week after the advance letter). Mailed study materials will also direct participants to a study website (http://www.cdc.gov/healthywater/study.html) where they can seek more information about the study or link to the secure study website to take the survey online. In addition, a refrigerator magnet depicting a calendar will be included with the designated two week period of interest highlighted as a reminder for the household to only answer about those days indicated. Two to three weeks is generally the maximum incubation period for pathogens known to cause waterborne AGI and ARI. Therefore, those persons who were exposed to these pathogens during the LPE will likely have developed symptoms by the time they receive the questionnaire. The letter will include information about the study and invite the recipient to participate. In order to maximize response rates, we will give survey participants the choice of two response modes: over the internet using a web-based survey tool, or via postal mail. The introductory letter will provide a study

identification number that the respondent will use as a pass code to enter on the survey website, should the respondent prefer to complete the survey online. This information will be duplicated in a label on the front of the survey booklet. We will include a postage-paid return envelope (addressed to CDC) for return of paper surveys. The calendar magnet included in the survey packet will have the dates of the 2-week period of interest highlighted. Some residents will be emailed a link to the web survey, when this information is available from the utility (Attachment H). By utilizing multiple survey methods, we hope to have equitable recruitment of subjects.

We will attempt to obtain survey information on all members of each household (limited to 6 in the paper version of the survey due to printing constraints). Study participants will receive a refrigerator magnet calendar from the CDC as a visual aid to improve recall and as a token gift designed to improve response rates.

The survey participants will indicate their consent to participate by completing the survey (in the postal and online versions of the survey). Instructions will clearly state that only individuals ≥ 18 years of age are eligible to fill out the questionnaire. No names will be collected at any time during the study.

Both versions of the study questionnaire will ask the same standardized questions from each household member concerning his or her recent gastrointestinal or respiratory illnesses, water service and exposures, and other activities. The questionnaire will focus on the two week time period after the LPE (referred to in the survey as the "two week period"). For households where children <18 years of age are present, we will ask a parent or guardian to answer questions and provide information on behalf of the child. Data from the paper-based surveys will be manually entered into the electronic database at CDC that will also house the data from the web-based surveys. Data from the web-based surveys will be saved automatically in the database after the participant answers a question.

One week after mailing the survey packet, we will send a personalized thank you/reminder notecard to the selected exposed and control households (Attachment L). The notecard will thank participants who already responded and request that those who have not yet responded do so. Participants will be reminded that they are important to the study, that their responses are useful, and that their participation is voluntary. The participant's individual login and passcode for the web survey will be included on the card, and participants will also be given a phone number to call with questions or to request a replacement questionnaire.

One week after mailing the thank you/reminder card, we will mail a replacement questionnaire packet to households that have not yet responded. The packet will contain essentially the same information as the first questionnaire mailing, but the letter will be altered slightly to indicate that we have not yet received their response, remind them that participation is easy and quick, and provides important information for public health planning (Attachment M), and a hand-written Post-It note encouraging them to respond.

For customers with available email address information, a reminder email with a survey link will also be sent (Attachment N).

One week after mailing the replacement questionnaire, the study team will place an outbound telephone call as a reminder before the final appeal letter is sent in an attempt to further increase the response rate (Attachment O).

One week after placing the reminder telephone call, we will send a final mailing to non-responders (Attachment P). This final letter will be in a different color envelope, contain a personalized letter on CDC letterhead with the signature of the principal investigator, and have a reminder of the participant's individual passcode for the web survey. This letter will convey a sense of urgency that time is running out for participation, and inform participants that all responses must be received within two weeks. Responses received by mail or internet more than 8 weeks from the low pressure events will still be accepted, and a sensitivity analysis will be used to determine whether these late responses introduce recall bias.

In addition to the information collected from the study participants, data will be collected from the water utility about the LPE using the LPE form (Attachment R). The utility repair crew will complete one LPE form per event, and a utility manager will review and approve the form before sending the form to CDC. Prior to the final selection of exposed and unexposed areas, the CDC study team will review the information and request follow-up from utility staff for missing or unclear information. The water utility will provide information on the condition of the distribution system infrastructure (e.g., age and type of pipe material), the duration of the LPE, the extent of the areas of the population affected (e.g., the number of households), location of wastewater and storm water collection systems in relation to the impacted water line, and procedures used to restore the distribution system after the LPE. Hydraulic model outputs and monitoring data including total coliforms, disinfectant residuals, turbidity and other information that relates to the microbiological status of the distribution system may also be collected.

We will also conduct supplemental environmental sampling. For each LPE, utility company personnel dispatched to perform repairs or maintenance will collect water samples from the distribution system in both the exposed area and the unexposed area from hydrants, sampling points, or commercial and residential hose bibs within 24 hours of the LPE. Water sample sites (selected by the water utilities) will not be linked to household survey addresses; rather, they will be assumed to represent water in the distribution system in the affected and unaffected areas. The following is a potential sampling schedule:
- Three, ~100-L drinking water samples from exposure area within 24 hours of LPE (exterior hose bibs)
- Three, ~100-L drinking water samples from unexposed area within 48 hours of LPE (exterior hose bibs)

Water samples will be collected using a published ultrafiltration technique (Smith and Hill 2009). Environmental sampling collection supplies and pre-paid shipping labels will

be provided to participating utilities. Depending upon their laboratory capacity, some of the environmental analyses will be conducted at the utility's own laboratory, but most samples will be shipped in coolers to the CDC Environmental Microbiology Laboratory for testing within 36 hours of collection. The CDC laboratory regularly responds to waterborne disease outbreaks across the U.S. and has extensive experience training partners. The team has monitoring systems, protocols, and training materials in place to ensure that the CDC team can monitor the water sample collection and mailing. Compliance with chain of custody procedures will be monitored and enforced. A utility manager will notify the CDC when an event occurs, and the CDC team will follow up in the event that there is a delay in receiving samples.

The drinking water samples will be analyzed for a suite of microbial indicators of fecal contamination including, total coliforms/*E. coli*, somatic coliphages, *Bacterodales spp.*, total aerobic spores, heterotrophic plate count (HPC) bacteria, Adenosine Triphosphate (ATP), and *Pseudomonas aeruginosa*.

Information on weather and climate variables (especially rainfall immediately preceding, during, and following the low pressure episode) will be collected and analyzed in context of the study by CDC staff using the National Oceanic and Atmospheric Administration's (NOAA) national climate database (http://www.climate.gov/).

Several quality control procedures will be implemented as part of this study:
- The web-based questionnaire has built-in skip patterns and internal logic controls for efficiently routing the respondent to the relevant questions based on their prior responses. The paper-based questionnaire has the same questions and the same skip patterns as the web-based version. However, participants completing the paper-based questionnaire must manually follow these skip patterns, which may increase the risk for data entry errors. Additionally, the web-based questionnaire will employ a variety of prompts to encourage survey completion, whereas the paper-based questionnaire will have no such prompts.
- The web-based questionnaire has data entry validation to limit data entry errors and reduce data cleaning efforts. Furthermore, data entry into the database is automatic thereby eliminating the need for manual data entry, which also limits potential data entry errors.
- The CDC study team will manually clean the database at the end of the data collection period.
- The Low Pressure Event form will be reviewed by the CDC study team following each event before participants are selected, and the utility staff will be contacted for completion of missing information.
- Utilities will be trained in following the CDC lab team's standard chain of custody procedures, and the CDC team will follow up with the utility if samples are not received in the allotted time following LPEs.

3. Methods to Maximize Response Rates and Deal with Noresponse
Participation in our study will be completely voluntary and no financial incentives will be offered. Still, we have attempted to maximize response rates through the following:

- Utilizing both mailed and web-based survey methods and giving potential participants an option of how they would prefer to respond to the survey;
- Keeping the survey length to a minimum, and using a respondent-friendly survey booklet. We anticipate that it will take approximately 12 minutes for each participant to complete either version (web or paper) of the questionnaire.
- Incorporating additional aspects of the Dillman Tailored Design Method, which have been shown to improve response rates (Dillman 2007): sending an advance letter; sending a personalized cover letter with the mailed questionnaire; including a token gift (refrigerator magnet calendar); sending a thank you/reminder postcard one week after mailing the survey packet; sending a follow-up survey mailing to households who have not responded to the survey after two weeks and including a personalized post-it note; and sending a final mailing in a contrasting envelope.

To avoid making persons feel pressured or coerced into consenting to participate in the study, all communications, including the advance letter, cover letter, follow-up mailings, and the information provided at the beginning of the web-based questionnaire, will advise readers that participation in the study is voluntary.

In the pilot study, the overall survey response rate was 37%, the response rates for each event ranged from 32-43% and did not appear to vary by event type or size. Additionally, the response rates were similar in LPE (38%) and non-LPE areas (36%). The non-response rate was < 10% for the main survey items. There was higher non-response for follow-up questions for the illness details onset date and number of days with symptoms. Other illness detail questions such as the number of days of school or work missed because of the illness, or healthcare seeking behavior were nearly complete. Skip patterns were built into the web survey interface; for the paper surveys, data from the mail surveys indicated that the majority of respondents correctly skipped or answered the appropriate questions. System data demonstrated that the web-based survey took customers a median of 11 minutes to complete, suggesting that the time burden for participation was low and in line with the anticipated time of 12 minutes. Most (94%) of the web surveys were submitted successfully, demonstrating there was limited survey break-off.

To improve response rates in the full study, the study team has modified the survey procedures and will increase efforts to promote the study in the participating communities. Nearly half of the page views for the study website (46%) occurred in the 30 days following the press release, before any survey materials were mailed out, suggesting that additional ongoing publicity in the study communities has potential to motivate participants. The study team will implement additional community outreach throughout the study period to improve community acceptance of the study and to boost response rates. Possible methods include implementing local public service announcements or periodic advertisements in local print media. Improving the study response rates is a priority for CDC branch-level management, who has prioritized obtaining resources for targeted communications efforts, including contract time with a health communicator.

The multiple mailing contact strategy encouraged participation, evidenced by boosts in survey response following the mail prompts. The majority of respondents chose to return the survey by postal mail (70%), using the provided return envelope. The data quality of

the web surveys was higher than the paper surveys because data verification rules and question skip patterns were built into the survey interface. Since the web survey instructions and access information were printed on the survey materials, rather than sent electronically to customers, it might have been inconvenient for respondents to access a computer, type the link to the website, and log-in to take the survey. To encourage web survey participation and reduce respondent burden in hopes of increasing response, the study team will send the survey link electronically to customers that have email addresses on file (approximately 10% of pilot utility customers). Additionally, the study team will add an outbound phone call as a reminder before the final appeal letter is sent.

Several recent studies have indicated that a response rate of 80% is not feasible for this study. A recent postal survey of households in a community affected by drinking water-associated outbreak of salmonellosis that also asked about acute gastrointestinal illness had a response rate of 33% (Ailes, Budge et al. 2010). A survey of backcountry hikers that allowed respondents to choose to respond via web-based or paper-based survey had a higher response rate of 48% (S. Roy, personal communication). Response rates obtained by some large continuing health surveys that provide key national health statistics (such as the FoodNet population studies) have response rates of 33% in a past survey cycle (Jones, McMillian et al. 2007). In spite of these low response rates, such surveys still provide valuable data to inform public health recommendations, policies, and programs, and we believe that our study will have similar value. Compared to these prior studies, we are implementing several procedures to attempt to encourage study participation, as outlined above, so a participation rate of 60% will be possible. Notably, the household survey for this study is substantially shorter than the other surveys mentioned in this section, and includes a token gift.

In the contingency that we do not reach 60% participation, we will still have 80% statistical power to identify an odds ratio of 1.6 in an unstratified analysis if we have at least 40% participation. With 40% participation, we would have 51% power to detect an odds ratio of 1.4 and 18% power to detect an odds ratio of 1.2 in an unstratified analysis. In a stratified analysis, with 40% participation, we would have over 85% power to detect an odds ratio of 2.0 and over 70% power to detect an odds ratio of 1.8.

Our primary concern regarding obtaining low response rates is the potential for bias engendered by differential response based on exposure and outcomes. If the response is low, we are also concerned that, overall, the study population may not be representative of the target population (i.e., non-response bias). However, recent research indicates that the magnitude of non-response bias is generally small (Galea and Tracy 2007).

Following the pilot, we assessed item non-response, and minor changes were made to the survey instrument to improve the quality of the data and facilitate ease of participation, which were approved in a nonsubstantive change request. The pilot survey helped the study team identify items that could be eliminated, information gaps that could be filled, and response options that should be added to existing questions. The survey recall time period was shortened from 3 weeks to 2 weeks, and the illness questions and water use questions were simplified. The survey supplemental materials were modified slightly to make them easier and faster to read. For example, the advance letter was edited to

improve clarity and readability, and the web survey has fewer questions per screen to make the survey easier to read. The study team does not have the resources to provide financial incentives. Although 94% of the web survey respondents successfully completed the survey, indicating survey break-off was not a major concern, additional measures were taken to encourage respondents to finish the web survey, including reformatting the survey and simplifying instructions to make the survey appear shorter and to reduce reading time.

To assess differential response following the multi-site study, we will compare responders and non-responders with respect to their *a priori* exposure status. If the response rates in exposed areas differ from response rates in unexposed areas by >20% (contrary to findings from the pilot study), we will plan to implement sensitivity analyses to estimate how greatly these might have affected results regarding the primary research question. We will compare demographic characteristics of respondents to census data at the Public Use Microsample Area (PUMA) level. While this level of data is not granular enough to determine whether data from each LPE cluster is representative of its respective census block group, it should be able to identify gross deviations from expected distributions of demographics across all events.

Next, we will evaluate item non-response and incomplete surveys (i.e. break-offs). Items that are missing for more than 10% of respondents will be evaluated for wording, structure, and placement in the survey. Incomplete surveys (break-offs) will be evaluated to determine where the break-off occurs, and whether break-offs differ on the paper and web versions of the survey. We will evaluate whether missing data is differential by exposure status or demographic characteristics, including household size and composition. We anticipate that outcome data may sometimes be missing for adults who share a household but may not be relatives (i.e. roommates), but expect that outcome data for related household members (such as the respondent's children) should be available and will be valuable to the study.

4. Tests of Procedures or Methods to be Undertaken
The data collection instruments and the participant information materials were piloted at one water utility site and evaluated by CDC. Overall, the pilot demonstrated that the study design and procedures will allow the study team to collect the data needed to meet the study goal and aims.

Prior to the pilot, the data collection instruments and the participant information materials (e.g., letters, websites) were reviewed (1) for content by three senior epidemiologists at CDC; (2) for question design, flow, clarity, and timing by five CDC staff members; and (3) for content by experts in the field. The questionnaire and participant information materials were also pilot-tested by six CDC staff members, and three members of the general public.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data
Statistical consultants for study design:

- Anna Blackstock, CDC – ph. 404-718-1228, email: hyp9@cdc.gov
- Tracy Ayers, CDC – ph. 404-639-0543, email: eyk6@cdc.gov
- Gordana Derado, CDC – ph. 404-639-0393, email: uwx8@cdc.gov
- Xin (Lucy) Lu, CDC/ORISE (Tracy Ayers is supervisor) – ph. 404-718-4868, email: xjp2@cdc.gov

Persons who designed the data collection:
- Joan Brunkard, CDC – ph. 770-488-7711, email: feu4@cdc.gov
- Julia Gargano, CDC – ph. 404-718-4893, e-mail: igc5@cdc.gov
- Vincent Hill, CDC – ph. 404-718-4151, email: veh2@cdc.gov
- Joseph Carpenter, CDC – ph. 404-639-3900, email: jcc0@cdc.gov
- Elizabeth Ailes, formerly contractor for CDC[*] – ph 770-488-7563, email: eha0@cdc.gov
- Mark Lamias, contractor for CDC[**] – ph. 404-498-1151, email: bnz6@cdc.gov
- Bonnie Mull, CDC - ph. 404-718-4156, email: gtx1@cdc.gov
- Elizabeth Adam, CDC/ORISE (Julia Gargano is supervisor) – ph. 404-718-4873, email: wsi7@cdc.gov

Persons who will collect the data:
- Katie Fullerton, CDC – ph. 404-718-; email: kgf9@cdc.gov
- Julia Gargano, CDC – ph. 404-718-4893, email: igc5@cdc.gov
- Elizabeth Adam, contractor for CDC (Katie Fullerton is supervisor)– ph. 404-718-4873, email: wsi7@cdc.gov
- Mark Lamias, contractor for CDC[**] – ph. 404-498-1151, email: bnz6@cdc.gov
- Vincent Hill, CDC – ph. 404-718-4151, email: veh2@cdc.gov
- Chandra Schneeberger, contractor for CDC (Vincent Hill is supervisor) - ph. 404-718-4154 email: ipq7@cdc.gov
- Alison Stargel, contractor for CDC (Vincent Hill is supervisor) - ph. 404-718-4154 email: kna8@cdc.gov

Persons who will analyze the data:
- Julia Gargano, CDC – ph. 404-718-4893, email: igc5@cdc.gov
- Elizabeth Adam, contractor for CDC (Julia Gargano is supervisor)– ph. 404-718-4873, email: wsi7@cdc.gov

[*] Elizabeth Ailes' CDC supervisor was Joan Brunkard, ph. 770-488-7711, email: feu4@cdc.gov
[**] Mark Lamias' CDC supervisor is Robert Fish, ph. 404.498.2646, email: bnz6@cdc.gov

References

Ailes, E., P. Budge, et al. (2010). Economic and Health Impacts Associated with a Salmonella Serotype Typhimurium Drinking Water Outbreak -- Alamosa, CO, 2008. International Conference on Emerging Infectious Diseases, Atlanta, GA.

Dillman, D. A. (2007). Mail and Internet Surveys: The Tailored Design Method. New Jersey, John Wiley & Sons, Inc.

Galea, S. and M. Tracy (2007). "Participation Rates in Epidemiologic Studies." Annals of Epidemiology **17**(9): 643-653.

Ghosh, T. S., J. L. Patnaik, et al. (2008). "Internet- versus telephone-based local outbreak investigations." Emerg Infect Dis **14**(6): 975-977.

Griffiths, S. L., R. L. Salmon, et al. (2010). "Using the internet for rapid investigation of an outbreak of diarrhoeal illness in mountain bikers." Epidemiol Infect **138**(12): 1704-1711.

Jones, T. F., M. B. McMillian, et al. (2007). "A population-based estimate of the substantial burden of diarrhoeal disease in the United States; FoodNet, 1996-2003." Epidemiol Infect **135**(2): 293-301.

Smith, B., T. C. Smith, et al. (2007). "When epidemiology meets the Internet: Web-based surveys in the Millennium Cohort Study." Am J Epidemiol **166**(11): 1345-1354.

Smith, C. M. and V. R. Hill (2009). "Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes from water." Appl Environ Microbiol **75**(16): 5284-5289.