

B. Collection of Information Employing Statistical Methods

Central America Water and Sanitation Program Sustainability Evaluation and Qualitative Interview

August 30, 2011

Raquel Sabogal, MSPH
Epidemiologist
770 488 3432
zkq3@cdc.gov
Fax: 770 488 3450

Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Environmental Health Services Branch
Global Water, Sanitation and Hygiene Education (GWASH) Team
4770 Buford Highway, NE
Atlanta, GA 30341-3724

Table of Contents

1. Respondent Universe and Sampling Methods.....	2
2. Procedures for the Collection of Information.....	3
3. Methods to Maximize Response Rates and Deal with Nonresponse.....	5
4. Test of Procedures or Methods to be Undertaken.....	5
5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data.....	6
References.....	8

1. Respondent Universe and Sampling Methods

CDC will survey communities that are part of the American Red Cross (ARC) post-Hurricane Mitch project so that their water, sanitation and hygiene education (WASH) program can be evaluated for sustainability. In addition, households will be surveyed to quantify four indicators related to WASH and to gather their thoughts and opinions from select households that have been a part of the program for at least 10 years.

Community Survey. There will be 15 to 16 communities in the study and 15 to 16 community surveys will be conducted. Expected response rate is 100% based on previous studies done by ARC in these communities. Household surveys will be carried out after the community survey is completed. Program evaluation is targeted at the community level.

Statistical Justification for Sample Size

The statistical justification for selecting 15 to 16 communities is based on work by Bell et.al. (2010). The rule of thumb for multi-level studies where individuals (households) are nested within groups (communities) recommends a minimum of 15 to 30 units. The criteria for a community to be included in this study are for it to have received the entire ARC program-water system, sanitation facilities and hygiene education. Currently there are eight communities that have been periodically surveyed. An additional seven to eight communities will be included. The additional eligible communities will be identified by the local Red Cross (RC) societies selected from an existing list of post-Hurricane Mitch communities that were part of the ARC program. In total, 15 to 16 communities are sufficient for program evaluation to determine sustainability.

Quantitative Survey. After the community survey is completed, so that the number of households in the community is confirmed, a systematic but random system will be used to select households. Households that access water from the ARC water distribution system will be eligible to be interviewed. Every xth house in each community (the total number of households in the community divided by 15 or 16) will be included until the appropriate number of households is reached. If no one is present at a selected home, then the next house to the right or left will be approached. Expected response rate is 99 to 100% based on ARC's past experience doing studies in these communities.

Statistical Justification of Sample Size

The statistical justification for selecting 15 to 16 households per community is also based on the work by Bell et.al. (2010). This is a multi-level study where the number of individuals (households) is nested within groups (communities). The recommended minimum is 15 to 30 units, so we will include 15 to 16 households per community. The criteria for a household to be included in this study are that it had received the ARC WASH interventions. In total, 240 to 256 households will be surveyed.

Qualitative Interviews. Qualitative interviews require a smaller number of households to understand common perceptions and experiences among a population. Currently there are eight communities in this study. These communities are:

Country	Communities	
El Salvador	La Ceiba	La Pozas
Guatemala	Plan Shalagua	Guayabo
Honduras	Las Lomas	Marcovia
Nicaragua	Dipilto Nuevo	Dipilto Viejo

In addition, seven or eight communities that have never been surveyed before by CDC will be selected. Two randomly selected households will be approached in the 15 to 16 communities in four countries. This will be determined along with the household survey so that every xth house, depending on community size, will be selected for this type of interview. If a household is approached for the quantitative survey it is not eligible to be interviewed again. There will be 30 to 32 key informant interviews completed. Expected response rate is 99 to 100% based on ARC’s past experience doing studies in these communities.

Water samples will also be collected by the CDC interviewer from households and community water sources/water systems.

Statistical Justification of Sample Size

The statistical justification of this sample size is based on the paper by Guest et.al (2006). Results presented in this paper indicated that six to 12 interviews may be sufficient to achieve the desired research objective if the goal is to identify overarching themes based on opinions and perceptions (Guest et. al., 2006). As in the work by Guest et.al, the populations in these communities are similar in demographics and relatively homogenous. All communities received the same type of ARC programming, only varying by date of completion and how well each intervention has been maintained through the years. Results will be presented on a regional basis.

All surveys/interviews will be conducted in Spanish.

Type of Survey	Number	Response rate
Community survey	15-16 communities	100%
Quantitative survey	256 households (maximum) face-to-face interviews with female head of household	98-100%
Qualitative interviews	32 (maximum) face-to face interviews with female head of household	98-100%

2. Procedures for the Collection of Information

This study is a cross-sectional design at the household level and program evaluation at the community level. Data will be collected from households that are in the communities that are

part of the ARC post-Hurricane Mitch program from 1998. Many of these communities have had population growth through the years and only homes that are currently or have been connected to the ARC constructed water distribution system will be included in the survey. CDC personnel from the Global Water Sanitation and Hygiene (GWASH) team from the National Center for Environmental Health will collect data. The GWASH team is experienced in field data collection and knowledgeable in the areas of epidemiology and environmental health, which includes water, sanitation, hygiene education programming, water quality testing, and data analysis. Team members will also pre-test all surveys that will be administered in the field prior to actual data collection. Pre-testing of surveys will be done in Spanish by team members with persons not familiar with the survey (**Attachment 4**). Use of personal data assistants (PDAs) will be practiced by administering the survey prior to actual data collection.

Through our Red Cross (RC) collaborator and local RC societies, communities will be informed in advance that the team from CDC will be coming to the community to conduct this survey. Once in country, CDC will meet each of the local RC societies. The team will go to each community accompanied by our RC partners and meet with the community leaders and/or water committee members. The community survey will be conducted in a small group setting with one CDC person conducting the interview and taking notes. Other GWASH team members will also assist by taking handwritten notes during the meeting. The community survey is first completed so that a clear understanding of the status of each community is known prior to data collection at the household level. This interview should last about 45 minutes to 1 hour.

The size of each community will be taken into consideration when defining the systematic random sample of households. Households will be selected at the same time for both the quantitative survey and qualitative interviews and eligible to be interviewed only once. The surveys will cover the same topic areas such as water collection, water use, water storage, latrine use and maintenance, and if and when hygiene education was received.

Household survey data will be collected using a handheld PDA. The quantitative survey will primarily request dichotomous and categorical answers with only a few open ended questions. The qualitative key informant interview will request the opinions, thoughts, and perceptions of the interviewee. Tape recording the verbal response is the selected method for capturing all data from the interview and will be transcribed upon the team's return to Atlanta.

Each survey will have an observation portion. The CDC interviewer will observe how water is stored and handled in the home, evaluate the condition of the latrine - that it is being used by the family, and observe how the interviewee performs hand washing. The quantitative survey should take approximately 30 to 45 minutes and the face-to-face interview will take from 45 minutes to 1 hour.

Quality Control Procedures

Quality control will be done on a daily basis while in the field. All data collection will be conducted in Spanish. All team members are bilingual and have the necessary language skills so that all work on this project can be conducted independently without a need for interpreters.

The community survey will be done using a paper survey. There will be an additional note taker along with the person conducting the survey with the community. Information from the community survey will be compiled by the community survey interviewer and team member assigned to take notes in that community.

All data collectors are CDC personnel that will have pre-tested the quantitative and qualitative data collection instruments prior to going into the field. The quantitative data collection will be done using a PDA. This method reduces input error and allows for quicker data analysis. Team members will review other team member's data collection of the household survey on a daily basis. The qualitative interviews will be tape recorded and transcribed upon return to Atlanta. The qualitative interviewers will assist in transcribing the interview in Spanish and then to English for coding the text to identify major themes.

3. Methods to Maximize Response Rates and Deal with Nonresponse

Based on the ARC's past experience in these communities, response rates for this study are very high. Our approach is to use available community members to accompany each CDC interviewer to the selected household. Once the female head of household is encountered, the local community member introduces the CDC interviewer to the head of household. After the initial introduction, the recruiting scripts will be given to the head of household. Each head of household has the opportunity to decline participation in the study; however, since there is prior knowledge of CDC's arrival to the community and introduction of the CDC interviewer by a local community member, the response rate is almost always 100% participation. If contact is made and a person declines to participate, the person is thanked and the next home will be approached until there is a household that agrees to participate.

4. Test of Procedures or Methods to be Undertaken

The results of the community survey will be summarized by study area and compared to the community survey results from 2006 and 2009 studies done by ARC with CDC's technical assistance. Summaries are written paragraphs of each study areas and may include an Excel spreadsheet of key sustainability elements for comparison across the years.

Data from household interviews will be compiled from the PDAs and transferred into Epi Info 3.5.1 (CDC 2008) at the end of each sampling day. Data from individual study areas will be pooled, and descriptive statistics obtained from the data. Key demographic data and other frequency data of interest (e.g., primary water source, hand washing technique, sanitation practices) will be compared to historical data from 2006 and 2009 and compared on a regional basis. The following tables will be generated:

- Household Availability and Storage
- Household Sanitation Facility Evaluation
- Hand Washing Behaviors and Education
- Household Hygienic Sanitation Facility Evaluation
- Household Treatment of Drinking Water
- Qualitative Results for Household Water Samples
- Qualitative Results for Tap Water Samples
- Qualitative Results for Community Water Sources

- Free Chlorine Residuals in Water Systems Using Chlorine Treatment
- 2009 Community Questionnaire Results
- USAID FANTA Guide Indicators, Regional Results

Additional tables will be generated based on the results observed from this data collection. Further analyses performed on the household survey results will be done by the CDC investigators upon their return to Atlanta by using SAS Software version 9.2 (SAS Institute Inc, 2002–2008, Cary, NC).

The qualitative face-to-face interviews will be coded to identify major themes. The thematic areas that are expected are on water - use, availability, access, safe handling and/storage; sanitation - use and availability; and hygiene education - proper hand washing technique and hygienic latrine maintenance and use.

The data collection instruments used in this study has been used in previous studies completed in 2006 and 2009. The current survey instruments were based on the three-year health study completed in 2002 and revised to focus on the data needed to estimate four indicators that relate to WASH and questions that focus on the same themes. All 2012 data collection materials were reviewed again by CDC personnel. These revised documents were again pretested. The pretests helped to improve the data collection materials, include additional questions if needed and ensure proper coding for data analysis.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The Division’s statistician was used as the technical advisor in developing the study design. The GWASH team has participated in designing the data collection and revised the survey instruments and will also be responsible for data collection. The team members marked with an asterisk (*) are the primary persons responsible for data analysis.

Raquel I. Sabogal, MSPH*

Epidemiologist
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Environmental Health Services Branch
Global Water, Sanitation and Hygiene Education (GWASH) Team
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-3432

Richard Gelting, PhD

Environmental Engineer
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Environmental Health Services Branch
Global Water, Sanitation and Hygiene Education (GWASH) Team

4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-7067

Elizabeth Medlin, MPH*

Epidemiologist
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Environmental Health Services Branch
Global Water, Sanitation and Hygiene Education (GWASH) Team
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-0530

Victoria Cuellar, MPH

Epidemiologist
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Environmental Health Services Branch
Global Water, Sanitation and Hygiene Education (GWASH) Team
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-1476

Kristin Delea, MPH

Epidemiologist
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Environmental Health Services Branch
Global Water, Sanitation and Hygiene Education (GWASH) Team
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-7065

Sue Sloop

Statistician
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
404-639-7054

Shailen N. Banerjee, Ph.D

Statistician
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-1492

Abigail Boorstin

Student researcher
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Emergency and Environmental Health Services
Global Water, Sanitation and Hygiene Education (GWASH) Team
4770 Buford Hwy, NE MS F-60
Atlanta, GA 30341
770-488-0676

References

B Bell, GM Morgan, JA Schoenberger, BL Loudermilk (2010). Dancing the Sample Size Limbo with Mixed Models: How Low Can You Go? Paper 197.

G Guest, A Bunce, L Johnson. (2006). How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*. Vol 18 (1). Feb: 59-82.