Knowledge, Attitudes, and Practices related to a Domestic Readiness Initiative on Zika Virus Disease

Request for OMB approval of ICR

Submitted on: February 15, 2017

Supporting Statement B

Program Official/Project Officer
Fred Fridinger, DrPH
Health Communications Specialist
Office of Communication Science
Office of the Associate Director for Communication
2500 Century Blvd., MS-E21
Atlanta, GA 30345
Phone: (404) 639-0632
Email:fwf0@cdc.gov

TABLE OF CONTENTS

SECTION B – COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

This is an Information Collection Request (ICR) for approval to conduct a final survey (wave 3) evaluating the CDC Domestic Readiness Initiative for Zika Virus. ICR emergency approval was granted in October 2016 to conduct six months of data collection which included an initial survey at the start of the campaign (wave 1) and a mid-point assessment survey (wave 2). Following this same design, the third wave of data collection will include 600 respondents at each of four locations with differing levels of campaign activity, for a total of 2400 interviews.

1. Respondent Universe and Sampling Methods

The respondent universe will be participants within select areas of the campaign catchment area in the domestic U.S. and Puerto Rico, based on evolving epidemiological data and historical presence of primarily *Aedes aegypti* but also *Aedes albopictus* mosquitoes as well as updated campaign implementation plans. Based on Zika activity reported in the United States in 2016 and subsequent campaign activity, the data collection will occur in the following four locations:

1. an area that has had an <u>intense, long-running active campaign</u>, of which the Island of **Puerto Rico** is the only such location
2. a <u>current intense active campaign</u> that has been taking place in the **Miami DMA** (designated market area)
3. a <u>newer campaign area</u> with mass media plus digital of which the **Houston DMA** is typified as representing, and,
4. a <u>control area</u> that currently has had no significant campaign efforts implemented and of which the state of **Mississippi** would represent.

The sample designs for each of the four groups are for 600 completed interviews. Samples are stratified into two strata: adults ages 18-44 and adults ages 45 and older. The purpose of the stratification is to allow for a purposive oversample of the younger stratum so that 450 cases or 75% of the completed interviews (600 x 0.75 = 450) are below age 45, leaving the balance of 150 interviews with older adults. Since younger adults naturally make up between 47% and 46% of the adult population, (2014 American Community Survey estimates), a statistical design effect (Deff) of 1.3 will result when this stratified design is correctly weighted to compensate for the oversample (note: when there is no oversample and only a simple random sample is used, Deff = 1.0). Thus, the effective sample sizes (n/Deff) will have an approximate precision of ±4.6 percentage points. With the same sample size per area and stratification, a comparison of proportional estimates between waves will be able to detect differences in excess of 9.5 percentage points with 80% power at an alpha level of 0.05 for a proportion of 50%.

For estimates about the young adult oversample stratum of 450 cases in each area's sample, there will be a proportional sample precision of ±5.4 percentage points. Sub-analyses comparing this stratum between areas and per wave will be able to detect differences of 11.0 percentage points with 80% power at an alpha level of 0.05 for a proportion of 50%. For the smaller, older adult stratum of 150 cases, the precision of proportional estimates is ±9.2 percentage points. Sub-analyses comparing this stratum between areas and per wave will be able to detect differences of 19.0 percentage points with 80% power at an alpha level of 0.05 for a proportion of 50%.

The sample has been designed to optimize available resources for making the most precise estimates for younger adults (ages 18-44) and overall for the total area-defined domestic sample and the island of Puerto Rico. There are limitations, however, when this particular optimized design is used. In the U.S. sample, the number of completed cases are designed to generalize findings for each of the specific DMAs (Miami and Houston) and for the entire state of Mississippi. The Mississippi sample will distribute itself in the same way as the Mississippi population is distributed across the state with more sample in the more densely populated areas. This state sample is not optimally designed to make any sub-area analyses. In Puerto Rico, it will also not be optimal to make any sub-geographic area comparisons, i.e. comparing smaller administrative units, because the sample distributes itself in a pattern that mirrors the natural population distribution, i.e., locations with more population having more sample than areas with less population. Another limitation will be the likely less than adequate sample capture of women who are currently pregnant (this is about 4% of all women) and those who say they are planning to get pregnant (a percent likely larger than those who are pregnant). These are rare groups for which this sample and its approach will not support meaningful estimates. An assist to a solution may be to include husbands/partners of currently pregnant women and those planning to get pregnant as offering some insight into the knowledge, attitudes and behaviors being investigated. Other limitations are that persons without any telephone service (about 2% of U.S. adults) and those living in institutional settings will not be represented.

Random-digit dial samples will be drawn from cellular and landline telephone frames using an optimal allocation between these two frames. The optimal allocation of interviews between the landline and cell frames is a function of the cost per interview (CPI), coverage and non-response properties of each frame. With current landline and cell CPIs nearly identical (cell is slightly higher by about 10%-15%), the deciding factor is which allocation will reduce the amount of weighting required and, in turn, provide the best precision.

Using the data collected in wave 1, we were able to compute the optimal allocation of interviews between the cellular and landline telephone frames such that minimizes the design effect (Deff) due to the telephone service calibration:

- Puerto Rico: 75% cellular / 25% landline
- Miami MDA: 80% cellular / 20% landline
- Houston MDA: 90% cellular / 10% landline
- Mississippi: 85% cellular / 15% landline

While the design effect from telephone service calibration is just one component of the overall design, it is noteworthy because it is directly affected by the allocation choice. Improving the representativeness of sample demographics is another important reason to increase cell allocation. For example, young adults are significantly underrepresented in landline RDD samples but well represented in cell samples. Therefore, the proposed allocations with a higher allocation of interviews to the cell sample will both reduce the Deff and greatly assist in achieving the target oversample of adults aged 18-44 years.

2. Procedures for the Collection of Information

Once OMB approval is received, a random probability telephone survey will be conducted in the geographic target areas for the campaign.

Independent general population samples will be drawn separately for the U.S. domestic surveys and Puerto Rico surveys. For each survey, samples will be drawn from both the landline and cellular random digit dial (RDD) frames to represent people with access to either a landline or cell phone. All samples will be provided by Survey Sampling International, LLC.

Landline RDD Sampling Frame
The landline RDD frame for the U.S. domestic survey is constructed by compiling all telephone exchanges in the 3 target regions in the U.S. The frame is referred to as "list-assisted" because a complete file of directory-listed residential numbers is used to remove 100-banks from the frame if they contain zero residential listings [note: a "bank" consists of the numbers from 00 to 99 as the last two digits (xx) in a 10-digit telephone number including area code and exchange, i.e., (123) 456-78xx]. The remaining 100-banks are "working" and used to enumerate all the telephone numbers within the bank from which a sample is drawn. All landline numbers (directory-listed and unlisted) in the working banks are eligible to be randomly selected. The landline sample is drawn proportional to the share of landline telephone numbers in each of the market areas sampled.

For the Puerto Rico surveys, the landline RDD telephone frame is constructed by compiling all landline telephone exchanges in Puerto Rico and then drawing a random sample of telephone numbers from 1000-blocks that are known to be in service and residential [note: a 1000-blocks consists of the numbers from 0000 to 9999 as the last four digits (xxxx) in a 10-digit telephone number including area code and exchange, i.e., (123) 456-xxxx]. Every working residential landline telephone number in Puerto Rico will have an equal chance of being selected in the landline sample.

Cellular RDD Sampling Frame
For the U.S. domestic surveys, the cellular RDD telephone frame begins with 1,000-blocks constructed from exchanges that provide cellular telephone service in the 3 target regions in the U.S. The frame of 1,000-blocks is then expanded to the 100-block level to identify and remove shared or "mixed use" 100-blocks and those that include landline numbers. The result is a systematic sampling of cellular 100-blocks that is mutually exclusive of the list-assisted RDD

sampling frame described above. The cellular sample is then drawn proportional to the share of cellular telephone numbers within each of the 3 target regions.

The cellular RDD frame in Puerto Rico consists of 1000-blocks constructed from exchanges that provide cellular telephone service. The frame of 1,000-blocks is then expanded to the 100-block level to identify and remove unassigned 100-blocks and those that include landline numbers. The remaining cellular 100-blocks are then systematically sampled to provide a random sample of all "working" cellular telephone numbers in Puerto Rico.

Data will be collected using the Zika Readiness Initiative Survey (Attachment D (English) and Attachment E (Spanish)) 12 months post-launch to assess long term outcomes of the campaign initiative.

Twenty four hundred questionnaires will be implemented (600 in Puerto Rico; 1800 in U.S.). Interviewing will be conducted by telephone utilizing a computer assisted telephone interviewing (CATI) software in English and Spanish, based on respondent preference. Interviews will average approximately 14 minutes.

During recruiting, telephone numbers of potential and actual participants will be collected via the random digit dialing system to facilitate participation. These data will be maintained locally in the secure online scheduler to which only local research staff will have access. This system will not be linked to the screening or individual interview data in any way that could connect a participant's identity to his/her responses. The number of staff with access to this information will be kept at the minimum necessary. Contact information for study participants will be destroyed after recruitment is completed. Participants that are not reached by phone will be left a voice mail message (Attachment G) and there is a script for interviewers when incoming calls are received (Attachment H).

As discussed in more detail above, at least 75% of the completed interviews (n=1,800) will come from persons in the 18-44 year old age group, considered reproductive ages for women and equal-age partners. Survey data will be weighted to make results representative and generalizable to the target areas, adjusting for the dual-frame and oversample design.

3. Methods to Maximize Response Rates and Deal with Non-Response

As our data collection methodology involves random digit dialing, the pool of potential participants that we can draw from is large. The questionnaire was developed with focus on the length of the questionnaire and the amount of time required to complete the survey over the phone. We kept the timing to a maximum of 14 minutes to keep the burden low as well as to keep response rates high.

The current science of dual-frame telephone survey methodology is challenged to obtain double-digit response rates as calculated using the rigorous computational standards set by the American Association for Public Opinion Research. General population surveys are more likely to obtain response rates in the range of 9% to 20% depending on sponsor recognition, salience of the topic, the use of cash incentives, extended follow-up efforts and in some cases, where possible, the use of mailed advance notice.

The following methodology and best practices will be employed to obtain the best possible response rate:

(1) Informing respondents of what the study is asking and why it is being asked
(2) Using bilingual and bicultural interviewers and culturally and linguistically appropriate data collection instruments
(3) Providing easy access to research instruments by calling over the phone and completing the interview with the participant
(4) Addressing confidentiality and anonymity with respondents: since respondents who know their answers will not be linked to them in any way will be more likely to respond and more likely to provide truthful responses
(5) Minimizing study length while maximizing the richness of data that can be obtained. Respondents will be told how much time the questionnaire will take to complete so they know what to expect.

Once all the data are collected, weights are used to rebalance the sample across multiple geographic and demographic categories to reflect more accurately those same dimensions of the target population. The target benchmarks are derived from official U.S. Census data sources that also cover Puerto Rico. Using a raking procedure, these survey data will be calibrated to control totals that represent the corresponding domestic adult population for the pre-defined U.S. sample area and the Puerto Rico adult population for the Puerto Rico sample. Depending on what the final data will support, the sample will be weighted to controls within stratum on such dimensions as gender by age groupings, level of education by age group, race/ethnicity, marital status, metropolitan area or not, and a broad geographic location dimension. All these adjustments together are designed to minimize non-response bias and produced unbiased estimates from the collected data for each of the samples.

4. Test of Procedures or Methods to be Undertaken

The data collection tool was initially tested in-person with internal study staff to assess the question wording and time to complete the survey. After incorporating feedback from study staff, the survey was pretested with 9 participants by phone. One of the overarching goals of this stage of pretest was to field the comprehension of the questions and the ease with which participants had in answering them. Another goal of this stage of the pretesting was to determine the length of time required to complete the survey interview within the study population.

Based on pre-testing, no changes were identified to the wording of individual items. Many of the items used in the survey instrument were adapted from CDC's Health Message Testing System (HTMS) so no additional cognitive testing of items was conducted. The pre-testing findings indicated that the length of time required to complete the interviews exceeded expectations. As a result, several questions were removed to maintain the maximum time frame of 14 minutes.

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

The following individuals, including contractors, provided advice about the protocol design, sampling methods, and data collection tools:

Lynn Sokler, BS
Senior Communication Advisor
Center for Disease Control and Prevention
1600 Clifton Road N.E.
Atlanta, GA 30333
Phone: (404) 498-6617
Email: zsz0@cdc.gov

Fred Fridigner, DrPH
Health Communications Specialist
Center for Disease Control and Prevention
2500 Century Blvd.
Atlanta, GA 30345
Phone: (404) 639-0632
Email: fwf0@cdc.gov

Selena Ramkeesoon, MBA
Vice President, Strategic Communications
Abt Associates Inc.
4500 Montgomery Ave., Suite 800 North
Bethesda, MD 20814
Phone: (301)347-5789
Email: Selena_Ramkeesoon@abtassoc.com

Cynthia Klein, PhD
Principal Associate/Scientist
Abt Associates Inc.
2200 Century Center Pkwy
Atlanta, GA 30345
Phone: (404) 946-6310
Email: Cynthia_Klein@abtassoc.com

Ann Cloud, MPH
Associate
Abt Associates Inc.
2200 Century Parkway
Atlanta, GA 30345
Phone: 404-946-6303
Email: Ann_Cloud@abtassoc.com

Mark Morgan, MS
Senior Vice President
Abt SRBI
180 Maiden Lane, Suite 802
New York, NY 10038
Phone: 646-486-8415
Email: M.Morgan@srbi.com

Daniel Loew, MA
Senior Survey Director
Abt SRBI
275 Seventh Ave., Suite 2700
New York, NY 10001
Phone: (646)486-8421
Email: D.loew@srbi.com

ATTACHMENTS

Note: Attachments are included as separate files.

Attachment G - Voicemail Message

Attachment H - Incoming Call Script