Food Safety Survey

0910-0345—Reinstatement

SUPPORTING STATEMENT

B. STATISTICAL METHODS

1. Respondent Universe and Sampling Methods

The respondent universe for this survey is all telephone households in the United States. Eligible households are defined as those containing one or more adults who have the following characteristics: aged 18 years or older, speak English or Spanish, sufficiently good health for a telephone interview.

Households will be selected using a Random Digit Dialing (RDD) procedure by employing GENESYS, an in-house database-assisted sampling system. The GENESYS system uses a database of working residential telephone banks for the entire United States to produce a single-stage random sample of residential telephone numbers. RDD samples from the GENESYS system eliminate the reduction in precision caused by the multi-stage cluster designs of traditional RDD procedures. GENESYS samples are widely accepted because of their methodological rigor and efficiency.

The GENESYS database is constructed from three sources: a master list of area code-exchange combinations obtained from BELLCORE, a summary file of listed telephone numbers in the continental United States obtained from Donnelly, and a summary file obtained from CATI and other sources that cross-reference zip codes to telephone exchanges. The telephone numbers in these sources are matched and analyzed to produce a database of two-digit banks that contain at least 99 percent of the eligible telephone numbers in the continental U.S. (A two-digit bank consists of the first eight digits of a 10-digit telephone number within which up to 100 telephone numbers could be assigned, e.g. 703/790-90xx). The database is used to generate a random sample in which every telephone number, whether listed or not, has an equal probability of selection. The sample, unlike a traditional RDD sample, has no design effect associated with clustering of telephone numbers within telephone exchanges.

Using GENESYS, the U.S. population can be stratified into high incidence versus low incidence strata. This approach produces unique geo-demographic samples. FDA's contractor, RTI, will use this methodology to obtain an over-sample of Hispanics by generating a stratified probability sample.

As proposed, the sample design will use the GENESYS sampling system to generate random samples of telephone numbers. The GENESYS system uses listed telephone numbers to identify blocks of working telephone numbers. At the same time, it records the geographic location of the telephone and associates it with known U.S. Census information such as the Census Tract. With this assignment, the GENESYS system can provide estimates of demographics

characteristics for an area code and telephone exchange combination such as the number of households, population size, number of Hispanics, number of Blacks or African-Americans, etc. RTI will use this information to create the sampling plan. Although it is not finalized, the sampling plan is likely to include five strata – one stratum of all the area codes and exchanges with high concentrations of Hispanic population and four additional strata for the remainder of the U.S. population – one for each U.S. Census region. An advantage of the GENESYS system is that it does not introduce clustering into the design; therefore, the only contribution to the design effect will be the allocation of the sample to strata.

The survey will oversample Hispanics. The target number of Hispanics, 500 out 4,000 interviews, is 12.5%. This is the approximate proportion of Hispanics in the population. However, a simple random sample usually yields only 8-10% Hispanics in the sample due to differential response propensity, access to telephones, multiple lines per household, differential household composition (more than one family unit per household), among other reasons. The stratified approach allows us to control for this and ensure the proper number of Hispanic respondents by allocating slightly more than the amount needed for a proportionate allocation. This means that the sample design itself will have a design effect around 1.045 or the design will only be 95.5% efficient due to the Hispanic oversampling.

The process for accounting for the oversample in the weights is covered in the steps already documented for the Food Safety Study in 2002 – "*Food Safety Study Sample Weight Documentation -- 2002 Recalculation for the 1988, 1993, 1998 and 2001 Waves.*" Two final weights will be calculated. The first final weight in the FSS is a post-stratified weight to adjust the sample distribution to match the population distribution for

-- Gender (2 categories);
-- Educational attainment (4 categories); and
-- Race/ethnicity (4 categories).

The second final weight will also adjust the sample distribution to match the population distribution for age (3 categories). This second weight was first calculated for the 2006 Food Safety Survey after noticing that older people are more likely to respond to the survey.

The final weight is

$$FINALWT_j = K_{FW}\, d_j\, P_i / E_i \text{ for respondent } j \text{ in weight cell } i$$
$$\text{where } K_{FW} = n \Big/ \sum_{j=1}^{n} d_j\, P_i / E_i$$

Here P_i is the proportion of the population in weight cell i that contains respondent j, E_i is the proportion in the sample calculated using the design weight, and K_{FW} normalizes the weights to the sample size. FINALWT is interpreted as the design weight multiplied by (the proportion of population in weight cell i divided by the proportion of design-weighted sample in weight cell i) and normalized to the sample size n. Note that K_{FW} is a constant because both n and the divisor, which is a sum over the entire sample, are constants. The properties of these weights are that

they sum to the sample size and the average weight is one. The design weight, d_j , will be a combination of the household size, the number of telephone lines and the weight for the stratum.

A sample size of 4,000 is needed to adequately perform all of the sub-analyses that are needed for the proposed Healthy People 2020 objective relating to consumer practices. This objective requires that we look at disparities in safe food handling habits among different populations. Some of these sub-analyses of interest include showing change over time for: Blacks or African-Americans, older adults (60+), and households with young children (<5 years old). A power analysis was conducted to determine the sample size needed to analyze changes in these sub-groups.

Identification of the designated respondent (DR) will be achieved by using the last birthday method. Once household eligibility has been established, interviewers will ask the person who answers the phone to speak with the adult age 18 and over in the household has had the most recent birthday. This method has been used for all five waves of the Food Safety Survey.

The response rate for the 2006 Food Safety Survey was 33.8% based on the American Association for Public Opinion Research (AAPOR) Response Rate 5.

2. Procedures for the Collection of Information

The Food Safety Survey will be conducted using computer-assisted telephone interviewing (CATI) technology. The interview will consist of two parts: the household screener and the extended interview (Attachment A). The household screener will be used to locate eligible households and to identify a DR as described in section B.1 above. Only one respondent per household will be interviewed.

Data will be collected by experienced and specifically trained telephone interviewers. Quality control will be assured by periodic monitoring of on-going interviews throughout the study. This monitoring replaces the previously used validation interview, which required maintaining the name and telephone number of the respondent until the validation interview could be completed.

3. Methods to Maximize Response Rates

To help ensure that the response rate is as high as possible, RTI will employ all appropriate methods demonstrated in the research literature on survey methodology. These procedures include the following:

- Design a questionnaire that minimizes respondent burden (short in length, written in easy-to-understand language).
- Test the draft questionnaire using cognitive interviews to ensure that respondents can properly understand the questions and that the response options are robust.

- Test the draft questionnaire in a pre-test to ensure that it minimizes burden and refine as appropriate.
- After the sample is drawn, all households for which an address can be matched to the telephone number will be sent a letter letting them know that they have been selected to participate in the survey (Attachment B).
- In addition to general training, all interviewers and supervisors will be trained on the specifics of the survey by a member of the project's professional staff. This will include an explanation of the importance and purpose of the survey, as well as a thorough review and practice reading of the entire survey instrument.
- A Spanish-speaking interviewer will re-contact all households in which the interview could not be completed because of a language barrier. Households in which neither English nor Spanish is spoken sufficiently to allow for completion of the interview will be excluded.

- All interviews are continuously monitored by telephone supervisors who listen to a portion of each call to ensure that each interview is conducted properly. Production rates and sample dispositions will be monitored each day to detect and resolve any problems or discrepancies quickly.

- A reasonable number of call attempts will be made to determine whether an "initial contact"—the establishment of the identity of a telephone number (residential or non-residential)—is made. For example, if the first three attempts received no response and the fourth attempt received a busy signal, the number will be called for a few more times to try to make an initial contact because the fourth attempt suggests this number has the potential of being a residential number. Only when there is certainty that a number is not a residential number will the limit of five attempts be applied. If a voicemail or answering machine indicates the number is residential, then an initial contact is considered made.

- No-answers after these attempts at initial contact will be regarded as non-households and eliminated from the sample. Households that initially refuse to participate will be sent a letter acknowledging the initial contact and asking again for the household's participation. Addresses will be obtained through a commercial list of known telephone number/address combinations. The letter will identify FDA as the sponsor of the survey, give a brief explanation of the study topic, and stress the importance of participation. Refusal conversion calls will be scheduled several days after the letters are sent out, in order to give the letter ample time to arrive, but close enough to the arrival date to be remembered by the respondent (Attachment C).

- When possible, household screening and extended interviews with designated respondents will be completed during the same call. If the DR is not available at the time of the screening call, up to 25 callbacks will be made in an effort to complete the interview. DRs who are not reached will be included in the denominator for the calculation of the response rate. DR's who initially refused will be sent a letter encouraging participation if an address match can be made.

- To determine if there are any systematic differences between those who participate in the FSS and non-responders, FDA will conduct a non-response analysis. Two hundred initial refusals (i.e., non-respondents) who have refused to participate twice; once when first called and a second time when called back, will be asked to take a shortened questionnaire consisting of core questions and a subset of the demographic questions (Attachment D).

4. Test of Procedures or Methods

In May 2009, four focus groups were held to develop the new module on food recalls and new questions on microwave safety. Twenty cognitive interviews will be conducted shortly after OMB approval to further refine the module on food recalls and the other small changes to the questionnaire since the 2006 Food Safety Survey.

The Food Safety Survey will be pre-tested with up to 27 respondents shortly after OMB approval of the information collection is expected. Scheduling the pretest close to the beginning of data collection will gain efficiency by using interviewer training for both the pretest and the complete data collection. Because the survey is based largely on questions from the 2006 survey, 27 pretests should be adequate to estimate the time required to complete the interview, to assure clarity of the added or changed instructions, questions, and response categories, and to check the CATI programming for correct skips and other procedures.

Representatives of FDA and the contractor will monitor the pretest interviews. Few changes to the questionnaire are expected from the pre-test, because we want to compare responses with identical questions asked on previous surveys. OMB will be provided with copies of the final questionnaires prior to implementation of the study.

5. Individuals Involved in Statistical Consultation and Information Collection

The contractor, RTI, will collect the information on behalf of FDA as a task order under the Quick Turn Around Survey Data Collection contract. Ms. Sheryl Cates is the Project Leader for RTI and Dr. Karol Krotki is the senior statistician for the study.