

U.S. Food and Drug Administration  
Health and Diet Survey

OMB Control No. 0910-0545

SUPPORTING STATEMENT **Part B:** Statistical Methods

1. Respondent Universe and Sampling Methods

The respondent universe for this collection of information will be non-institutionalized adults 18 and older who speak English or Spanish. Within the sampling universe, respondents will be randomly selected from two sampling frames: a national landline telephone frame and a national cell phone frame. More specifically, the Health and Diet Survey will adopt the overlapping dual frame approach as described by the American Association for Public Opinion Research, i.e., we will select independent samples from random-digit-dialing frames that overlap in their coverage (e.g., a landline frame and a cell phone frame) and we will not screen out dual-frame prospective respondents (AAPOR 2010). Within a dual frame, we will include four groups of adults: (1) users of landline telephone only, (2) users of cell phone only, (3) dual users who use landline mainly, and (4) dual users who use cell phone mainly.

For the landline interviews, households will be selected using a random-digit-dialing (RDD) procedure by employing GENESYS, a database-assisted sampling methodology. 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 United States obtained from Donnelly, and a summary file obtained from CATI and other sources that cross-references 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 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. 123/456-78xx). 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.

Identification of the designated respondent (DR) will be achieved by the most recent birthday method. Once household eligibility has been established, interviewers will ask to speak with the adult household member who had the most recent birthday. The DR will be selected prior to any questions about at-home status or availability of potential DR, and no

substitutions will be allowed. If the DR will be unavailable throughout the study period, the household will become ineligible.

The survey will over-sample African-American and Hispanic populations by dividing the landline telephone exchanges into three strata: a stratum of geographic areas with high concentrations of African-American population, a stratum of geographic areas with high concentrations of Hispanic population, and a stratum with the remainder of the U.S. The first two strata will be sampled with higher rates. These sampling rates will be determined to achieve the desired numbers of African-American and Hispanics based on estimated incidences within each stratum. The final sample numbers for Hispanic and African-Americans will occur at random from the sample without screening. The geographic areas with higher African-American and Hispanic concentrations will be identified using GENESYS. In addition to the landline telephone exchanges, we may also oversample cell phone users because research indicates that proportionately more African-Americans and Hispanics are cell phone users than other racial groups.

## 2. Procedures for the Collection of Information

The survey will be conducted by a contractor using computer-assisted telephone interviewing (CATI) technology. Each landline interview will consist of two parts: the household screener and the survey questions. The household screener will be used to locate eligible households and to identify a designated respondent (DR) as described below. Only one respondent per household will be interviewed. Each cell phone interview will only include the survey questions and will not include any within-household eligibility screening.

Information 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 and Deal with Non-response

In an effort to increase response rate, we plan to take the following measures:

- send advance letters to those households whose addresses can be found to notify them of the impending interview;
- make as many call attempts as needed, up to 30 call attempts for a landline number and 5 call attempts for a cell phone number, to complete an interview;
- keep the length of the survey to an average of 15 minutes;
- allow a field period of 120 days;
- stagger calling in various times of the day and days of the week;
- send conversion letters to those who initially refuse the interview and use conversion specialists to conduct conversion calls;
- use a Spanish questionnaire to accommodate respondents who do not speak English;

- conduct a non-response study to identify potential non-response biases and adjust estimates statistically, if necessary; and
- regular and continuous monitoring of interviewer performance and address any operational issues in a timely manner.
- non-responders for whom an address match is available in advance of the conversion attempt.

#### 4. Test of Procedures or Methods to be Undertaken

Two types of tests of the collection procedure are planned to minimize collection burden on respondents and improve quality of collected information.

We plan to conduct 18 cognitive interviews. The primary purpose of these interviews is to understand the mental processes that respondents use to answer survey questions. We will use cognitive interview findings to revise the draft questionnaire, if necessary and appropriate.

The second type of tests is field pretests focusing more on the length of the questionnaire and respondent burden in an environment as close as possible to the real interviews. The data collection contractor will administer the full instrument by telephone to 200 randomly-selected adults shortly before OMB approval of the collection of information is expected. Scheduling the pretests close to the beginning of data collection will gain efficiency by using interviewer training for both the pretests and the complete data collection. The pretests will also serve the purposes of addressing problems in respondent selection, interviewer instructions, skip patterns, and design of the computer-assisted-telephone-interview program.

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

A contractor will collect the information on behalf of the FDA. The contractor will be tasked to provide statistical consultation to the survey. Analysis of the information will be conducted primarily by staff on the Consumer Studies Team, Center for Food Safety and Applied Nutrition, FDA.