SUPPORTING STATEMENT

HEALTH AND DIET SURVEY

OMB No. 0910-0545

B. Statistical Methods (used for collection of information employing 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 in households with telephones in the 50 states and the District of Columbia. As of 1999, 94 percent of American households have telephone service.[1]

A response rate of 34 percent was achieved in the collection of a subset of the information in this survey that was conducted in 2004. The agency expects to achieve a similar or higher response rate in this collection of information.

2. Procedures for the Collection of Information

2.1 Statistical methodology for collection and sample selection

The survey will be conducted using computer-assisted telephone interviewing (CATI) technology. The interview will consist of two parts: the household screener and the core 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.

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.

---

[1] U.S. Census Bureau. 2002. Table 1126, *Statistical Abstract of the United States: 2001*. Washington, D.C.

Identification of the 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.

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.

The survey will over-sample African-American and Hispanic households by dividing the population 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.

2.2 Estimation Procedure

Each interviewed person will receive a basic sampling weight equal to the reciprocal of his or her probability of selection. The basic sampling weight will account for (1) multiple telephone numbers in households, (2) household size, and (3) nonresponse. Households with more than one residential telephone number have a greater chance of selection; therefore, sampling weights will be adjusted by the reciprocal of the number of residential telephone numbers on which the household receives calls, excluding cell phone numbers. The weights will also reflect the differential probability of selection depending on household size. For example, a person living alone would be selected with certainty, whereas a person living in a household with four other adults would have a one in five chance of being selected.

To compensate for under-coverage and to reduce the mean square error of the estimates, the final base weights will further be adjusted to match recent Census totals for sex, education, and race.

2.3 Degree of accuracy needed for the purpose described in the justification

For analyses of the *General Topics* survey, the proposed sample size (3,000 adults) will provide a precision of approximately $\pm$ 1.1 to 1.8 percentage points at the 95 percent confidence level (Table 3). For analyses of subgroups, a standard error of $\pm$ 2.5 percentage points is usually acceptable. As shown in Table 3, this level of precision will also be achieved with the proposed sample size for major demographic classifications (e.g., age, gender, education, and race) as well as major subject-matter classifications of respondents (e.g., dietary supplement users, vitamin/mineral users). For instance, suppose the collected information from 3,000 respondents yields an estimate that 80 percent (proportion = 0.8) of the sampled adults have taken one or more vitamins or minerals in the

past 12 months. We will then expect that, if the sample were drawn 100 times, in 95 times the true percentage of users will fall somewhere between 81.4 percent (80+1.4) and 78.6 percent (80-1.4).

For analyses of the *Dietary Guidelines* Supplement, the proposed sample size (1,200 adults) will provide a precision of approximately ± 1.7 to 2.8 percentage points at the 95 percent confidence level. Over-sampling of African-American and Hispanic households will achieve 200-300 completed interviews in each category. So if the collected information yields an estimate that 40 percent (proportion = 0.4) of the sampled adults report familiarity with some government information about nutrition, we will expect that in 95 times out of 100 times of sampling, the true percentage of African Americans that are familiar with said information will fall somewhere between 46.8 percent (40+6.8) and 33.2 percent (40-6.8).

Table 3. Sampling Error (± percentage points) at the 95 Percent Confidence Level for Different Sample Sizes

	Proportion				
Sample Size	0.1 (0.9)	0.2 (0.8)	0.3 (0.7)	0.4 (0.6)	0.5 (0.5)
3000	1.1%	1.4%	1.6%	1.8%	1.8%
2000	1.3%	1.8%	2.0%	2.1%	2.2%
1800	1.4%	1.8%	2.1%	2.3%	2.3%
1600	1.5%	2.0%	2.2%	2.4%	2.5%
1400	1.6%	2.1%	2.4%	2.6%	2.6%
1200	1.7%	2.3%	2.6%	2.8%	2.8%
1000	1.9%	2.5%	2.8%	3.0%	3.1%
800	2.1%	2.8%	3.2%	3.4%	3.5%
600	2.4%	3.2%	3.7%	3.9%	4.0%
400	2.9%	3.9%	4.5%	4.8%	4.9%
200	4.2%	5.5%	6.4%	6.8%	6.9%

2.4 Use of specialized sampling procedures

No specialized sampling procedures are required.

2.5 Use of periodic data collection cycles to reduce burden

This is a one-time data collection.

3. Methods to Maximize Response Rates

In an effort to increase response rate, the agency plans to take the following measures:

- send advance letters to those households whose addresses can be found to notify them the impending interview;
- make as many call attempts as needed, up to 35 call attempts, to complete an interview;
- extend data collection period from 75 days to 120 days; and
- conduct a non-response study to identify potential non-response biases and adjust estimates statistically, if necessary.

Advance letters and a longer data collection period have often been used by survey organizations as part of an effort to increase telephone survey response rates. Studying non-response may help the agency in identifying significant non-response biases. Existing research, however, has shown that non-response biases in random-digit-dialing national telephone survey may not be significant. For example, Keeter et al. (2000) found no measurable differences in findings between a survey with a response rate of 36% and an identical survey with a response rate of 61%, even though potential respondents in the latter were sent advance letters and a $2 incentive.[2]

The agency plans to make as many call attempts as needed, up to 35 call attempts, to complete an interview; the 35 attempts include a maximum of 25 attempts to complete the interview after an eligible respondent is identified. Recent research has suggested that any effort beyond 24 attempts does not change national estimates of a random-digit-dialing telephone survey and does not improve response rates by a significant degree.[3]

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 3 attempts received no response and the fourth attempt received a busy signal. Then 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 4 attempts be applied. If a voicemail or answering machine indicates the number is residential, then an initial contact is considered made.

Calls will be staggered over times of day and days of the week to maximize the chances of making contact with a household. No-answers after these attempts at initial contact will be regarded as non-households and eliminated from the sample. Whenever possible, household screening and extended interviews with designated respondents (DR's) will be completed during the same call.

In addition to the measures mentioned above, the data collection contractor will implement the following procedures to obtain the highest possible response rate:

- 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 collection of information as well as a thorough review and practice reading of the entire information collection instrument.

- Respondents who initially refuse to participate will be assigned to conversion specialists, who will attempt to complete the interview on a different day. Conversion letters acknowledging a contact attempt and describing the purpose of the study will be sent to non-responders for whom an address match is available in advance of the conversion attempt.

---

[2] Keeter, S., Miller, C., Kohut, A., Groves, R.M., and Presser, S. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." Public Opinion Quarterly 64: 125-148.

[3] Dennis, M., Mathiowetz, N.A., Saulsberry, C., Frenkel, M., Srinath, K.P., Roden, A.-S., Smith, P.J., and Wright, R.A. 1999. "Analysis of RDD Interviews by the Number of Call Attempts: The National Immunization Survey." Presented at the Annual Meeting of the American Association for Public Opinion Research.

- A Spanish speaking interviewer will recontact all households in which the interview could not be completed because of a language barrier.

- All interviewers will be monitored by a supervisor during the first day of interviewing and intermittently throughout the course of the collection of information thereafter. Production rates and call dispositions will be monitored each day to detect and resolve any problems or discrepancies quickly.

- The contractor will provide detailed descriptions of procedures for assuring quality control, for identifying interviewers who are having difficulties, and for dealing with problems.

To ensure quality control, the contractor will maintain complete call disposition records on every household contacted. In no case will telephone numbers be abandoned prior to achieving one of the following: (1) completed interview, (2) completed conversion attempt or refusal, (3) exhaustion of callbacks, (4) determination that a household is not eligible, and (5) exhaustion of initial contact attempts.

When a household is determined to be ineligible, the basis for the determination will be recorded.

The response rate for this study will be defined as follows: completed interviews / (completed interviews + terminations + interview refusals + screening refusals).

4. Tests of Procedures or Methods

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

The first type of tests is cognitive interviews; the primary purpose of these interviews is to understand the mental processes that respondents use to answer survey questions. Nine randomly selected adults were asked dietary supplements questions contained in earlier drafts of the instrument, and probed the mental processes they went through in providing the answers. The focus of analysis was on (1) comprehension of the meaning of certain questions or words, and (2) strategies used to recall information and to arrive at an answer.

In producing the submitted instrument (Appendix B), the agency has considered findings from the cognitive interviews.

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 nine 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 Involved in Statistical Consultation and Information Collection

For the General Topics survey, the contractor, Westat, Inc. will collect the information on behalf of the FDA as a task order under the Quick-Turn-Around Research Services contract. Pat Dean Brick, Ph.D., is the Senior Study Director for Market Facts, telephone 301-251-4382. Analysis of the information will be conducted primarily by staff on the Consumer Studies Team, Division of Market Studies, CFSAN, FDA, and coordinated by Conrad J. Choinière, PhD, telephone (301) 436-1844. For the *Dietary Guidelines* Supplement, Synovate, Inc. will collect the information. Valerie Fuller, Ph.D., is the Project Director. Analysis of the information will be conducted by Porter Novelli under the supervision of Adams Burn.