United States Food and Drug Administration

Web-Based Pilot Survey to Assess Allergy to Cosmetics in the United States

OMB Control No. 0910-NEW

SUPPORTING STATEMENT

Part B.   Statistical Methods

1. <u>Respondent Universe and Sampling Methods</u>

   The sampling frame for the study is GfK's online consumer panel, KnowledgePanel (KP). KP is designed to be representative of the U.S. adult population. KP was selected because of the extensive number of panelists, the inclusion of panelists from non-Internet households, and the ability to generalize study findings to the larger U.S. adult population.

   Currently, KP consists of approximately 55,000 adults who were randomly selected and invited to participate as panelists. Panelists typically serve for 2 to 3 years and complete two to three surveys per month during their tenure. To ensure full representation of U.S. adults, GfK equips panelists from non-Internet households with access to the Internet and a tablet computer if needed. The panel also includes households that have listed and unlisted telephone numbers and those without a landline telephone.

   GfK recruits KP members using random sampling strategies, which is distinct from other online panels that use opt-in recruitment methods and allow individuals to volunteer as panelists. Random digit dialing (RDD) was originally GfK's primary sampling strategy. GfK began supplementing RDD with address-based sampling (ABS) in 2008 and eliminated RDD sampling altogether in late 2009.

   With ABS methodology, households are randomly selected based on mailing address, including post office boxes and rural route addresses. Business and institutional addresses (e.g., dormitories, nursing homes, group homes, and jails) are removed from the frame, as is military housing. Specifically, GfK randomly samples mailing addresses from the U.S. Postal Service's Delivery Sequence File (which covers 97% of U.S. homes) and invites sampled addresses to join KP through a series of mailings, including an initial invitation letter, a reminder postcard, and a subsequent follow-up letter. About 5 weeks after the initial mailing, GfK conducts telephone recruitment calls if an address and telephone number could be matched (about 45% of the physical addresses can be matched to a corresponding landline telephone number). After receiving the mailing and/or recruitment call, households that accept the invitation join KP through one of three methods:

   1. Completing a paper enrollment form and returning it in a postage-paid envelope,

2. Calling GfK's toll-free hotline and completing an enrollment interview, or

3. Visiting a secure GfK Web site and completing an online enrollment form.

After an individual accepts GfK's invitation to be a panel member, s/he is instructed to log on to a secure GfK Web site and complete individual profile questionnaires. The questionnaires capture essential demographic information (e.g., sex, age, race, income, education) and health information (e.g., diagnoses, health status, health behaviors). These profiles enable GfK to prescreen potential participants for eligibility and to eliminate demographic questions on every survey, and can also be used in a nonresponse bias analysis. Each panelist updates his or her profile annually.

Between 2009 and 2011, GfK intentionally oversampled Hispanic and non-Hispanic households (aged 18 to 29) to help ensure sufficient representation. In 2012, GfK again oversampled Hispanic and non-Hispanic households (aged 18 to 29) but also included the age group 30 years and older. The panel's weighting procedures adjust for the oversampling carried out to improve the demographic composition of the panel.

For selection of general population samples (i.e., adults 18 and older) from the KP, GfK uses a patented methodology that ensures the resulting sample represents an equal probability of selection method sample. First, the entire KP (i.e., the sampling frame) is weighted to the latest Current Population Survey (CPS) (March 2016) benchmarks to compensate for any minor misalignments that may result due to differential attrition rates among hard-to-reach subgroups. This step helps to ensure that the weighted distribution of KP aligns with the U.S. population of adults. The geo-demographic dimensions used for weighting the entire KP include the following:

- gender (male/female)

- age (18 to 29, 30 to 44, 45 to 59, and 60+)

- race/Hispanic ethnicity (white/non-Hispanic, black/non-Hispanic, other/non-Hispanic, 2+ races/non-Hispanic, Hispanic)

- education (less than high school, high school, some college, bachelor's and beyond)

- Census region (Northeast, Midwest, South, West)

- household income (under $10,000, $10,000 to <$25,000, $25,000 to <$50,000, $50,000 to <$75,000, $75,000+)

- home ownership status (own, rent/other)

- metropolitan area (yes, no)

- Internet access (yes, no)

Using the above weights as the measure of size (MOS) for each panel member, in the next step a probability proportional to size (PPS) procedure is used to select study-specific samples. It is the application of this PPS methodology with the above MOS values that produces fully self-weighting samples from KP. For this study, an equal probability of selection method sample of 1,667 KP members will be selected to achieve 1,000 completed surveys.

Once the study sample has been fielded, and all the survey data are cleaned, design weights are adjusted to compensate for any differential nonresponse that may have occurred during the data collection process. Final analysis weights are produced using an iterative proportional fitting (raking) procedure to ensure that the resulting weights are aligned with respect to all study benchmark distributions simultaneously. In the last step, calculated weights are examined to identify and, if necessary, trim outliers at the extreme upper and lower tails of the weight distribution. The resulting weights are then scaled to sum to the total sample size of all eligible respondents.

2. Procedures for the Collection of Information

Once the sample for this study is randomly selected, GfK will send sampled panelists an e-mail invitation to invite them to participate in the study. The e-mail invitation will contain a unique link, specifically for that panelist, to the survey. After clicking on the link in the e-mail invitation, panelists will be directed to the online instrument. On the first screen, panelists will be provided information on informed consent and asked if they would like to proceed with the survey. If panelists accept, they will proceed with the survey. If panelists decline, they will be categorized as nonrespondents.

We estimate that it will take respondents no more than 20 minutes to complete the online survey. Data collection will take up to 4 weeks to complete.

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

Based on experience conducting 20-minute online surveys with general population samples (i.e., adults 18 and older), GfK estimates that about 60% of the randomly selected panelists will complete the online survey. GfK will send a maximum of three automatic e-mail reminders to nonresponding panelists during the course of data collection. If panelists did not complete the study, they will be categorized as nonrespondents.

To maximize participation, we will conduct cognitive interviews and pretests to help improve the understandability and usability of the questionnaire, reduce participant burden, and enhance administration.

In addition, to encourage participation, each e-mail invitation and reminder will state the study purpose and identify FDA as the study sponsor (see Appendix B for e-mail invitation and reminders) and provide an e-mail address and toll-free number for selected

panelists to acquire additional information about the study or verify the authenticity of the study. GfK will monitor all phases of sampling and data collection and resolve any problems immediately throughout the course of the study.

Nonresponse bias

Because we anticipate that the response rate will be less than 80%, a nonresponse bias analysis will be conducted to evaluate the potential for bias due to nonresponse. Nonresponse may cause bias in survey estimates if sample members who chose not to respond would have provided answers that differ systematically from answers provided by sample members who chose to respond.

At the conclusion of the survey, GfK will provide demographic data of all panel members selected for the survey, including nonrespondents. The demographic information (i.e., age, gender, race/ethnicity, educational attainment) of all sampled panel members will be compared with U.S. Census benchmarks and the bias due to nonresponse estimated. This analysis will be conducted using the final weights. The results of these analyses will be reported alongside summary statistics and other information resulting from this proposed data collection.

4. Test of Procedures or Methods to be Undertaken

After the online instrument is programmed, RTI will conduct cognitive interviews with nine adults to identify survey questions and response items that are confusing or difficult to understand. If necessary, the online instrument will be refined based on the cognitive interview findings.

To ensure that the programming logic, sample distribution and fulfillment, and data compilation are functioning correctly, GfK will conduct a pretest with a total of 100 randomly selected panelists. If necessary, the online instrument will be refined based on the pretest findings.

If changes are made to the survey instrument based on the cognitive interview and pretest findings, a revised survey instrument will be submitted to OMB for approval. Data collection will not commence until OMB approval is obtained for the final survey instrument.

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

The contractor, RTI, will collect the information on behalf of the Agency. Dallas Wood, Ph.D., is the project lead at RTI. Katherine Kosa is the task leader for the Web-based survey, and Sheryl C. Cates is the Senior Advisor for the project. Data analysis and dissemination will be led by John Gasper, CO.