Supporting Statement B For:

Family Life, Activity, Sun, Health, and Eating (FLASHE) Study **(NCI)**

August 30, 2013

Linda Nebeling, PH.D., MPH, RD, FADA
Division of Cancer Control and Population Sciences
National Cancer Institute
9609 Medical Center Drive, Room 3E102
Bethesda, MD 20892-9671
Tel: 240-276-6855
nebelinl@mail.nih.gov

Table of Contents

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.....1

B.1 RESPONDENT UNIVERSE AND SAMPLING METHODS.....1

B.2. PROCEDURES FOR THE COLLECTION OF INFORMATION.....8

B.3 METHODS TO MAXIMIZE RESPONSE RATES AND DEAL WITH NONRESPONSE.....11

B.4 TEST OF PROCEDURES OR METHODS TO BE UNDERTAKEN.....12

B.5 INDIVIDUALS CONSULTED ON STATISTICAL ASPECTS AND INDIVIDUALS COLLECTING AND/OR ANALYZING DATA.....14

LIST OF ATTACHMENTS

Attachment A – Privacy Impact Assessment (PIA)

Attachment B – List of Individuals Consulted

Attachment C – Privacy Act Memo

Attachment D – Westat Confidentiality Pledge

Attachment E – IRB Approvals

Attachment F – Parent Self-Enrollment and Consent Instrument

Attachment G1 – Parental Enrollment and Consent Instrument for Adolescent

Attachment G2 – Adolescent Assent Instrument

Attachment H – Adult Physical Activity Survey Instrument (includes demographics)

Attachment I – Adolescent Physical Activity Survey Instrument (includes demographics)

Attachment J – Adult Diet Survey Instrument

Attachment K – Adolescent Diet Survey Instrument

Attachment L – Wear Log Instrument

Attachment M - FLASHE Study Flow Chart

Attachment N - Email Correspondence Related to Enrollment, Website, and FAQs

Attachment O – Correspondence Related to Survey

Attachment P – Correspondence Related to Motion Sensing Study

Attachment Q – References

B. STATISTICAL METHODS

B.1 Respondent Universe and Sampling Methods

Sample Frame

The Family, Life, Activity, Sun, Health and Eating (FLASHE) study will be conducted among 12-to-17-year-old adolescents paired with their parent/caregiver in the United States. In this study, parent is defined as an adult who takes care of the adolescent and lives with the adolescent at least 50% of the time during the week. This could be a birth parent, an adoptive parent, or a guardian (an adult relative or an adult who is not related to the adolescent).

Westat/NCI will purchase a sample of 4,155 parent/adolescent dyads from Ipsos Interactive Services (Ipsos). Ipsos will use their Consumer Opinion Panel (COP) to create the dyadic sample according to Westat/NCI specifications. Each dyad provided by Ipsos will consist of a parent (selected from the COP) and an adolescent aged 12-17 (selected from eligible children residing with the parent). Ipsos's existing COP, which makes up the sampling frame for selecting the parents, has been operated by Ipsos since the 1970's and covers 45 counties. The US panel includes over 600,000 panel members over the age of 18. The panel members have volunteered to take part in surveys and Ipsos maintains extensive profiles of the panel members. Approximately 25% of the panel is replaced annually and the average length of tenure is 5.9 years. On average, panel households are contacted about once a month to participate in Ipsos web-based surveys. Because Ipsos maintains contact with panel members via the Internet, 100% of Ipsos members are Internet users.

Use of panel methodology

The sample for this study will be drawn from the Ipsos Consumer Opinion Panel (COP), and thus based on panel members rather than a probability sample drawn from the US population. FLASHE will be selecting a sample that reflects the distribution of the general US population with regard to certain characteristics by using specific balancing techniques described below. FLASHE seeks to explore correlates and relationships between cancer preventive behaviors (e.g., diet and physical activity) and psychosocial constructs, parenting, and home and neighborhood environments among dyads of adolescents and parents versus a surveillance study which would require a nationally representative sample to identify population-level behavioral point estimates. Generalizing to the US population is not the goal of the FLASHE study.

Within current government budget constraints, identifying parent-adolescent dyads via random selection (e.g., Random Digit Dial--RDD) would be cost prohibitive. The cost of randomly calling households to try to identify households with eligible respondents would increase the cost of the study by several hundred thousand dollars. Another cost efficiency in using the COP approach is that Ipsos will be providing a working email address for each sampled respondent, which provides an economical way for study staff to contact study participants and invite them to the web survey. These efficiencies would not be the case with a RDD sample and so it is unlikely that web-based data collection would be feasible without the use of the COP.

Using this alternative sample frame will also likely increase the response rate (see further discussion in B.3). Because the COP members are accustomed to communicating with Ipsos via

the Internet and completing web surveys for Ipsos, the sampled respondents will be familiar with and comfortable completing surveys using the web survey mode.

In addition, the panel provides information that has previously been collected from panel members. This information provides extensive information for sample design and weight adjustments for nonrespondents. In this study, the race information of the panel member will be used for oversampling non-Hispanic blacks. In addition, the demographic information to be provided by Ipsos will include characteristics of the panel member (e.g. age, gender, education level, employment status and occupation) and household information (e.g. Geographic region, MSA status, population density, household income, household size, dwelling type and home ownership.) These variables will be available for both survey respondents and non-respondents and will be used as covariates in adjusting analysis weights for non-respondents.

Panel methodology has been used successfully by NCI in the past. For example, the 5 A Day Customized Survey (OMB No. 0925-0560, Approved 3/31/2006) was conducted using COP methodology in 2007. Use of the panel enabled NCI to make the most efficient use of federal funding while still obtaining significant and useful scientific results to inform behavioral nutrition research across the disciplines of nutrition, epidemiology, public health, psychology, and nursing. Key findings from 5 A Day data analyses include the role of the neighborhood and individual self-efficacy for eating fruits and vegetables, evaluation of awareness of the Federal Fruit and Veggies More Matters Campaign (Erinosho, et al, 2012), the role of drinking water on health (Goodman, et al, 2013), examination of behavioral health theory and explaining gender differences in fruit and vegetable consumption (Emanuel, et al, 2012), worry in predicting

nutrition behavior (Ferrer, et al 2012), survey instrument validation (Yaroch, et al, 2012), grocery shopping behaviors and farm-to-consumer vender use (Blank,et al, 2001). The 5 A Day data and survey items have also been used in grant applications and for student training and research. These examples of research using 5 A Day data as well as other ongoing research illustrate the value of behavioral and psychosocial data collection and the value of COP sampling methodologies, in particular, to allow for efficient data collection in behavioral health studies to inform health behavior intervention, theory development and measurement. The Centers for Disease Control and Prevention's on going use of a COP for the Health Styles Survey is another example of the use of panel methodology to conduct survey research and collect behavioral information data to inform intervention and public health programming (Masse, et al, 2012).

Sampling Plan

For the survey, the initial sample size is expected to be approximately 4,155 to achieve 2,500 completed dyadic surveys (both adult and adolescent) at a 60 percent response rate. Ten years of past experience with this particular panel has yielded person-level response rates ranging from the high-50's to 80 percent with a median response rate of 65 percent. However, because the adolescents in the dyad are not part of the COP and are therefore an unknown component, we anticipate this could slightly lower the overall response rate for the FLASHE study. The 4,155 dyads will be stratified by race, oversampling non-Hispanic Blacks. Using the sample instructions provided by Westat/NCI (and shown on Table B.1), Ipsos will select a set of potential participants from each stratum. A subsample of 1,500 adolescents will be randomly selected by Westat for participation in the Motion Sensing Study (a component of this request). The same stratification and oversampling will be applied to this subsample, as shown in Table B.2.

Table B.1 Sample size and expected numbers of participating dyads in the sample completing all the web survey measures by race (assuming a 60% response rate)

Race	Non-Hispanic Black	Others	Overall
Sample size (percentage)	1,039 (25%)	3,116 (75%)	4,155 (100%)
Sample participants (percentage)	625 (25%)	1,875 (75%)	2,500 (100%)

Table B.2 Sample size and expected numbers of participating dyads in the sample completing the motion sensing study by race (assuming a 60% response rate)

Race	Non-Hispanic Black	Others	Overall
Sample size (percentage)	375 (25%)	1,125 (75%)	1,500 (100%)
Sample participants (percentage)	225 (25%)	675 (75%)	900 (100%)

Sample balancing

The sampling goal for FLASHE is to achieve a sample that demographically matches the US population on key demographic characteristics. Ipsos will use balancing techniques to ensure the sample of adults mirrors the distribution of the US population for the key demographic characteristics to be specified by NCI/Westat: geographic region, household income and household size. In addition, some characteristics of the adolescent will be used: age of the adolescent child, sex of the adolescent child. To draw such a sample, a target distribution based on the US population will be specified by Westat for each demographic characteristic to which the sample is to be balanced. These target distributions are marginal distributions on each of the demographic characteristics and reflect the distribution of the target population on each characteristic.

The target distribution for the cells defined by the cross-classification of all demographic characteristics is not directly specified. A sample-balancing algorithm is used to determine the number of households that are to be selected from every cell. The balancing algorithm first determines the distribution of available panel households for the cells across all characteristics to which the sample is to be balanced. The algorithm adjusts this distribution of households across the cells so the marginal distributions then match the targets. The resulting adjusted number in each cell is then the "target" for that combination of all the characteristics and is the number of households that will be selected from that cell.

Iterative proportional fitting is employed in the algorithm to adjust the distribution of available households one characteristic at a time until all the marginal distributions of the sample match the target distribution. This is the same technique that is typically used in marginal weighting programs. Once the distributions are matched, the number of households to be selected from every cell is determined and then the actual selection of households takes place. Random selections are made from those households available in the cell.

Sample size and Power

The anticipated sample of 2,500 completed dyads will provide adequate statistical power for examining correlates of cancer preventive behaviors, mainly diet, activity and sedentary behaviors, but also examining other behaviors such as sleep, sun-safety, and tobacco. The oversampling of non-Hispanic Blacks will provide a larger sample size for examining this sub-domain and will allow comparisons of non-Hispanic white vs. non-Hispanic black households.

Statistical power was calculated based on selected research questions and hypotheses that were guided by the FLASHE conceptual framework. The minimum sample size required to test the significance (alpha=0.05) of neighborhood characteristics in predicting food intake or physical activity in a linear regression was calculated. A 60% response rate was assumed based on prior work and results from the NCI 5 A Daysurvey (Erinosho et. al., 2012), which also collected data from a COP, and a design effect of 1.5 was estimated based on oversampling non-Hispanic Blacks.

FLASHE provides the opportunity for multi-level analyses (parent-adolescent dyads and examining the effects of the neighborhood environment on behavior). Because the intent of data collection is to develop a public use dataset, we have not identified all potential statistical analyses. Rather, we conducted power analyses to ensure future analyses could be powered to a multitude of potential research questions and analyses. The following describes the minimum sample size needed with at least 80% statistical power for testing different hypotheses given partial correlations between proposed dependent and independent variables.

Home environmental characteristics, family characteristics and fruit and vegetable intake and/or physical activity (partial correlation = 0.10-0.30): To achieve at least 80% statistical power for detecting a significant relationship between variables with an assumed 0.1 partial correlation, a minimum sample size of 1,949 would be required, after taking into account the 60% response rate and a design effect of 1.5. To achieve at least 80% statistical power for detecting a significant relationship between variables with an assumed partial correlation of 0.2, a minimum

sample size of 476 would be required after taking into account the 60% response rate and design effect of 1.5.

The examination of the association between neighborhood environment and fruit and vegetable intake or physical activity (partial correlation 0.05-0.1) is another area of analysis. To achieve at least 80% statistical power for detecting a significant relationship between variables with an assumed 0.065 partial correlation, a minimum sample size of 4,155 (2,500 complete dyads) would be required after taking into account the 60% response rate and a design effect of 1.5.

The minimum sample size required for analysis with sufficient statistical power depends on the presumed partial correlation between the dependent variable and independent variables being analyzed. The selected sample size of 2,500 completed dyads should provide sufficient power to support most of the selected potential hypotheses as described above and in Section A.

B.2 Procedures for the Collection of Information

There are three main phases of FLASHE data collection: enrollment, survey, and motion sensing study. Each of these steps are described below as well as illustrated in **Attachment M.**

Enrollment

The parent/caregiver in each dyad will be contacted by email and invited to participate in the FLASHE study. They will be provided a Personal Identification Number (PIN), and asked to log into the FLASHE website and enroll in the study. By assigning a PIN to each potential participant the study is able to link contact information to the PIN. This will allow the enrollment

form to be prepopulated with the contact information provided by Ipsos Interactive Services (Ipsos) who will be providing the sample.

As part of the enrollment, parents will be asked to confirm the contact information provided by Ipsos (**Attachment F**). In addition, they will be asked to provide an email address for the selected adolescent so that communications (assent forms, survey invitations, etc.) can be sent directly to the adolescent. Parents will be asked to complete an online consent form for both him/herself and a permission form for his/her selected adolescent **(Attachment G-1).** Once the parent has provided permission, the adolescent will be contacted by email, provided with his/her own PIN, invited to participate in FLASHE, and asked to completed an assent for **(Attachment G-2)**. Participants that do not enroll and/or do not provide consent/permission will be sent reminders (up to three times). FAQs will be available on the site. Email correspondence related to enrollment, screenshots of the FLASHE website, and the FAQs, can be found in **Attachment N**.

Web Survey

Once fully consented/permitted/assented into the study, study participants (both members of each dyad) will be contacted by email and each asked to complete two on-line surveys. Email correspondence related to the web survey can be found in **Attachment O**. There will be a total of four survey instruments, each with about 120 items:

- Adult Diet Survey;
- Adolescent Diet Survey;
- Adult Physical Activity Survey; and
- Adolescent Physical Activity Survey.

The survey instruments can be found in **Attachments H, I, J and K**. All survey data collection will be conducted via the Internet. Dyads will be randomized to determine which survey (diet or physical activity) they will receive first. The first screen of the first survey for the adolescent will include an assent form. The second survey will not become available until both members of the dyad have completed the first one. The demographics section will be included in whichever survey is provided first. For the purposes of this submission, the demographics have been included with the Physical Activity surveys.

Email or text message reminders will be sent to nonrespondents (every two weeks, up to three times). It is anticipated that the two surveys will be administered no more than 120 days apart.

Motion Sensing Study

A subsample of 1,500 adolescents will be selected for participation in the motion sensing study. This subsample will be taken only from adolescents:

- That have been provided with permission by the parent to participate in the motion sensing study; and
- That have provided assent for their own participation in the motion sensing study.

The parents will not be invited to participate in the motion sensing study. Correspondence related to the motion sensing study can be found in **Attachment P.**

Selected adolescents will be sent a package containing:

- an accelerometer and wrist strap; a cover letter with instructions for use of the accelerometer;
- a wear log; and
- a padded return envelope.

The accelerometer will be fully charged and initialized prior to the mailing. The adolescent will be asked to wear the device for seven consecutive days beginning on a pre-assigned day and to fill out a log daily. The Wear Log instrument can be found in **Attachment L**. After the seven-day 'wear' period, participants will be asked to return the accelerometer and wear log to Westat in the pre-paid padded envelope provided. The participant will receive a reminder email or text message. If the device is not returned within one week of the wear period, the adolescent will receive additional reminders. Up to three additional email reminders will be sent at 2 week intervals to request the return of the device.

As described in Statement A.9, an experiment related to the tokens of appreciation will be conducted in conjunction with the Motion Sensing Study. Participants will be randomized to receive either $20 or $40 if they return the motion sensing device. The proportion of respondents in each treatment group has not yet been finalized and will be budget-dependent. The token of appreciation will be sent within 5 days of the return of the device to reduce the risk of device loss. It is anticipated that 900 adolescents will complete the motion sensing study.

B.3 Methods to Maximize Response Rates and Deal with Nonresponse

In addition to sampling errors, all surveys are subject to other sources of errors, including non-coverage, non-response, and measurement errors. It is anticipated that use of the COP will reduce nonresponse because respondents have voluntarily joined the panel and agreed to participate in web surveys. Also, unlike other web surveys, a working email address will be provided by Ipsos, thereby reducing loss to contact. In addition email or text message reminders will be sent to nonrespondents throughout the study (see **Attachment M).**

As discussed earlier, the estimated response rate for this study is 60% based on prior work using a COP. This response rate is relatively high compared to other large national health surveys, which have documented lower and declining response rates. According to the Behavioral Risk Factor Surveillance System 2011 Summary Data Quality Report (2013) from CDC, the response rates for six major health related surveys range from 11.1% to 61.7% for landline telephone surveys. These surveys include California Health Interview Survey (CHIS), The Commonwealth Fund 2010 Biennial Health Insurance Survey, National Immunization Survey, Pew Research Center for People and the Press, Pew Internet and American Life Project6 2012, Behavioral Risk Factor Surveillance System.

Base analysis weights will be calculated that reflect the proportion of the sample in each of the balancing cells used to select the sample. These weights will be the ratio of known population counts, as derived from the most current available data (e.g., the American Community Survey), to the number of sampled cases in cells defined by demographic variables used in sampling. These base analysis weights will then serve as the basis for making nonresponse weighting adjustments. Further adjustments will also be applied to calibrate the nonresponse-adjusted weights to population totals using post stratification techniques such as ratio raking. The methodology for nonresponse and post-stratification adjustments are standard in survey research (e.g., see Potter et al 2004, Battaglia et al 2004, Johnson and McMahon 2002, Mantel et al 2000, Goksel et al 1992, Ezzati and Khare 1992, and Little 1986). Nonresponse adjustment is used to reduce bias related to survey nonresponse and post-stratification adjustments are used to assure that sums of weights for the sample respondents for key demographic variables match known population totals.

B.4 Test of Procedures or Methods to be Undertaken

Cognitive Testing

Cognitive testing was an important component of the instrument development process for FLASHE since some of the questions are newly developed and/or not previously used on any survey, not previously tested, or not previously-used with an adolescent population. This testing (conducted under OMB number 0925-0589-04, Approved 7/5/2012) was essential in identifying problems in question wording, context or order effects, as well as response difficulties resulting from the design and layout of the survey.

Two rounds of cognitive testing were conducted in 2012. The instruments were tested with adult and adolescent respondents to help identify and remove potential causes of response error. In each round, the surveys were tested with adolescent respondents (aged between 12 and 17) and the parent surveys were tested with a matching set of adult respondents who were parents of those adolescents. Significant effort was made to recruit diverse participants with regard to age, gender, educational background, and level of regular physical activity. Ten adolescent and parent pairs were tested in the first round and eight pairs were tested in the second round.

The two rounds of testing resulted in a large volume of useful information which was used to revise the instruments with regard to topics such as the overall length of the instruments, the formatting of certain question types (especially the matrices), respondents' ability to work with the question stems and apply them to multiple sub-items, and respondents' ability to recall activities in different time periods (i.e., times per day or times per week). The instruments submitted with this package are the final result of this testing.

Embedded Experiment

As discussed in section A.9, FLASHE intends to conduct an experiment on the tokens of appreciation related to the Motion Sensing Study. It is unclear what level of token is appropriate or needed to ensure the adolescents return the accelerometers to the study. FLASHE will randomly assign adolescent participants to be offered either $20 or $40 for their participation in the Motion Sensing Study. It is anticipated that the results of this experiment will inform future NCI studies that include the use of accelerometers being sent to respondents via mail.

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

Representatives from both contracting organizations (Westat and Ipsos) were consulted for their expertise in survey design, measurement, and administration. **Attachment B** has a list of additional individuals that were critical in developing the survey instrument, sampling strategies, and research plan in this study. Many in this national collaborative group may be consulted for the analysis phase, once the data are collected, due to their experience with the measures and their interest and expertise in survey data and behavioral health.