

**SUPPORTING STATEMENT FOR THE  
NATIONAL ADULT TOBACCO SURVEY**

**0920-0828  
Expiration 10/31/2010**

**Reinstatement with Changes**

**PART B**

**Submitted by:**

Sean Hu, MD, MS, DrPH  
Senior Epidemiologist/Project Officer  
Office on Smoking and Health - Epidemiology Branch  
National Center for Chronic Disease Prevention (NCCDPHP)  
Centers for Disease Control and Prevention (CDC)  
4770 Buford Highway, MS K-50  
Atlanta, GA 30341  
Email: [shu@cdc.gov](mailto:shu@cdc.gov)  
Phone: 770-488-5845  
Fax: 770-488-5848

**Centers for Disease Control and Prevention  
Department of Health and Human Services**

**Revised: April 26, 2012**

## Table of Contents

B. Collections Of Information Employing Statistical Methods.....	1
B.1 Respondent Universe And Sampling Methods.....	1
B.2 Procedures For Collection Of Information.....	2
B.3 Methods To Maximize Response Rates And Deal With Nonresponse.....	11
B.4 Tests Of Procedures Or Methods To Be Undertaken.....	14
B.5 Individuals Consulted On Statistical Aspects And Individuals Collecting And/Or Analyzing Data.....	15
REFERENCES .....	17

## **LIST OF APPENDICES**

- A. Authorizing Legislation
  - A-1. The Family Smoking Prevention and Tobacco Control Act
  - A-2. Public Health Service Act
- B. Logic Model as Framework for FDA’s Data Needs
- C. 60-Day Federal Register Notice
- D. Summary of Public Comments Received during the 60-Day Public Comment Period
- E. Item-Level Justification Related to FDA’s Mission in Regulating Tobacco Products
- F. Item-Level Justification Related Supplemental Uses of NATS Data
- G. Consultants on the Development of the NATS
- H. National Adult Tobacco Survey Questionnaire (English)
  - H-1. Screener for Landline Users
  - H-2. Screener for Cell Phone Users
  - H-3. Main Questionnaire
  - H-4. Cover and Table of Contents
- I. National Adult Tobacco Survey Questionnaire (Spanish)
  - I-1. Screener for Landline Users
  - I-2. Screener for Cell Phone Users
  - I-3. Main Questionnaire
  - H-4. Cover and Table of Contents
- J. Non-Disclosure Agreement Signed by Interviewers
- K. Sample Table Shells
- L. Advance “Dear Resident” Letter

## **B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS**

The NATS will develop national estimates of tobacco use attitude and behaviors and of exposure to pro- and anti-tobacco influences among non-institutionalized adults residing in the United States. There will be sufficient sample sizes to provide national estimates for many subpopulations, including domains defined by gender, race/ethnicity, and age group, and subdomains defined by smoking behavior.

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

The universe for the study will consist of non-institutionalized adults (age 18 and over) residing in the 50 States and the District of Columbia (DC). Interviews will not be conducted with adults who live in group quarters. Respondents will be selected through Random Digit Dialing (RDD) from two sampling frames: one for landlines and one for cell phones. The cell-phone frame will only be used to find households which are cell-phone only, i.e., ones that rely exclusively on cell phones because they do not have a landline that they use for receiving calls<sup>1</sup>.

As indicated in Part A, the sample supplier is a private company, Marketing Systems Group (MSG), which maintains the Genesys database of landline and cell telephone numbers, with landline numbers cross-referenced when feasible to addresses (approximately 60% of the time). . MSG supplies samples for thousands of research studies, including a variety of national studies conducted by CDC or conducted by states for their own purposes using protocols developed by CDC.

A probability sampling proportional to state population size will be used for achieving good estimates in national level. A Table later in Part 5 describes the sampling frame of respondents by demonstrating that the sample will be stratified by state to ensure that the sample by state is proportionate to the nationwide population. The data collection will be conducted using Computer Assisted Telephone Interviewing (CATI).

It is widely known that response rates for CATI surveys gradually have declined over the past 20 years due to increased use of caller ID and wariness of presumed telemarketers on the part of the public. While OMB sets a high bar for anticipated response rates, it also has been understanding of the continued value of CATI surveys as the most efficient, and often the only realistic means of conducting large-scale surveys. OMB generally has recognized that realistic expectations need to be set for CATI surveys. For the proposed NATS, our expectations regarding response rates are based largely on the experience of the 2009/2010 NATS.

In describing the response rates for the prior and upcoming cycles of NATS, we consider both the CASRO response rate and the CASRO cooperation rate. The CASRO response rate assumes that the proportion of eligible cases in the cases with unknown eligibility is equivalent to the proportion of eligible cases in the sum of cases in the sample of which the eligibility or ineligibility could be determined. The CASRO response rate in equation form is

---

<sup>1</sup> Landline numbers that are only used for modems, business purposes, etc., will not be counted in deciding whether households are cellphone only.

$$\frac{\text{(Number of completes)}}{\left( \frac{\text{Number of telephone numbers definitely reaching an eligible respondent}}{\text{Number of telephone numbers definitely reaching an eligible respondent}} + \frac{\text{Number of telephone numbers not definitively identified as reaching or not reaching an eligible respondent}}{\text{Number of telephone numbers not definitively identified as reaching or not reaching an eligible respondent}} \right) \times \left( \frac{\text{Number of telephone numbers definitely reaching an eligible respondent}}{\text{Number of telephone numbers definitely reaching or not reaching an eligible respondent}} \right)}$$

The CASRO cooperation rate is a proportion with the number of completed interviews in the numerator and the number of selected respondents who spoke with an interviewer in the denominator.

On the 2009/2010 NATS, the CASRO response rate was 37.6% and the CASRO Cooperation rate was 62.3%. For landlines, the response rate was 40.4% compared to 24.9% for cell phones. Conversely, for landlines, the cooperation rate was 61.9% but for cell phones the cooperation rate was 68.7%. For the upcoming cycles of NATS, we anticipate maintaining or improving on the rates achieved on the 2009/2010 NATS based on improvements in the methods for maximizing response, discussed below in B.3.

## **B.2 PROCEDURES FOR COLLECTION OF INFORMATION**

### **B.2.a Statistical Methodology for Stratification and Sample Selection**

*Landline Samples.* The NATS landline sample will consist of a list-assisted RDD sample of telephone numbers. To build the list-assisted frame, all possible telephone numbers are divided into blocks (or banks) of 100 numbers (e.g., 617-492-1200 to 617-492-1299). A 100-block is the series of 100 phone numbers defined by the last two digits of a 10-digit phone number. For phone numbers with the first eight digits in common, there are 100 possible combinations of the last two digits (ranging from 00-99). To enhance efficiency and reduce costs, the frame excludes zero-blocks, i.e., those 100-blocks with zero listed phone numbers.

Telephone numbers will be stratified into state-based strata according to the primary state served by the (area code and prefix). Within each state, telephone numbers will be further stratified into the *high-density* substratum or the *low-density* substratum based on whether the number is listed in local residential telephone directories or not. Telephone numbers listed in residential directories are most often working residential numbers, whereas unlisted telephone numbers include large numbers of non-working and nonresidential telephone numbers. To leverage this information, the high-density stratum will be oversampled at a 1.5-to-1 ratio relative to the low-density stratum. This oversampling increases the sampling efficiency by raising the percentage of working residential numbers selected in the sample. The sample will be selected in independent replicates to facilitate the control of the final number of completed interviews.

*Cell Phone Sample.* The cell phone sample will be an RDD sample of phone numbers from cell phone and cell/landline exchanges. The exchanges originate from the Telecordia® TPM™ Data Source. The cell phone exchanges and mixed-use exchanges are identified from exchange type. The NATS cell phone sample will be stratified explicitly by state to help control the geographic distribution of the sample.

## B.2.b Estimation and Justification of Sample Size

This section provides justification for the sample sizes. Our focus is on national estimates which can provide sufficient power for a wide range of population domains within the US adult population. These domains are defined in terms of gender, race/ethnicity, and age group, as well as subdomains by tobacco use behaviors within these basic demographic domains. We wish to provide estimates of population values within these domains and subdomains with a high degree of precision.

Table B-1 presents standard errors for a range of domain sizes and a range of population-level questionnaire percentages. The precision is presented in terms of standard error of estimated prevalence rates (percentages or proportions). The standard errors assume a binomial distribution with probability  $p$  (equal to the questionnaire percentage), a sample size equal to 75,000 multiplied by the domain proportion, and a design effect (DEFF)<sup>2</sup> of 2.0<sup>3</sup>. In Table B-1 the standard errors have stars associated with them as follows:

- \*\*\*: Coefficient of variation less than 10% (standard error less than 10 percent of the mean)—Excellent precision<sup>4</sup>;
- \*\*: Coefficient of variation between 10% and 20% (standard error between 10 percent and 20 percent of mean)—Very good precision<sup>5</sup>;
- \*: Coefficient of variation between 20% and 30% (standard error between 20 percent and 30 percent of mean)—Good precision<sup>6</sup>;
- Coefficient of variation greater than 30% (standard error greater than 30 percent of mean)—Fair precision;

The confidence intervals are two-sided 95 percent confidence intervals around sample percentages equal to the question percentage<sup>7</sup>.

Table B-1 95% Confidence Intervals for Selected Domain Sizes.

<sup>2</sup> The design effect (DEFF) is defined as the actual sampling variance divided by the variance that would be attained for a simple random sample of the same size. The DEFF, which equal 1.0 for simple random sampling, is a measure of the extra variability induced by complex sampling designs. The sources of this design effect will be the difference in the landline and cellphone sampling fractions (the cellphone sampling fraction being smaller due to the greater cost of these per completed interview), differential household weights and nonresponse and coverage adjustments, and the effect of having a minimum sample size of 1,000 for small states.

<sup>3</sup> The standard error is  $\sqrt{2 * p * (1 - p) / (75,000 * d)}$ , where  $p$  is the question percentage and  $d$  is the domain percentage.

<sup>4</sup> The 95% confidence intervals are tighter than the mean plus-or-minus 20%.

<sup>5</sup> The 95% confidence intervals are between the mean plus-or-minus 20% and the mean plus-or-minus 40%.

<sup>6</sup> The 95% confidence intervals are between the mean plus-or-minus 40% and the mean plus-or-minus 60%.

<sup>7</sup> The question percentage  $p$  plus or minus 1.96 times the standard error.

Question percentage	Std error 50% domain	95% Confidence Interval for 50% domain	Std error 25% domain	95% Confidence Interval for 25% domain	Std error 10% domain	95% Confidence Interval for 10% domain
50%	0.37%** *	(49.3%, 50.7%)	0.52%** *	(49.0%, 51.0%)	0.82%** *	(48.4%, 51.6%)
25%	0.32%** *	(24.4%, 25.6%)	0.45%** *	(24.1%, 25.9%)	0.71%** *	(23.6%, 26.4%)
10%	0.22%** *	(9.6%, 10.4%)	0.31%** *	(9.4%, 10.6%)	0.49%** *	(9.0%, 11.0%)
5%	0.16%** *	(4.7%, 5.3%)	0.23%** *	(4.6%, 5.4%)	0.36%** *	(4.3%, 5.7%)
2%	0.10%** *	(1.8%, 2.2%)	0.14%** *	(1.7%, 2.3%)	0.23%**	(1.6%, 2.4%)
1%	0.07%** *	(0.9%, 1.1%)	0.10%**	(0.8%, 1.2%)	0.16%**	(0.7%, 1.3%)
Question percentage	Std error 5% domain	95% Confidence Interval for 5% domain	Std error 2% domain	95% Confidence Interval for 2% domain	Std error 1% domain	95% Confidence Interval for 1% domain
50%	1.15%** *	(47.7%, 52.3%)	1.83%** *	(46.4%, 53.6%)	2.58%** *	(44.9%, 55.1%)
25%	1.00%** *	(23.0%, 27.0%)	1.58%** *	(21.9%, 28.1%)	2.24%** *	(20.6%, 29.4%)
10%	0.69%** *	(8.6%, 11.4%)	1.10%**	(7.9%, 12.1%)	1.55%**	(7.0%, 13.0%)
5%	0.50%**	(4.0%, 6.0%)	0.80%**	(3.4%, 6.6%)	1.13%*	(2.8%, 7.2%)
2%	0.32%**	(1.4%, 2.6%)	0.51%*	(1.0%, 3.0%)	0.72%	(0.6%, 3.4%)
1%	0.23%*	(0.5%, 1.5%)	0.36%	(0.3%, 1.7%)	0.51%	(0.0%, 2.0%)

Percentages for persons answering particular questionnaire items can vary through all possible percentages (1% through 99%). Percentages  $p$  from 50% to 100% have standard errors equal to those of the complement percentage  $1 - p$ , so only those  $p$  values 50% and below are presented.

We define ‘good precision’, as a standard error less than 30% of the mean value (corresponding to 95% confidence intervals which are at least the mean plus or minus 60% of the mean). “Good precision” can be achieved either with the current sample size of 75,000 and a). domains of size 2% with questionnaire percentages between 2% and 98%, or b). domains of size 1% with questionnaire percentages between 5% and 95%. Similarly, ‘very good precision’ could be achieved with questionnaire percentage ranges of (2%, 98%) for 5% domains; questionnaire percentage ranges of (5%, 95%) for 2% domains, and questionnaire percentages ranges of (10%, 90%) for 1% domains. There are many domains of interest, which are in these ‘cutoff’ ranges (current tobacco users are roughly 20%-25% of the adults), and breaking them further into

subdomains will result in many domains of size 2% (10% of current smokers), and domains of size 1% (5% of current smokers).

Examples of domains which are roughly 10% of the adult population (and will also be roughly 10% of the sample) are as follows:

- Blacks (all adults) or Hispanics (all adults);
- Hispanics (all adults);
- Males age 45 to 54;
- Male current smokers<sup>8</sup>.
- Current smokers who have tried to quit this year<sup>9</sup>;

Examples of domains which are roughly 5% of the adult population (and will also be roughly 5% of the sample) are as follows:

- Black males or Black females;
- Current smokeless tobacco users;
- Cigar/cigarillo use;
- Smokeless use (age 18-24);

Examples of domains which are roughly 2% or 1% of the adult population (and will also be roughly 2% or 1% of the sample respectively) are as follows:

- Black current smokers or Hispanic current smokers (2% each);
- Males age 45 to 54 current smokers (2%);
- Smokeless use among Blacks (1%);
- Asian males (2%);
- Black males 35 to 44 years old (1%);
- Black male current smokers (1%).

Tables B-2 through B-4 present domain percentages for the full range of gender, race/ethnicity, and age group adult population subgroups. Current tobacco users as a subgroup of these should be roughly 1 in 5 in most cases. These domain percentages are from population estimates for the United States adult population from the 2010 American Community Survey (US Bureau of the Census).

We plan to modify the purely proportional design to assign a minimum sample size to the smaller states. States which receive an expected sample size less than 1,000 have their sample sizes reset up to 1,000, with remaining sample sizes then allocated proportionally. Table B-5 presents a proposed breakdown for an overall sample size of 75,000, using 2010 population figures to define the allocation. Larger states are assigned sample sizes completely proportionally to their population.

Effective sample size, defined as  $n/DEFF$  where  $n$  is the actual sample size (number of respondents), is the size of a simple random sample with equivalent precision. The effective sample size from the Table B-5 sample design for a national estimate is 63,816. This is a design

---

<sup>8</sup> Current smokers are roughly 20% of the adult population.

<sup>9</sup> Roughly one-half of current smokers try to quit during a given year.



effect of 1.175. The design effect from assigning these minimal state sample sizes is subsumed into the overall assumed design effect of 2 in Table B-1 (i.e., in the variance calculations in Table B-1 we assume that design effects from all effects together will be 2.0). The targets in Table B-5 would be allocated to cellphone and landline using cellphone-only allocations for each state. The allocations to the two strata would be based on the percentages of households that are cellphone-only and non cellphone-only, and also on the relative costs of completing the survey. The final allocations would be proportional to the percentage divided by the square root of the cost ratio, in a way that maximizes precision for a fixed cost.

Table B-2. Population Percentages for Gender by Age Domains

Domain	2010 ACS Population Estimate	% total adult U.S. pop
Total U.S. adult pop	235,184,324	100.00%
Male U.S. adult pop	114,105,885	48.52%
18 and 24 years	15,801,040	6.72%
25 to 34 years	20,553,204	8.74%
35 to 44 years	20,494,598	8.71%
45 to 54 years	22,100,095	9.40%
55 to 64 years	17,723,303	7.54%
65 and over	17,433,645	7.41%
Female U.S. adult pop	121,078,439	51.48%
18 and 24 years	15,094,351	6.42%
25 to 34 years	20,418,879	8.68%
35 to 44 years	20,697,730	8.80%
45 to 54 years	22,828,938	9.71%
55 to 64 years	19,038,661	8.10%
65 and over	22,999,880	9.78%

Table B-3. Population Percentages for Race by Gender by Age Groups<sup>10</sup>

Domain	White		Black		Asian	
	Population estimate	Percent of adults	Population estimate	Percent of adults	Population estimate	Percent of adults
Total Adults	178,894,402	76.07%	28,120,054	11.96%	11,487,356	4.88%
Male Adults	87,276,798	37.11%	13,034,091	5.54%	5,350,710	2.28%
18 and 24 years	10,988,750	4.67%	2,288,916	0.97%	775,000	0.33%

<sup>10</sup> Each race group includes Hispanics and nonHispanics. Multi-race persons (who claimed multiple races on the ACS questionnaire) are not included in any of the three groups.

25 to 34 years	14,650,556	6.23%	2,528,345	1.08%	1,154,561	0.49%
35 to 44 years	15,012,749	6.38%	2,517,408	1.07%	1,158,303	0.49%
45 to 54 years	17,199,909	7.31%	2,574,973	1.09%	960,092	0.41%
55 to 64 years	14,451,021	6.14%	1,778,983	0.76%	695,799	0.30%
65 years plus	14,973,813	6.37%	1,345,466	0.57%	606,955	0.26%
Female Adults	91,617,604	38.96%	15,085,963	6.41%	6,136,646	2.61%
18 and 24 years	10,468,450	4.45%	2,300,492	0.98%	747,952	0.32%
25 to 34 years	14,277,929	6.07%	2,815,204	1.20%	1,311,659	0.56%
35 to 44 years	14,836,493	6.31%	2,816,252	1.20%	1,312,838	0.56%
45 to 54 years	17,426,733	7.41%	2,910,818	1.24%	1,104,004	0.47%
55 to 64 years	15,186,153	6.46%	2,147,641	0.91%	854,084	0.36%
65 years plus	19,421,846	8.26%	2,095,556	0.89%	806,109	0.34%

Table B-4. Population Percentages by Gender by Age Groups for Hispanics.

Hispanic domain	Population estimate Males	Percent of adults	Population estimate Females	Percent of adults
All Adults	16,937,454	7.20%	16,604,976	7.06%
18 and 24 years	3,300,862	1.40%	2,927,572	1.24%
25 to 34 years	4,396,484	1.87%	4,018,027	1.71%
35 to 44 years	3,757,039	1.60%	3,602,355	1.53%
45 to 54 years	2,751,200	1.17%	2,742,437	1.17%
55 to 64 years	1,544,067	0.66%	1,694,048	0.72%
65 years and over	1,187,802	0.51%	1,620,537	0.69%

### B.2.c. Estimation and Statistical Testing Procedures

Weights will be computed following the general formulation:

$$\text{FINALWT} = \text{BASEWT} \times (1/\text{NPH}) \times \text{NAD} \times \text{POSTSTRAT}$$

FINALWT is the final weight assigned to each respondent.

BASEWT accounts for differences in the basic probability of selection among strata (high density vs. low-density, landline vs. cellphone). It is the inverse of the sampling fraction of each stratum.

1/NPH is the inverse of the number of residential telephone numbers in the respondent's household.

NAD is the number of adults in the respondent's household.

POSTSTRAT is the number of people in an age-by-sex or age-by-race/ethnicity-by-sex category divided by the sum of the preceding weights for the respondents in the same age-by-sex or age-by-race/ethnicity-by-sex category. It adjusts for noncoverage and nonresponse and forces the sum of the weighted frequencies to equal population estimates for these subgroups. This will effectively adjust for differential nonresponse between these subgroups.

Variance estimates can also be directly computed for the national sample estimates using any of the software available for survey data analyses; e.g. SUDAAN or SAS survey procedures (such as Survey Means.)

#### B.2.d. Use of Less Frequent Than Annual Data Collection to Reduce Burden

Clearance is being sought for the annual administration of NATS. Annual data collection is needed to support FDA's efforts to monitor, evaluate, and adjust as needed the implementation of its regulatory activities related to tobacco use, described in Part A of the supporting justification. Given the speed at which changes can be experienced and the need for rapid feedback and adjustments in implementation methods, less frequent than annual data collection would sharply impair FDA's effectiveness.

**Table B-5. Sample Allocation by State**

State Name	7-1- 2010 Pop	Pop Pct	Total Target	Pct of sample
Total	309,050,816	100%	75,000	100%
Alabama	4,729,656	1.530%	1,000	1.33%
Alaska	708,862	0.229%	1,000	1.33%
Arizona	6,676,627	2.160%	1,255	1.67%
Arkansas	2,910,236	0.942%	1,000	1.33%
California	37,266,600	12.058%	7,005	9.34%
Colorado	5,095,309	1.649%	1,000	1.33%
Connecticut	3,526,937	1.141%	1,000	1.33%
Delaware	891,464	0.288%	1,000	1.33%
District of Columbia	610,589	0.198%	1,000	1.33%
Florida	18,678,049	6.044%	3,511	4.68%
Georgia	9,908,357	3.206%	1,862	2.48%
Hawaii	1,300,086	0.421%	1,000	1.33%
Idaho	1,559,796	0.505%	1,000	1.33%
Illinois	12,944,410	4.188%	2,433	3.24%
Indiana	6,445,295	2.086%	1,212	1.62%
Iowa	3,023,081	0.978%	1,000	1.33%
Kansas	2,841,121	0.919%	1,000	1.33%
Kentucky	4,339,435	1.404%	1,000	1.33%
Louisiana	4,529,426	1.466%	1,000	1.33%
Maine	1,312,939	0.425%	1,000	1.33%
Maryland	5,737,274	1.856%	1,078	1.44%
Massachusetts	6,631,280	2.146%	1,247	1.66%
Michigan	9,931,235	3.213%	1,867	2.49%
Minnesota	5,290,447	1.712%	1,000	1.33%
Mississippi	2,960,467	0.958%	1,000	1.33%
Missouri	6,011,741	1.945%	1,130	1.51%
Montana	980,152	0.317%	1,000	1.33%
Nebraska	1,811,072	0.586%	1,000	1.33%
Nevada	2,654,751	0.859%	1,000	1.33%
New Hampshire	1,323,531	0.428%	1,000	1.33%
New Jersey	8,732,811	2.826%	1,642	2.19%
New Mexico	2,033,875	0.658%	1,000	1.33%
New York	19,577,730	6.335%	3,680	4.91%
North Carolina	9,458,888	3.061%	1,778	2.37%
North Dakota	653,778	0.212%	1,000	1.33%
Ohio	11,532,111	3.731%	2,167	2.89%
Oklahoma	3,724,447	1.205%	1,000	1.33%
Oregon	3,855,536	1.248%	1,000	1.33%
Pennsylvania	12,632,780	4.088%	2,375	3.17%
Rhode Island	1,056,870	0.342%	1,000	1.33%
South Carolina	4,596,958	1.487%	1,000	1.33%
South Dakota	820,077	0.265%	1,000	1.33%
Tennessee	6,338,112	2.051%	1,191	1.59%
Texas	25,213,445	8.158%	4,739	6.32%
Utah	2,830,753	0.916%	1,000	1.33%
Vermont	622,433	0.201%	1,000	1.33%
Virginia	7,952,119	2.573%	1,495	1.99%
Washington	6,746,199	2.183%	1,268	1.69%
West Virginia	1,825,513	0.591%	1,000	1.33%

Wisconsin	5,668,519	1.834%	1,065	1.42%
Wyoming	547,637	0.177%	1,000	1.33%

### B.2.e Survey Instrument

The NATS questionnaire (Appendix H/I) contains 141 items, including landline and cell phone screening questions. The NATS comprehensively assesses use of many tobacco products. In addition, it contains questions regarding demographics. Exhibit B-7 outlines the questionnaire topics and the number of questions in each topic area. All but two of the questions are in a multiple-choice format. The other two questions ask for open-ended responses to a time interval.

### EXHIBIT B-7 QUESTIONNAIRE TOPICS AND TOTAL NUMBER OF QUESTIONS PER TOPIC

Section	Number of Questions
Introductory Questions	2
Cigarette Smoking	15
Susceptibility—Cigarettes	3
Other Tobacco Products	40
Addiction	7
Cessation	8
Marketing/Public Education	11
Purchasing	6
Demographics	12
Cell Phone Usage (landline survey only)	4
Knowledge/Attitudes/Perceptions	14
Secondhand Smoke	2
Closing Questions	2
Landline or Cell Phone Screener	8 or 7
Total	134 for landline or 129 for cell phone

### B.2.f Data Collection Procedures

The data collection procedures for the NATS are made up of several components, the application of which varies between landline vs. cell phone surveys. The components include: (1) advanced mailings; (2) loading the sample; (3) managing call attempts; (4) conducting the interview; (5) handling busy and no-answer; (6) attempting call backs; (7) managing refusals and interrupted interviews and; (8) recording call dispositions.

*Advanced mailings:* Advance letters (Appendix L) addressed to “Dear Resident” will be sent to all sampled households for which addresses can be obtained. All envelopes will be

stamped with first class postage, sealed, and have the official CDC logo. Because cell phone numbers cannot be reverse-matched for addresses, cell phone sample will not receive advance letters

*Loading the Sample:* CDC will provide Westat with 12 sets of monthly sample files, each of which they will complete in approximately six weeks. Sample records will have been pre-screened to identify business and non-working numbers and these records will not be called..

*Managing Call Attempts:* Call attempts will occur in respondent time, from 9:00 a.m. to 9:00 p.m. Monday through Friday, 10:00 a.m. to 6:00 p.m. on Saturday, and 2:00 p.m. to 9:00 p.m. on Sunday. For landline sample records, each non-final record will receive a minimum of 15 attempts. Non-final cell phone sample records will be limited to 12 attempts because random-respondent selection is not conducted on cell phone sample and more interviews are completed on the first contact.

Until contact is made, call attempts will occur across a variety of “timeslices” (day of the week and time period combinations) to maximize the likelihood of contact. Sample records that encounter a contact but do not result in a final result (requested call-backs, etc.) will be called according to the time requested by the respondent.

*Conducting the Interview:* A screener will be conducted at the beginning of each call. The screener consists of (1) verification of private residence; and (3) random respondent selection. The screener will be modified for cell phone respondents to ensure we do not: (1) jeopardize safety (e.g., respondent is driving); (2) make duplicate calls (e.g., respondent has a land line that could be in the landline sample); (3) interview someone who is underage.

*Dealing with Busy and No-Answer:* Lines that are busy will be called back 4 times at 15-minute intervals. If the line is still busy after the fourth attempt, the number will be attempted again on different calling occasions until the record is resolved. Calls that ring without being answered or connected will be allowed to ring 6 times (as heard by the interviewer) before hanging up and being attempted again on the next calling occasion.

*Attempting Call-backs:* The CATI scheduler used for NATS automatically assigns scheduled appointment cases as the highest priority and scheduled call-backs are always met when there is an available interviewer (and called as soon as possible after the scheduled time if all interviewers are busy at the exact appointment time.. If the respondent is not available at the scheduled time, one additional attempt is made after 20 minutes before the scheduler treats the case as a missed appointment. This improves the likelihood of reaching a respondent if, for example, they were late getting home from work. The effective management of call-backs will increase the response rate and population coverage. Perhaps more importantly, scheduling an appointment that is convenient for the respondent, and ensuring that the appointment is kept, offers a basic courtesy to someone who has agreed to assist us with a study.

*Managing Interrupted Interviews:* Interrupted interviews with receptive respondents will be restarted using a definite call-back strategy. A definite call-back for an exact time can be set and the interview can begin where it left off. If the interviewer who began the survey is available at the prescribed time, the system will send the call back to that station.

*Recording Call Dispositions:* Dispositions of each call attempt on all records in the

sample will be automatically stored in the CATI system. This provides a complete call history for each record in the sample. The call history is displayed on the interviewer's screen during each new attempt.

#### B.2.g Informed Consent

Before each interview, the interviewer will read the scripted informed consent text to each participant. The consent form describes the interview, the types of questions that will be asked on the actual survey, the risks and benefits of participation, and participants' rights, and provides information on whom to contact with questions about any aspect of the study. The consent form also indicates that participation is completely voluntary and that participants can refuse to answer any question or discontinue the interview at any time. The interviewer will enter a code via the keyboard to signify that the participant was read the informed consent script and agreed to participate.

#### B.2.h Quality Control

The following table identifies and describes the major means of quality control. As shown, the task of collecting quality data begins with clear and explicit testing of CATI programs, continues with thorough initial and follow-up training and ongoing monitoring of project interviewers and ends with procedures for the cleaning, coding, and verification of collected data.

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

Response rates are an important indicator of data quality. OMB generally regards studies with higher response rates as offering more representative data. At the same time, OMB has acknowledged repeatedly in its own guidance documents that the range of likely feasible response rates are largely a function of the objectives of a study and of the methodology required. OMB also sets no pre-determined minimum required response rate across surveys of all types, recognizing that some types of surveys, such as population-based CATI surveys, necessarily should be expected to achieve lower response rates than surveys involving many other data collection methods. Moreover, OMB has recognized that CATI survey response rates have been declining in recent years for a variety of reasons, but serve an important purpose and need to be included in the mix of methods used to gather population-based data.

#### B.3.a Maximizing Response Rates

As indicated above, CATI surveys of all types, including NATS, have experienced declining response rates over the past 20 years. There are a limited number of accepted best practices available to help offset this decline. The following best practices that will be used to maximize response rates in the NATS:

- Advance letters - when sampled landline phone numbers can be matched to an address
- Using highly trained interviewers (including bilinguals) with effective interviewing techniques
- Using a sample management approach that ensures a high number of contact attempts (15 for landline numbers and up to 12 for cell phone numbers)



- Well scheduled call attempts to maximize contact likelihood (with higher staffing levels during evenings and weekends when more respondents tend to be reachable)

### Quality Control Procedures

Survey Step	Quality Control Procedures
Testing of CATI program	<ul style="list-style-type: none"> <li>• Test each response to each question, and each path through the survey (100%)</li> <li>• Review frequencies from testing data to ensure that the program is organizing data properly and recording values according to the survey specification (100%)</li> <li>• Develop skip check program to check data against defined conditions specified in the provided questionnaire (100%)</li> <li>• Provide CDC with an electronic test version of the programmed survey (100%)</li> </ul>
CATI pretest	<ul style="list-style-type: none"> <li>• Pretest of 100 interviews to ensure the CATI program is working properly and to verify questionnaire content, skip patterns, value verification, consistency of answers across questions, interviewer and supervisor training, and sample management procedures</li> </ul>
Advance letters	<ul style="list-style-type: none"> <li>• Verify that envelopes are stamped with first class postage, sealed, and have the official CDC logo (5% sample)</li> </ul>
CATI quality assurance	<ul style="list-style-type: none"> <li>• Monitor at least 10% of all interviews (10% sample)</li> <li>• Monitor each interviewer at least once per week and provide coaching and feedback. (100%)</li> <li>• Review call center shift reports and internal project tracking reports daily (100%)</li> </ul>
Preparation of data files	<ul style="list-style-type: none"> <li>• Identify incomplete interviews and merge back into the main data file (100%)</li> <li>• Clean and, when applicable, back-code open ended responses (100%)</li> <li>• Assign a final disposition to each record (100%)</li> <li>• Produce frequency tabulations of every question and variable to detect missing data or errors in skip patterns (100%)</li> </ul>

- Reviewing and adjusting introductory scripts to maximize the effectiveness of the survey language (to minimize length, quickly gain respondent's attention, and provide respondents with a reason to participate)
- Providing respondents (via the advance letter and via telephone contact) with information to validate the legitimacy of the study (informational letters, web sites, email addresses, and/or toll-free numbers)
- Employing refusal conversion to re-contact respondents who initially refuse the interview

Attempting to complete surveys with people or households that initially refuse can be a sensitive issue. However, refusal conversion can significantly improve response rates with minimal risk of respondent frustration when done carefully. If a respondent refuses to participate

in the NATS, the interviewer will indicate the respondent's level of hostility. Cases that indicate a high level of hostility (e.g., one in which the respondent is abusive) are moved to a review queue for a supervisor to address and are usually not attempted for conversion. Cases that indicate no hostility are automatically moved to a queue for a designated number of days. The number of days for this "cool-off" period will be 13 days when possible, but may be reduced near the end of an interviewing period. Cases that receive a second refusal will immediately become final.B.3.b

### B.3.b Dealing with Nonresponse

Survey nonresponse bias occurs when respondents are substantively different from the nonrespondents. Response rates are often used as a measure of data quality because they are thought to reflect the degree to which non-response bias exists in the data, but this connection is tenuous.<sup>11,12</sup> Instead, response rates are a measure of the risk of nonresponse. High response rates reflect low risk of nonresponse bias while high response rates increase the risk of nonresponse. In the absence of high response rates, a nonresponse analysis helps to justify the accuracy of the survey data.

As a whole, the field of survey research has been experiencing declining response rates over recent years. Bias will be present in NATS if the nonrespondents are different from the respondents in terms of the statistics of interest. In 2008, ICF Macro conducted a nonresponse follow-up (NRFU) to the Maryland Adult Tobacco Survey (MATS) on behalf of the Maryland Department of Health. The justification for the research was to analyze if the nonrespondents were different than the respondents. The research concluded that the respondents and nonrespondents were different in terms of smoking statistics, but much of the difference was explained by demographic differences between the respondents and nonrespondents. In turn, the weighting algorithm which corrected for known demographic biases in RDD surveys corrected for the differences in smoking characteristics between the nonrespondents and respondents.<sup>13</sup>

For the NATS, a NRFU survey is not planned. Instead, we intend to evaluate the extent of nonresponse bias using external data sources. Many of these comparisons are naturally inherent in the process of poststratification and weighting for nonresponse and noncoverage. For the weighting process, the comparisons typically focus on age, sex, race, Hispanic origin, education status, and marital status within each state. The data for these comparisons will be based on the American Community Survey (ACS).

The landline sample records contain two variables that could be used to explicitly adjust for nonresponse—listed or not-listed status and a metropolitan status code. We will weight for differential non-response for the categories of each of these variables and determine their effect, individually and jointly, on the bias of selected demographic characteristics and on the estimates and variances of key substantive variables. A decision as to whether or not to make explicit nonresponse adjustments using listed or not-listed status and the metropolitan status code will be based on these analyses.

---

<sup>11</sup> Curtin, R., Presser, S., & Singer, E. (2000). The Effects of Response Rate Changes on the Index of Consumer Sentiment. *Public Opinion Quarterly*, 413-428.

<sup>12</sup> Groves, R. (2006). Non-response Rates and Non-response Bias in Household Surveys. *Public Opinion Quarterly*, 646-675

<sup>13</sup> Freedner, N. R. ZuWallack, J. Dayton, J. Ross. (2009) Effects of Nonresponse by Smokers in Lowering Adult Tobacco Survey vs. Behavioral Risk Factor Surveillance System Smoking Estimates. Presentation at the 64<sup>th</sup> Annual Conference of the American Association of Public Opinion Research (AAPOR), May 14-19, Hollywood, FL.

The NATS has a limited set of survey questions that overlap with other data sources including the National Health Interview Survey (NHIS) and the Current Population Survey Tobacco Use Survey (CPS-TUS). Both the NHIS and CPS-TUS are valuable in quantifying nonresponse, but both have limitations. The NHIS and CPS-TUS include computer assisted personal interviewing and achieve extremely high response rates. These surveys are less susceptible to bias due to nonresponse, but observed differences when comparing to NATS may be confounded with the mode of survey administration.

The CPS-TUS has overlapping content with NATS including smoking status, quit attempts and cessation, smoking in the home. Further, the CPS-TUS can be analyzed at the state level. However, CPS-TUS is only conducted every three years and focuses on secondhand smoke policies and cessation assistance, areas not related to FDA authorities. Observed differences may be confounded with trends in tobacco behaviors and attitudes. The NHIS is limited to smoking status and quit attempts.

Through the use of auxiliary variables and demographic and limited substantive comparisons with ACS, NHIS, and CPS-TUS, we will assess the risk of nonresponse bias in the NATS. Despite the stated limitations, these data sources provide valuable benchmarks for NATS. Substantial deviations from these benchmarks will be explored further to 1) better understand the nature of the differences (e.g. do they vary across subgroups); 2) evaluate whether the differences are caused by nonresponse (and/or noncoverage) error or if there are other reasons that could explain the differences; and 3) if necessary, develop additional weighting adjustments to mitigate the risk of nonresponse bias on NATS estimates. Ultimately the nonresponse analyses will inform the survey weighting and identify limitations in the data that will be communicated to stakeholders.

#### **B.4 TESTS OF PROCEDURES OR METHODS TO BE UNDERTAKEN**

The NATS questionnaire was developed originally in the summer of 2008 based on eight years of experience with the ATS by 25 States, with technical guidance from CDC. The ATS was significantly reconfigured to create the NATS. As part of this process, in accord with OMB guidelines, NATS questionnaire items were subjected to cognitive interviewing and analyses by the contractor in the Fall 2008 and Winter 2009. This cognitive analysis resulted in the revision, addition, or deletion of response options and the revision or deletion of certain questions, with the overall effect of improving the clarity of questions and lowering respondent burden.

The NATS questionnaire was revised extensively to meet the needs of FDA in the summer and fall 2011. The revised questionnaire was subjected to cognitive interviewing and analyses by the contractor in February and March 2012. The cognitive analyses again resulted in revision, addition, or deletion of response options and the revision or deletion of certain questions. Again, this process improved the clarity of questions and lowered respondent burden. Following cognitive interviewing, the revised questionnaire underwent a limited pretest in Prince George's County, Maryland in accord with OMB guidelines. The pretests sharpened the articulation of certain survey questions and confirmed the empirical estimate of the survey burden.

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

B.5.a Statistical Review

Statistical aspects of the study have been reviewed by the individuals listed below.

Sean Hu, Ph.D.  
Centers for Disease Control and Prevention (CDC)  
National Center for Chronic Disease Prevention and Health Promotion  
Office on Smoking and Health  
4770 Buford Highway Mailstop K-50  
Atlanta, Georgia 30341  
(770) 488-5845  
[fik4@cdc.gov](mailto:fik4@cdc.gov)

Lou Rizzo, Ph.D.  
Westat, Inc.  
1600 Research Blvd.  
Rockville, MD 20850-3195  
(301) 294-4486  
[RizzoL1@westat.com](mailto:RizzoL1@westat.com)

B.5.b Agency Responsibility

Within the agency, the following individual will be responsible for receiving and approving contract deliverables and will have primary responsibility for data analysis:

Sean Hu, Ph.D.  
Centers for Disease Control and Prevention (CDC)  
National Center for Chronic Disease Prevention and Health Promotion  
Office on Smoking and Health  
4770 Buford Highway Mailstop K-50  
Atlanta, Georgia 30341  
(770) 488-5845  
[fik4@cdc.gov](mailto:fik4@cdc.gov)

B.5.c Responsibility for Data Collection

The representative of the contractor responsible for conducting the planned data collection is:

James Ross, M.S.  
Westat, Inc.  
1600 Research Blvd.  
Rockville, MD 20850-3195  
(240) 453-5620  
[JamesRoss@westat.com](mailto:JamesRoss@westat.com)

## REFERENCES

Abreu, D.A., and Winters, F. (1999) "Using Monetary Incentives to Reduce Attrition in the Survey of Income and Program Participation." *Proceedings of the Survey Research Methods Section of the American Statistical Association*.

American Cancer Society. *Cigar Smoking*. Atlanta: American Cancer Society, 2010 [Available at <http://www.cancer.org/Cancer/CancerCauses/TobaccoCancer/CigarSmoking/index>]. Accessed 22 December, 2011

Burns, David M., "Cigar Smoking: Overview and Current State of the Science," *Smoking and Tobacco Control Monograph No. 9*. Available at [http://cancercontrol.cancer.gov/tcrb/monographs/9/m9\\_1.PDF](http://cancercontrol.cancer.gov/tcrb/monographs/9/m9_1.PDF). Accessed 22 December 2011

CDC (2006). *National Strategic Plan for Tobacco Control - FY2006-FY2008*. Atlanta, GA: CDC.

CDC (2007). *Best Practices for Comprehensive Tobacco Control Programs – 2007*. Atlanta, GA: U.S. Department of Health and Human Services, CDC, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

CDC (2008). Smoking-attributable mortality, years of potential life lost, and productivity losses – United States, 2000-2004. *MMWR*; 57(45):1226-1228.

CDC (2011) Vital Signs: Current Cigarette Smoking Among Adults Aged  $\geq$  18 Years. United States--2005-2010. *MMWR*; 60(35):1207-1212

CDC, NCHS (2008). Public use data file and documentation: multiple causes of death for ICD-10 2005 data [CD-ROM].

National Institute of Drug Abuse (2009). Topics in Brief: Smokeless Tobacco. Available at <http://drugabuse.gov/tib/smokeless.html> Accessed 21 December 2011

Ryan, H., Wortley, P.M., Easton, A., Pederson, L., Greenwood, G. (2001). Smoking among lesbians, gays, and bisexuals: A review of the literature. *American Journal of Preventative Medicine*; 21: 142-149.

Substance Abuse and Mental Health Services Administration (SAMHSA). *Results from the 2009 National Survey on Drug Use and Health: Detailed Tables*. Available at <http://www.oas.samhsa.gov/NSDUH/2k9NSDUH/tabs/cover.pdf> Accessed 22 December 2011

SAMHSA. 2008 National Survey on Drug Use and Health: National Findings. DHHS NSDUH, Office of Applied Studies. Available at <http://www.oas.samhsa.gov/nsduh/2k8nsduh/2k8results.pdf> Accessed 22 December 2011  
Starr, G, T Rogers, M Schooley, S Porter, E Wiesen & N Jamison (2005). *Key Outcome Indicators for Evaluating Comprehensive Tobacco Control Programs*. Atlanta, GA: CDC.

US Department of Agriculture (USDA). Expenditures for Tobacco Products and Disposable Personal Income, 1989-2006. Department of Commerce, Bureau of Economic Analysis.

US Department of Health and Human Services (DHHS) (2004). *The Health Consequences of Smoking: A Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, CDC, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

USDHHS (2010). Office of Disease Prevention and Health Promotion. *Healthy People 2020*. Washington, DC. Available at <http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=41> Accessed 18 December, 2011.

USDHHS (2010). *Ending the Tobacco Epidemic: A Tobacco Control Strategic Action Plan for the U.S. Department of Health and Human Services*. Washington, DC: Office of the Assistant Secretary for Health.

US Federal Trade Commission (FTC), Smokless Tobacco Report for the Year 2006. Available at <http://www.ftc.gov/os/2009/08/090812smokelesstobaccoreport.pdf> Accessed 22 December 2011.

USDDHHS (2010) *Ending the Tobacco Epidemic: A Tobacco Control Strategic Action Plan for the U.S. Department of Health and Human Services*. Washington, DC: Office of the Assistant Secretary for Health.