

Health  
Information  
National  
Trends  
Survey

hints

# Health Information National Trends Survey 4 (HINTS 4)

## Cycle 1 Methodology Report

June 2012

**Prepared for**  
National Cancer Institute  
6120 Executive Boulevard  
Bethesda, MD 20892-7195

**Prepared by**  
Westat  
1600 Research Boulevard  
Rockville, MD 20850



## Table of Contents

<b>Chapter</b>		<b>Page</b>
1	Cycle 1 Overview.....	1
	1.1 Embedded Experiment: Respondent Selection.....	1
	1.2 Enhancements to Increase Response .....	2
2	Sample Selection .....	3
	2.1 Sampling Frame.....	3
	2.2 Stratification.....	4
	2.3 Selection of Address Sample .....	4
	2.4 Within-Household Sample Selection .....	5
3	Data Collection .....	5
	3.1 Mailing Protocol.....	5
	3.2 All-Adult Follow-up .....	7
	3.3 Spanish Language Households .....	7
	3.4 In-Bound Telephone Calls.....	8
	3.5 Incoming Questionnaires.....	9
4	Data Management.....	10
	4.1 Scanning .....	11
	4.2 Cleaning and Editing.....	12
	4.3 Imputation.....	13
	4.4 Determination of the Number of Household Adults.....	13
	4.5 Survey Eligibility.....	14
5	Weighting and Variance Estimation .....	15
	5.1 Household Base Weights .....	16
	5.2 Household Nonresponse Adjustment .....	16
	5.3 Initial Person-Level Weights .....	18
	5.4 Calibration Adjustments .....	19
	5.5 Replicate Variance Estimation .....	20
6	Response Rates .....	22
	6.1 Household Response Rate.....	22
	6.2 Within-Household Response Rate .....	23
	6.3 Overall Response Rate .....	24
	References .....	25

<b><u>Tables</u></b>	<b><u>Page</u></b>
2-1 Cycle 1 Sample Sizes by Stratum and Treatment Group.....	5
3-1 Mailing Protocol .....	6
3-2 Number of Packets per Mailing.....	7
3-3 Spanish Mailings .....	8
3-4 Telephone Calls Received.....	9
3-5 Final Household Status.....	10
3-6 Survey Response by Date.....	10
6-1 Household Response Rate .....	23
6-2 Household Response Rate Calculations for the All Adult Group by Strata.....	23
6-3 Within-Household Response Rate Calculations for All Adult Group by Strata.....	24
6-4 Final Response Rates by Selection Method.....	24
6-5 Final Response Rates for HINTS 4 Cycle 1 by Strata .....	24

**Appendices**

- A: Cover Letters in English
- B: Cover Letters in Spanish
- C: Frequently Asked Questions, English and Spanish
- D: Targeted Cover Letter for All Adult Household
- E: Variable Values and Edits
- F: Data Editing Procedures

The Health Information National Trends Survey (HINTS) is a nationally-representative survey which has been administered every few years by the National Cancer Institute since 2003. The HINTS target population is all adults aged 18 or older in the civilian non-institutionalized population of the United States. The most recent version of HINTS administration (referred to as HINTS 4) includes four mail-mode data collection cycles over the course of three years. The first of these cycles (Cycle 1) was conducted from October 2011 through February 2012. This report summarizes the methodology, sampling, and procedures of Cycle 1. Data cleaning and weighing procedures as well as response rates are discussed.

## Cycle 1 Overview

# 1

Cycle 1 was conducted by mail using a protocol similar to that used in the 2007 HINTS data collection with a goal of obtaining 3,500 completed questionnaires. Some enhancements were made to procedures to encourage response and an embedded experiment on respondent selection methods was included.

### 1.1 Embedded Experiment: Respondent Selection

There is little literature about the best way to have a sampled household select which member of the household will complete the questionnaire in an all-mail survey. In order to possibly inform future cycles of HINTS data collection, an embedded experiment was included in Cycle 1. There were two conditions:

- **The Next Birthday Method** asks that the adult with the next birthday complete the questionnaire. This condition results in one completed questionnaire per household. However, it is unclear from existing literature whether respondents identify the correct person from their household to complete the questionnaire.
- **The All Adult Method** (which was used in HINTS 2007) asks that all the adults in the household fill out a questionnaire. This can result in multiple completed questionnaires per household. It can also result in partially complete households where a completed questionnaire from every reported adult in the household is never received.

Households were randomly assigned to one of the two respondent selection conditions. A detailed description of the sampling methods and assignment can be found in Chapter 2.

## 1.2 Enhancements to Increase Response

A number of enhancements relative to HINTS 2007 were made to the data collection protocol in order to try to increase response rates.

**Follow-up to Partially Complete Households:** It was anticipated that some households assigned to the All Adult method would not send back questionnaires for all adults in the household. In HINTS 2007, these households were left as partially complete. However, in Cycle 1 of HINTS 4, follow-up procedures were developed to try to gain full participation. Questionnaires targeted at the nonresponding adults in the household were sent in subsequent mailings.

**Targeting Potential Spanish-Speaking Households:** In an effort to increase participation by Hispanic respondents, all materials (letters, FAQs, questionnaires) were translated into Spanish using the TRAP method (Harkness, et.al., 2010). Households were flagged as potentially Spanish-speaking in one of three ways:

- Linguistically isolated areas: The US Census Bureau defines linguistically isolated households as those in which everyone over 14 years old speaks a language other than English and does not speak English very well. Sampled households from Census tracts with relatively high proportions of linguistically isolated Spanish households were flagged as potentially Spanish-speaking.
- Hispanic surname match: The surnames provided by the address vendor were compared to typical Hispanic surnames. Households who were identified as having a Hispanic surname were flagged as potentially Spanish-speaking.
- Respondent request: Respondents who called Westat to request Spanish materials were flagged as Spanish-speaking.

An additional mailing was added to the protocol to further encourage the return of Spanish questionnaires. This mailing was sent only to households flagged as potentially Spanish-speaking.

**Short Version of the Instrument:** To try to encourage households that did not respond to early mailings, a shortened version of the questionnaire was sent in the third and fourth mailings. The shortened instrument reduced the number of items from 205 to 134.

A description of the operationalization of the mailing procedures can be found in Chapter 3.

The sample design for the HINTS Cycle 1 survey consisted of a two-stage design. In the first stage, a stratified sample of addresses was selected from a file of residential addresses, and in the second-stage, a sample of adults within the sampled households was selected.

## 2.1 Sampling Frame

The sampling frame consisted of a database of addresses used by Marketing Systems Group (MSG) to provide random samples of addresses. All non-vacant residential addresses in the United States present on the MSG database, including post office (P.O.) boxes, throwbacks (i.e., street addresses for which mail is redirected by the United States Postal Service to a specified P.O. box), and seasonal addresses, were subject to sampling.

Rarely are surveys conducted with a sampling frame that perfectly represents the target population. The sampling frame is one of the many sources of error in the survey process. The sampling frame used for the address sample contained duplicate units because some households receive mail in more than one way. To permit adjustment for this duplication of households in the sampling frame, we included a question on the mail questionnaire that asked how many different ways respondents receive mail.

In rural areas, some of the addresses do not contain street addresses or box numbers. Simplified addresses contain insufficient information for mailing questionnaires. Consequently, alternative sources of usable addresses were used when a carrier route contained simplified addresses. This partially ameliorated the frame's known undercoverage of rural areas, but the coverage and undeliverable rates for this portion of the frame is not known.

## **2.2 Stratification**

The sampling frame of addresses was grouped into three explicit sampling strata: the first stratum consisted of addresses in areas with high concentrations of minority population, the second stratum consisted of addresses in areas with low concentrations of minority population, and the third stratum consisted of addresses located in counties comprising Central Appalachia regardless of minority population. The high and low minority strata were formed using the block group level characteristics from the 2005 to 2009 American Community Survey (ACS) 5-year Summary File. Addresses in census block groups that had a population proportion of Hispanics or African Americans that equaled or exceeded 40 percent were assigned to the high-minority stratum. All the remaining addresses were assigned to the low-minority stratum. Addresses in counties comprising Central Appalachia were assigned to the Central Appalachia stratum regardless of minority status.

The purpose of creating high- and low-minority strata and then oversampling the high-minority stratum is to increase the precision of estimates for minority subpopulations. The gains in precision stem from the increase in sample sizes for the minority subpopulations produced by the oversampling. Central Appalachia is an add-on area with separate precision requirements. It is in a separate stratum to ensure an adequate sample size to meet these precision requirements.

## **2.3 Selection of Address Sample**

An equal-probability sample of addresses was selected from within each explicit sampling stratum. The total number of addresses selected for Cycle 1 was 12,385; 6,730 from the high minority stratum, 5,475 from the low minority stratum, and 180 from the Central Appalachia stratum. The high-minority stratum's proportion of the sampling frame was 23.4 percent, and it was oversampled so that its proportion of the sample was 54.3 percent. The Central Appalachia stratum's proportion of the sampling frame was 0.8 percent, and it was moderately oversampled so that its proportion of the sample was 1.5 percent. Conversely, the low minority stratum comprised 74.5 percent of the sampling frame, but made up just 44.2 percent of the sample.

## 2.4 Within-Household Sample Selection

Data collection implemented two approaches to select persons within sampled households: (1) all adult approach, and (2) next birthday approach. The former takes all adults in a sampled household, whereas the latter selects the adult in a sampled household having the next birthday. About 1/6 of all sampled addresses were assigned to the all-adult selection group, and the remaining 5/6 of the sample addresses were assigned to the next-birthday group. This assignment rate was made uniformly across all three strata.

The table below summarizes the sample selection for Cycle 1. It shows the total sample size of addresses broken out by sampling stratum and by within-household selection method.

Table 2-1. Cycle 1 Sample Sizes by Stratum and Treatment Group

Stratum	Total	Selection Method 1: Take All Adults	Selection Method 2: Next Birthday Adult
High minority areas	6,730	1,100	5,630
Low minority areas	5,475	888	4,587
Central Appalachia	180	29	151
Total	12,385	2,017	10,368

## Data Collection

# 3

Data collection for Cycle 1 started on October 25, 2011 and concluded on February 21, 2012. The survey was conducted exclusively by mail with a \$2 monetary incentive to encourage participation. The specific mailing procedures and outcomes are described in detail below.

### 3.1 Mailing Protocol

A total of five mailings were sent out as part of Cycle 1. All households in the sample received the first mailing, while only non-responding households received subsequent mailings. Households assigned to the Next Birthday (NB) method were sent one survey instrument per mailing with instructions that only one adult in the household was to respond. Households assigned to the All

Adult (AA) method were sent two instruments per mailing with instructions that all adults in that household should respond and with instructions to call Westat for more surveys if needed.

The mailing protocol followed a modified Dillman approach (Dillman, et.al., 2009) with a total of four mailings: an initial mailing, a reminder postcard, and three follow-up mailings, each of which was slightly different with the objective of increasing response. The second mailing was sent via USPS Priority Mail, while all other mailings were sent First Class. The third mailing contained the shortened version of the questionnaire. The second and third mailings targeted potentially Spanish-speaking households. An additional fourth mailing was sent to households targeted for the Spanish questionnaire. The contents of the mailings are further described in Table 3-1 below. Cover letters in English can be found in Appendix A and coverletters in Spanish are in Appendix B. All cover letters include a list of Frequently Asked Questions (FAQs) on the back. These FAQs in both English and Spanish are in Appendix C.

Table 3-1. Mailing Protocol

<b>Mailing</b>	<b>Date(s) Mailed</b>	<b>Mailing Method</b>	<b>Materials</b>	<b>Type of Recipients</b>
<b>Mailing 1</b>	Oct 25-27, 2011	1 <sup>st</sup> Class Mail	Cover letter with FAQs English questionnaire (1 or 2) \$2 bill Return envelope (1 or 2)	All sampled addresses
<b>Reminder</b>	Nov 4-7, 2011	1 <sup>st</sup> Class Mail	Reminder/thank you postcard	All non-responding households
<b>Mailing 2</b>	Nov 29, 2011	USPS Priority Mail	English cover letter with FAQs English questionnaire (1 or 2) Return envelope (1 or 2)	Non-responding households not identified as possibly Spanish-speaking
			English and Spanish cover letters with FAQs English and Spanish questionnaires (1 of each) Return envelope (1)	Non-responding households identified as possibly Spanish-speaking
<b>Mailing 3</b>	Dec 15, 2011	1 <sup>st</sup> Class Mail	English cover letter with FAQs English questionnaire, short version (1 or 2) Return envelope (1 or 2)	Non-responding households not identified as possibly Spanish-speaking
			English and Spanish cover letters with FAQs English and Spanish questionnaires, short version (1 of each) Return envelope (1)	Non-responding households identified as possibly Spanish-speaking
<b>Mailing 4</b>	Jan 10, 2012	1 <sup>st</sup> Class Mail	English and Spanish cover letters with FAQs English and Spanish questionnaires, short version (1 of each) Return envelope (1)	Non-responding households identified as possibly Spanish-speaking

The number of packets sent per mailing is outlined in Table 3-2 below. Households who sent in completed questionnaires were removed from further mailings. In addition, households whose packets were returned by the Postal Service as “undeliverable” were removed from any further mailings.

Table 3-2. Number of Packets per Mailing

Mailing	Number of Packets Mailed		
	Next Birthday	All Adult	Total
Mailing 1	10,368	2,018	12,386
Mailing 2	7,632	1,484	9,116
Mailing 3	6,899	1,357	8,256
Mailing 4	1,235	2	1,237
Total	26,134	4,861	30,995

### 3.2 All-Adult Follow-up

As noted above, households assigned to All Adult were asked to have each adult in the household return a questionnaire. As completed AA questionnaires were returned, the data from the screener page of the instrument and the household enumeration (question K12) were reviewed to see if all reported adults in the household returned a survey. AA households that returned less than the expected number of questionnaires were coded as “partially complete.” Targeted mailings were sent to these partially complete households to encourage full household participation. The cover letter included:

- How many questionnaires had been received from the household at that time;
- The age and gender of respondents that were still missing (based on the screener and Question K12); and
- The number of questionnaires included in that mailing. The number of questionnaires sent equaled the number of missing respondents.

The cover letter for AA targeted mailings can be found in Appendix D. The total number of AA targeted mailings conducted during the field period was 241.

### 3.3 Spanish Language Households

Extra efforts were made to elicit responses from Spanish-speaking households. The first mailing was sent in English only, but subsequent mailings were sent in both English and Spanish. Mailings

that included both English and Spanish materials were sent to households who:

- Were flagged as living in a linguistically-isolated area;
- Were flagged as having a potentially Hispanic surname; or
- Called to request materials in Spanish.

Any household that called to request Spanish materials then received materials in Spanish for all subsequent mailings regardless of any other flag status.

Because the initial return of Spanish-language surveys was low, a fourth mailing was conducted to specifically target potentially Spanish-speaking households (see table 3-1). This additional mailing included a sticker on the outside envelop in both English and Spanish asking that the questionnaire be returned within 2 weeks.

The number of Spanish packets sent during the field period is outlined in Table 3-3 below.

Table 3-3. Spanish Mailings

Mailing	Number of Packets Mailed		
	Both English and Spanish Materials	Spanish Materials Only	Total
Mailing 1	0	0	0
Mailing 2	1,444	20	1,464
Mailing 3	1,345	18	1,363
Mailing 4	1,235	2	1,237
Total	4,024	40	4,064

### 3.4 In-bound telephone calls

Two toll-free 800 numbers were provided to respondents -- one was used for English calls and one was used for Spanish calls. These numbers were both provided in each mailing. Respondents were told that they could call the number if they had comments, concerns, or if they needed to request anything special (such as additional questionnaires or a questionnaire in Spanish). Each number had a HINTS-specific voicemail message that instructed callers to leave their contact information and the reason for the call, and then a study staff member would return their call. The Spanish line was staffed by a native Spanish speaker. When voicemails were received, they were logged into the Survey Management System (SMS) and the request was either processed (such as recording their desire for a Spanish questionnaire) or the respondent was called back to ascertain the respondent's need if it was not clear from the message. Callers stating they did not want to participate in the study were coded as "refusal" and removed from any subsequent mailings.

The two toll-free lines together received 55 calls throughout the Cycle 1 field period (Table 3-4). A majority of the in-bound calls were to request Spanish materials. The rest were refusals or respondents calling in with some form of a comment or a question. Six calls were never able to be resolved because the study staff were never able to reach the respondent.

Table 3-4. Telephone Calls Received

<b>Reason for Call</b>	<b>Number of Calls Received</b>
Request for Spanish questionnaires	27
Refusal	10
Request for English questionnaires	4
Caller wanted to ensure that nothing more would come after they responded	2
Privacy concerns	2
Comments about the questionnaire being confusing	1
A promise that the caller would be sending the survey soon	1
A concern about not knowing all the answers	1
Request for replacement of a damaged \$2 bill	1
Calls that were never resolved	6
<b>Total</b>	<b>55</b>

### 3.5 Incoming Questionnaires

Field room staff receipted into the SMS all received questionnaires using each questionnaire’s unique barcode. The SMS tracked each received questionnaire, as well as the status of each household (nonresponsive, complete, or partially complete). Once a household was deemed complete, it no longer received any additional mailings. Packages that came back as undeliverable were marked as such in the SMS and those addresses did not receive any further mailings.

In addition to refusing by calling the toll-free line, some respondents also refused by sending a letter stating that they did not wish to participate or asking to be removed from our mailing list. These households were marked in the system as refusals and were removed from subsequent mailings. Respondents who sent back a blank questionnaire were not considered refusals and continued to receive mailings.

The final status of all households can be found in Table 3-5 below.

Table 3-5. Final Household Status

Household-level Results	All Adult		Next Birthday		TOTAL	
	N	%	N	%	N	%
Complete	432	21.4	3,133	30.2	3,565	28.8
Partial Complete	130	6.5			130	1.0
Hard refusal	10	0.5	67	0.6	77	0.6
Undeliverable	257	12.7	1,241	12.0	1,498	12.1
Nonresponse	1,188	58.9	5,927	57.2	7,115	57.5
TOTAL	2,017	100.0	10,368	100.0	12,385	100.0

The number of questionnaires returned by date during the field period can be found in Table 3-6 below.

Table 3-6. Survey Response by Date

Date of Mailing	Period of Returns	Number of Returns
Mailing 1: October 25-27	October 26-November 4	497
Postcard: November 4-7	November 5-November 29	1,950
Mailing 2: November 29	November 30-December 15	693
Mailing 3: December 15	December 16-January 10	642
Mailing 4: January 10 (Spanish Households only)	January 11-February 21	177
Total		3,959

## Data Management

# 4

After being processed and entered into the SMS, each returned questionnaire was scanned, verified, cleaned and edited. Imputation procedures were also conducted. These procedures are described below. A list of the variables where initial values were edited, imputed, or recoded can be found in Appendix E (Variable Values and Edits).

## 4.1 Scanning

All completed questionnaires were electronically scanned to capture the survey data and images. Staff reviewed each form as it was prepared for scanning. The review included:

- Determining if the form was not scannable for any reason such as being damaged in the mail. Some questionnaires needed to be overwritten with a pen that was readable by the data capture software; and
- Documenting potential partial complete questionnaires or pertinent comments made by respondents in a decision log. Comments in Spanish were reviewed by a Spanish-speaking staff member.

The reviewed surveys were then sent through the high-speed TeleForm scanner to capture the responses. TeleForm read the form image files and extracted data according to HINTS Cycle 1 rules established prior to the field period. Scanned data were then subject to validation according to HINTS specifications. If a data value violated validation rules, such as marking more than one choice box in a mark-only-one question, the data item was flagged for review by verifiers who looked at the images and the corresponding extracted data and resolved any discrepancies. Spanish forms were verified by a Spanish-speaking staff member.

Decisions made about data issues were recorded in a data decision log. The decision log contained the respondent ID, the value triggering the edit, the updated value, and the reason for the update. A total of 408 entries were made into the data decision log during the course of data processing. Approximately 95 percent of these were attributed to multiple responses on a mark-only-one question.

A 10% quality control check was then conducted on the scanned data and the electronic images of the data forms. Quality Assurance (QA) staff compared the hard copy questionnaire to the data captured in the database item-for-item and the images stored in the repository page-for-page to ensure that all items were correctly captured. If needed, updates were made. In addition, QA staff closely reviewed frequencies and cross tabulations of the HINTS raw data to identify outliers and open ended items to be verified. ID reconciliation across the database, images, and the SMS was completed to confirm data integrity.

## 4.2 Data Cleaning and Editing

Once scanned, the data were cleaned and edited. General cleaning and editing activities are described briefly below, with more detailed information found in Appendix F, Data Editing Procedures.

- Customized range and logical inconsistency edits, following predetermined processing rules to ensure data integrity, were developed and applied against the data.
- Edit rules were created to identify and recode nonresponse or indeterminate responses. Appendix E (Variable Values and Edits) provides a list of the values and their definitions.
- Missing values were recoded for some responses to questions that featured a forced-choice response format and for filter questions where responses to later questions suggested a particular response was appropriate. Appendix E (Variable Values and Edits) provides details about when and how these recodings took place.
- Variables were designed to summarize the response for the race and cancer questions. These variables, `race_cat` and `cancer_cat`, indicated each response selected for respondents selecting only one response, and a multiple category for all of the respondents who answered multiple responses.
- Derived variables were created to reflect each response recorded for the occupation question in order to capture this information when respondents did not follow the instruction to mark only one response. For some variables for which a respondent was expected to select only one response, rules, as described in Appendix E (Variable Values and Edits), were used to determine which response was retained. For other variables, imputation, as described below, was carried out.
- Data cleaning was carried out for the two height variables: `Height_Feet` and `Height_Inches`. The rules that were applied minimized the number of out-of-range values by accounting for response measurements in incorrect boxes, responses using metric, responses using only one unit of measurement and other response errors. A list of the rules applied to clean these variables is included in Appendix F (Data Editing Procedures).
- “Other, specify” responses were examined, cleaned for spelling errors, categorized, and upcoded into preexisting response codes when applicable.

### **4.3 Imputation**

For two variables featuring a mark-only-one response instruction (WhereSeekHealthInfo and StrongNeedHealthInfo), imputation was carried out for the questionnaires in which multiple responses were recorded. The imputation process used was the same as that carried out for the HINTS 2007 survey. Responses where a missing value of -5 (multiple responses selected) was applied were imputed. This occurred for 466 respondents for WhereSeekHealthInfo and 160 respondents for StrongNeedHealthInfo. A respondent's multiple answers were replaced with a single imputed answer that had the same distribution over the multiple answers as occurred in the single-answer responses.

In addition, hot-deck imputation was used to replace missing responses with imputed data for items used in the raking procedure of the Cycle 1 weighting process. Hot-deck imputation is a data processing procedure in which a case with a missing value for specific variable is assigned the corresponding value of a "similar" case in the same imputation class. The data record that supplies the imputed value is referred to as the "donor." Under a hot deck approach, the resulting distribution preserves the distribution of values observed for respondents. Imputation classes are defined on the basis of variables that are thought to be correlated with the item with missing values. A donor is then randomly selected within an imputation class to supply the imputed value. Items imputed using the hot-deck approach were those involving the following characteristics: age, gender, educational attainment, marital status, race, ethnicity, health insurance coverage, and cancer diagnosis.

### **4.4 Determination of the Number of Household Adults**

For the purpose of applying weights, a measure of the number of adults in each household, 'R\_HHAdults,' was created using questionnaire responses. The initial measure was taken from responses to demographic section questions asking for the total number of people and the number of children in the household. Implausible or missing values that resulted from the answers to those questions were substituted with values to questions on the respondent-selection page of the questionnaire, and further substituted with data from the demographic section roster. Edits were carried out to reconcile different values reported within households and correct differences with the receipted number of returned questionnaires. A detailed list of the steps carried out to identify the number of adults in each household is included in Appendix F (Data Editing Procedures).

## 4.5 Survey Eligibility

Of the 3,998 questionnaires received, 22 were incompletely-answered and a further 17 were determined to be ineligible since they were duplicates from the same respondent or were from incorrect respondents (according to the ‘next-birthday’ rule). A total of 3,959 questionnaires were coded as eligible questionnaires. Details about the eligibility rules are below.

### Definition of a Complete and Partial Complete Questionnaire

The procedures for determining whether or not a returned questionnaire was complete were similar to those used for the HINTS 2007 mail survey. A complete questionnaire was defined as any questionnaire with at least 80% of the required questions answered in Sections A and B. A partial complete was defined as when between 50% and 79% of the questions were answered in Sections A and B. There were 52 partially complete questionnaires. Both partially-completed and completely-answered questionnaires were retained. The 22 questionnaires with fewer than 50% of the required questions answered in Sections A and B were coded as incompletely-filled out and discarded.

### Eligibility of Multiple Questionnaires from a Household

Some households returned more than the requested number of completed questionnaires. In these cases, procedures were developed to identify which of the questionnaires were eligible for the survey. These procedures differed for the two respondent selection methods and are described here:

- **Next Birthday Households:** A single completed survey was expected from each Next Birthday household.
  - If the same respondent returned multiple questionnaires, the first questionnaire received was retained unless it was a reduced version and a subsequent return was a full version.
  - If a return date was unavailable for questionnaires from the same respondent that were the same version, questionnaires with fewer substantive questions omitted were retained.
  - If different respondents returned a questionnaire and the ages of household members listed in the roster were in agreement (or differed by only one year), the questionnaire that complied with the next birthday rule was retained.<sup>1</sup>
  - If, in the above situation, compliance for one or both questionnaires from a household was unclear, the first questionnaire returned was retained
  - If different respondents returned a questionnaire and the ages of household members listed in the roster question were not substantively in agreement, the earliest questionnaire received that complied with the next birthday rule was retained.

---

<sup>1</sup> Compliance was determined by whether the person listed in the roster who matched the respondent’s age and gender had a month of birth that was the first to follow the month in which the questionnaire was returned.

- **All Adult Households:** The household could return one or more questionnaires, depending on information provided by the respondents.
  - If the ages of household members were in general agreement for some of the returned questionnaires, the questionnaire that did not have any of the same ages listed in the roster question as the other returned household questionnaires was coded as ineligible and the questionnaires that had other household members' ages in the roster question cross-listed were retained.

Following cleaning and editing, a detailed codebook including frequencies was created for both the weighted and unweighted data. The codebooks define all variables in the dataset, provide the question text, list the allowable codes, and explain the inclusion criteria for each item. All versions of the both the English and Spanish instruments were annotated with variable names and allowable codes to support the usability of the delivery data.

## Weighting and Variance Estimation

# 5

Every sampled adult who completed a questionnaire in Cycle 1 received a full-sample weight and a set of 50 replicate weights. The full-sample weight is the weight which is used to calculate population and subpopulation estimates from the HINTS data collected in Cycle 1. Replicate weights are used to compute standard errors for these estimates.

The weighting process encompasses the procedures used to create the final full-sample and replicate weights for the survey respondents. The use of sampling weights is done to ensure valid inferences from the responding sample to the population, correcting for nonresponse and noncoverage biases to the extent possible. The computation of the full-sample weights for Cycle 1 consisted of the following steps. Each of these steps will be described in more detail in the sections 5.1 through 5.4.

- Calculating household-level base weights;
- Adjusting for household nonresponse;
- Calculating person-level initial weights;
- Calibrating the person-level weights to population counts (also known as control totals).

The replicate weights were calculated using the jackknife method. The steps used to derive these weights were aimed at reflecting the features of the sample design, so that when the jackknife variance estimation procedure is implemented, approximately unbiased estimates of sampling variance are obtained. In addition, the various weighting procedures carried out on the full-sample weight were repeated on each set of replicate weights to appropriately reflect the impact of the weighting adjustments on the sampling variance of a survey estimate. Details of replication used for variance estimation can be found in section 5.5.

## **5.1 Household Base Weights**

The initial step in the weighting process was calculating the household-level base weight for each household in the sample. The household base weight is the reciprocal of the probability of selecting the household for the survey, which depends on the stratum the household was selected from. Generally, base weights for units in oversampled strata are smaller than those in strata that were not oversampled. In Cycle 1, the base weights for households in the high minority and Central Appalachia strata were roughly  $1/4$  and  $1/3$ , respectively, the size of those in the low minority stratum.

If two different addresses led to the same household – for example, if a household receives mail via both a street address and a post office box – that household had twice the chance of selection of a household with only one address (and should therefore receive half the normal weight). An additional adjustment was made to the base weights of households that had multiple ways of receiving mail (as determined by the answers to a survey question about this).

## **5.2 Household Nonresponse Adjustment**

Nonresponse is generally encountered to some degree in every survey. The first and most obvious effect of nonresponse is to reduce the effective sample size, which increases the sampling variance. In addition, if there are systematic differences between the respondents and the nonrespondents, there also will be a bias of unknown size and direction. This bias is generally adjusted for in the case of unit nonrespondents (nonrespondents who refuse to answer any part of the questionnaire) with the use of a weighting adjustment term multiplied to the base weights of sample respondents. Item nonresponse (nonresponse to specific questions only) is generally adjusted for through the use of imputation. This section discusses weighting adjustments for unit nonresponse.

The most widely accepted paradigm for unit nonresponse weighting adjustment is the quasi-randomization approach (Oh & Scheuren, 1983). In this approach, nonresponse cells are defined based on those measured characteristics of the sample members that are known to be related to response propensity. For example, if it is known that males respond at a lower rate than females, then sex should be one characteristic used in generating nonresponse cells. Under this approach, sample units are assigned to a response cell, based on a set of defined characteristics. The weighting adjustment for the sample unit is the reciprocal of the estimated response rate for the cell. Any set of response cells must be based on characteristics that are known for all sample units, responding and nonresponding. Thus questionnaire items on the survey cannot be used in the development of response cells, because these characteristics are only known for the responding sample units.

Under the quasi-randomization paradigm, Westat models nonresponse as a “sample” from the population of adults in that cell. If this model is in fact valid, then the use of the quasi-randomization weighting adjustment eliminates any nonresponse bias (see, for example, Little & Rubin (1987), Chapter 4).

The weighting procedure for Cycle 1 used a household-level nonresponse adjustment procedure based on this approach. The base weights of the households that did return at least one questionnaire were adjusted to reflect nonresponse by the remaining eligible households. A search algorithm<sup>2</sup> was used to identify variables highly correlated with household-level response, and these variables were used to create the nonresponse adjustment cells. The variables used to define nonresponse cells for Cycle 1 were:

- Sampling stratum (High Minority; Low Minority; Central Appalachia)
- Treatment group (All Adult method; Next Birthday method)
- Census region (Northeast; South; Midwest; West)
- Route type (Street address; Other addresses such as PO Box, Rural Route, etc)
- Metropolitan Status (county in Metro areas; county in Non-Metro areas)

---

<sup>2</sup> An inhouse macro WESSEARCH, which calls the Search software, a freeware product developed by the University of Michigan (<http://www.isr.umich.edu/src/smp/search/>.)

Nonresponse adjustment factors were computed for each nonresponse cell  $b$  as follows:

$$HH\_NRAF(b) = \frac{\sum_{S(b)} HH\_BWT_i}{\sum_{C(b)} HH\_BWT_i},$$

where  $HH\_BWT_i$  is the base weight for sampled household  $i$ ,  $S(b)$  is the set of all eligible sampled households) in nonresponse cell  $b$ ,  $C(b)$  is the set of all cooperating sampled households in cell  $b$ , and  $HH\_NRAF(b)$  is the household nonresponse adjustment factor for nonresponse cell  $b$ .

The household nonresponse adjustment factors ranged from a low of 1.60 to a high of 4.41, and averaged 2.67 across all nonresponse adjustment cells.

### 5.3 Initial Person-Level Weights

Each sampled adult in responding households was assigned an initial person-level weight. The initial person-level weight was calculated by multiplying the nonresponse-adjusted household weight by the reciprocal of the sample person's within-household probability of selection. If all adults within a household were selected to participate in the survey, the initial weight was identical to the nonresponse-adjusted household weight. However, if only one adult was selected from a household, then the initial weight for the sampled adult was equal to the nonresponse-adjusted weight times the number of eligible adults in that household. For example, if a household contains three adults and only one adult was selected, the initial weight for the selected adult is equal to the nonresponse-adjusted household weight times three.

The initial person weight also included an additional adjustment for within-household nonresponse for households where all adults in the household were selected, and only some, but not all, adults participated in the survey. The adjustment was made at the household level, and an adjustment factor was calculated by dividing the number of eligible adults in the households by the number of responding adults. For example, if a household contains four adults but only two responded, an additional weighting adjustment of 2 ( $=4/2$ ) was applied to the person-level weight. This adjustment is very similar to the calculation of the within-household weight for households where one adult was selected.

## 5.4 Calibration Adjustments

The purpose of calibration is to reduce the sampling variance of estimators through the use of reliable auxiliary information (see, for example, Deville & Sarndal, 1992). In the ideal case, this auxiliary information usually takes the form of known population totals for particular characteristics (called *control totals*). However, calibration also reduces the sampling variance of estimators if the auxiliary information has sampling errors, as long as these sampling errors are significantly smaller than those of the survey itself.

Calibration reduces sampling errors particularly for estimators of characteristics that are highly correlated to the calibration variables in the population. The extreme case of this would be the calibration variables themselves. The survey estimates of the control totals would have considerably higher sampling errors than the “calibrated” estimates of the control totals, which would be the control totals themselves. The estimator of any characteristic that is correlated to any calibration variable will share partially in this reduction of sampling variance, though not fully. Only estimators of characteristics that are completely uncorrelated to the calibration variables will show no improvement in sampling error. Deville and Sarndal (1992) provide a rigorous discussion of these results.

### Control Totals

The American Community Survey (ACS) of the U.S. Census Bureau has much larger sample sizes than those of HINTS. The ACS estimates of any U.S. population totals have lower sampling error than the corresponding HINTS estimates, making calibration of the survey weights to ACS control totals beneficial. The ACS estimates are available via the Internet. Westat used the 2010 ACS estimates that were available on the Census Bureau web site.

Calibration variables were selected among those that were on the ACS public-use file and were found to be well correlated to important HINTS questionnaire item outcomes (i.e., Westat wanted ACS-available characteristics that tend to have differing mean values for HINTS questionnaire item outcomes). The following ACS characteristics correlate well with HINTS questionnaire items:

- Age
- Gender
- Educational Attainment
- Marital Status
- Race
- Ethnicity
- Census Region

In addition to characteristics from the ACS, two health-related variables were used. These variables came from the 2010 National Health Information Survey (NHIS) and correspond to questions asked in the HINTS survey. They were:

- Percent With Health Insurance
- Percent Ever Had Cancer

Raking to the control totals for these variables (either alone or cross-classified with each other) was then performed. As a result of the raking HINTS weights to the control totals, estimates calculated from HINTS data for the control-total variables agree with those calculated from the source data for the control totals. For example, the national-level estimate of Percent Ever Had Cancer calculated from HINTS data agrees with the estimate calculated from NHIS 2010 data.

## **5.5 Replicate Variance Estimation**

In addition to the full-sample weight, a set of 50 replicate weights were provided for each adult. These replicate weights are used to calculate standard error of estimates obtained from the HINTS data, using the jackknife replication method.

The jackknife technique is compatible with the sample design and weighting procedures for HINTS. The jackknife variance estimation technique takes carefully selected subsets of the data for each “replicate,” and for each respondent in the replicate subset and determines a sampling weight, as if the replicate subset were in fact the responding sample. (This replicate subset is usually almost the entire sample, except for a group of respondents that are “deleted” for that replicate.) The resulting weights are called replicate weights.

The jackknife variance estimator requires the use of replicate weights. For the Cycle 1 data set, a set of 50 replicate weights was assigned to each responding adult. To illustrate how the replicate weights

are computed, suppose  $P$  is a percentage of adults in the U.S. population having a particular characteristic (e.g., answering one of the HINTS questions in a particular way). A nationally representative estimator  $p$  can be computed by aggregating the adult sampling weights of all responding adults with this characteristic (e.g., all responding adults in the survey answering the survey question in a particular way). A jackknife variance estimator of the sampling variance of  $p$  can be computed in two steps:

- Step 1. Recompute estimators  $p(r)$ ,  $r=1,\dots,50$ , by aggregating the replicate sampling weights corresponding to replicate  $r$  for all responding adults with the characteristic.
- Step 2. Compute the jackknife variance estimator

$$v(p) = \frac{R-1}{R} \sum_{r=1}^{50} (p(r) - p)^2$$

The replicate weights are computed by systematically deleting a portion of the original sample, and recomputing the sampling weights as if the remaining sample (without the deleted portion) were the actual sample. These deleted sample units should be first-stage sampling units, which in HINTS are households. The remainder of the sample with the deleted portion removed is called the replicate subset, and it should mirror the full sample design, as if it were a reduced version of the original sample.

For the purposes of jackknife variance estimation, each household was assigned to one of 50 replicate “deletion” groups  $D(r)$ ,  $r=1,\dots, 50$ . Each replicate sample is the full sample minus the deletion group (i.e., it is roughly 49/50 of the original sample).

The replicate sampling weights were generated in a series of steps that parallel the steps computing the full-sample sampling weights. The replicate base weight for each sampled household or adult and each replicate is either equal to  $R/(R-1)$  times the full sample base weight (if the household is contained in the replicate subset) or equal to 0 (if the household is not contained in the replicate subset, but instead is contained in the “deleted” set for that replicate).

Nonresponse and calibration adjustments were then computed for each set of replicate weights, using the replicate weights in the computation of nonresponse and calibration adjustments in place of the original weights. These calculations generated a set of replicate nonresponse and

poststratification adjustments for each responding adult. The final replicate weights were products of the replicate weights, nonresponse adjustments, and calibration adjustments.

## Response Rates

# 6

Response rates were calculated separately for each of the respondent selection methods using the RR2 formula of the American Association of Public Opinion Research (AAPOR). For the Next Birthday (NB) procedure the response rate was computed as a single stage survey. For the All Adult (AA) method, the rate was computed in two stages. For the first stage a household level response rate was calculated based on those cases that at least one survey was returned for a sampled address. The NB response rate and the AA household response rate is described in Section 6.1.

The second stage for the AA method was calculated using the total number of surveys that were returned divided by the total number of adults in those households that returned at least one survey. As noted in prior sections, the AA method asked that each adult fill out a questionnaire. In addition, as a part of the survey, the respondent was asked to record the number of adults in the household. Comparing the number of surveys returned to the number of adults in the household (as recorded in the survey) gives the within-household response rate. This is discussed in Section 6.2. The final section provides the final response rate, once combining across the two selection methods and stages of calculation.

### 6.1 Household Response Rate

Table 6-1 shows the response rate calculation for the NB method. These data have been weighted to account for the oversampling of addresses in high-minority areas. The percent of undeliverable letters ranged from 11 percent to 16 percent, with the Central Appalachia strata having the highest. The overall response rate for NB was 37.91 percent, however this differed significantly by strata. The High Minority strata had the lowest response rate (29.14%) and the Low Minority had the highest (40.61%).

Table 6-1. Response rate calculations for Next Birthday Group by Strata

Response class	High minority	Low minority	Central Appalachia	Overall
Total sample	26,163,191	85,018,513	949,393	112,131,096
Respondents	6,668,593	30,674,872	56,568	37,613,822
Nonrespondents	16,218,390	44,853,892	528,139	61,600,421
Undeliverable	3,276,208	9,489,749	150,897	12,916,854
Households	22,886,983	75,528,764	798,496	99,214,243
Percent undeliverable	12.52%	11.16%	15.89%	11.52%
Household response rate	29.14%	40.61%	33.86%	37.91%

A similar pattern emerges for the household-level response rate for the AA selection method. The undeliverable rates were slightly higher, but the pattern by strata was the same as for the NB methodology. The household-level response rate was also slightly lower than for NB.

Table 6-2. Household response rate calculations for the All Adult Group by Strata

Response class	High minority	Low minority	Central Appalachia	Overall
Total sample	5,111,813	16,458,783	182,334	21,752,931
Respondents	1,194,306	5,504,796	56,586	6,755,688
Nonrespondents	3,215,795	9,100,521	88,023	12,404,340
Undeliverable	701,713	1,853,467	37,724	2,592,903
Households	4,410,101	14,605,317	144,610	19,160,027
Percent undeliverable	13.73%	11.26%	20.69%	11.92%
Household response rate	27.08%	37.69%	33.86%	35.26%

## 6.2 Within-Household Response Rate

For the AA method, all adults were asked to respond. Table 6-3 shows the person-level rates for this methodology. The sum of the weights of those who completed a questionnaire are compared to the sum of the weights of the self-reported number of adults in a household to arrive at the within-household response rate. As with the household-level response rate, the within-household response

rate varies across stratum in a similar way as the household-level rate. The highest rate is for the Low Minority strata and the lowest is for the Central Appalachia strata.

Table 6-3. Within-household response rate calculations for All Adult Group by Strata

Response class	High minority	Low minority	Central Appalachia	Overall
Respondents	1,619,515	8,284,996	56,586	9,961,097
Adults	2,046,817	9,644,050	88,023	11,778,890
Within-household response rate	79.12%	85.91%	64.29%	84.57%

### 6.3 Overall Response Rate

The overall response is computed by taking the product of the household and within-household (person-level) response rates for the AA method and then combining the rates for the two methods. Once multiplying the household and within-household response rates for the AA, the final rate is 29.82 percent. This is approximately 8 percentage points below the rate for the NB (37.91%).

Table 6-4. Final response rates by Selection Method

Response type	Next Birthday Method	All Adult Method
Household	37.91%	35.26%
Within-household	NA	84.57%
Overall response rate	37.91%	29.82%

The final response rate for the study is a combination of the final NB and AA rates. This combination is done in proportion to the size of the sample allocated for each methodology. As shown in Table 6-5, this rate is 36.67 percent.

Table 6-5. Final Response rates for HINTS 4 Cycle 1 by Strata

	High minority	Low minority	Central Appalachia	Overall
Final Response Rate	27.97%	39.34%	32.62%	36.67%

## References

- Deville, J.C., and Sarndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.
- Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: John Wiley & Sons.
- Harkness, Janet A., Ana Villar and Brad Edwards (2010) “Translation, Adaptation, and Design” in *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* J. Harkness, M. Braun, B. Edwards, T. Johnson, L. Lyberg, P. Mohler, B. Pennell, and T. Smith (eds), Chapter 7, pgs 117-140.
- Little, R., and Rubin, D.B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.
- Oh, H., and Scheuren, F. (1983). Weighting adjustments for unit response. In W.G. Madow, I. Olkin, and D. B. Rubin (Eds.), *Incomplete data in sampling surveys, Vol. II: Theory and annotated bibliography*. New York: Academic Press.

**Appendix A:**  
**Cover Letters in English**



**ALL ADULT COVER LETTER – FIRST MAILING**

Dear {City} Resident:

We are writing to invite you to take part in an important national survey sponsored by the U.S. Department of Health and Human Services, the Health Information National Trends Survey (HINTS).

The goal of HINTS is to learn what health information people want to know and where they try to find it. By completing this questionnaire, you will help us learn what health information you need and how to make that information available to you, your family and your community.

**We ask that each adult in this household complete a questionnaire and return it to us in the next two weeks.**

Your participation is voluntary and your responses will not be linked to your name. We have enclosed \$2 as a token of our appreciation for your participation.

You can find out more about HINTS at [hints.cancer.gov](http://hints.cancer.gov). Westat, a research firm, will conduct the survey. If you have any questions about HINTS or if you need more questionnaires, or if you would like to complete this survey in a language other than English or Spanish, please call Westat toll-free at 1-888-738-6805.

Thank you in advance for your cooperation.

Sincerely,



Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.



**NEXT BIRTHDAY COVER LETTER – FIRST MAILING**

Dear {City} Resident:

We are writing to invite you to take part in an important national survey sponsored by the U.S. Department of Health and Human Services, the Health Information National Trends Survey (HINTS).

The goal of HINTS is to learn what health information people want to know and where they try to find it. By completing this questionnaire, you will help us learn what health information you need and how to make that information available to you, your family and your community.

**In order to make sure we get responses from a random sample of people, we ask that the adult in your household with the next birthday complete and return this questionnaire in the next two weeks.**

Your participation is voluntary and your responses will not be linked to your name. We have enclosed \$2 as a token of our appreciation for your participation.

You can find out more about HINTS at [hints.cancer.gov](http://hints.cancer.gov). Westat, a research firm, will conduct the survey. If you have any questions about HINTS, or if you would like to complete this survey in a language other than English or Spanish, please call Westat toll-free at 1-888-738-6805.

Thank you in advance for your cooperation.

Sincerely,

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.

### ALL ADULT POSTCARD TEXT

A few days ago you should have received a questionnaire packet asking for your household's participation in the Health Information National Trends Survey. By completing the questionnaire, you can help the U.S. Department of Health and Human Services determine the best ways of communicating important health information to members of your community.

**We are inviting each adult in the household to complete a questionnaire.** If all adults have already completed a questionnaire and returned it to us, please accept my sincere thanks. If all adults in this household have not yet completed and returned a questionnaire, we ask that they do so as soon as possible.

Your household's participation is important and will help inform efforts to provide timely and relevant health information to you, your family, and your community.

Sincerely,



Bradford W. Hesse, Ph.D.  
HINTS Project Officer  
National Institutes of Health  
U.S. Dept. of Health and Human Services

---

### NEXT BIRTHDAY POSTCARD TEXT

A few days ago you should have received a questionnaire packet asking for your household's participation in the Health Information National Trends Survey. By completing the questionnaire, you can help the U.S. Department of Health and Human Services determine the best ways of communicating important health information to members of your community.

**We are inviting one adult in the household to complete a questionnaire.** If the adult identified on the questionnaire has already completed it and returned it to us, please accept my sincere thanks. If that adult has not yet completed and returned the questionnaire, we ask that he or she please do so as soon as possible.

Your household's participation is important and will help inform efforts to provide timely and relevant health information to you, your family, and your community.

Sincerely,



Bradford W. Hesse, Ph.D.  
HINTS Project Officer  
National Institutes of Health  
U.S. Dept. of Health and Human Services



**ALL ADULT COVER LETTER – SECOND MAILING**

Dear {City} Resident:

We recently invited you to participate in an important national survey sponsored by the U.S. Department of Health and Human Services (HHS), the Health Information National Trends Survey (HINTS). By completing the HINTS survey, you will help us keep you and your family better informed on the health issues that matter to you.

If you already sent back your survey and it crossed in the mail with this letter, thank you for the time you took to help make this study a success.

If you have not yet had a chance to complete the HINTS survey, we understand that it is sometimes difficult to find the time to participate in a study like HINTS. In order to reduce the amount of time it will take, we have enclosed a shorter version of the survey. This shorter version focuses on issues that tell us how individuals like you find and use health information.

**We ask that each adult in this household complete a questionnaire and return it to us in the next two weeks.**

Additional information about HINTS is available at: [hints.cancer.gov](http://hints.cancer.gov). If you have any questions or would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Thank you in advance for contributing to this important national study.

Sincerely,

A handwritten signature in blue ink that reads "Bradford W. Hesse".

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.



**NEXT BIRTHDAY COVER LETTER – SECOND MAILING**

National Institutes of Health  
Bethesda, Maryland 20892

Dear {City} Resident:

We recently invited you to participate in an important national survey sponsored by the U.S. Department of Health and Human Services (HHS). The goal of the Health Information National Trends Survey (HINTS) is to learn what health information people want to know and where they go to find it. Your responses will help us keep you, your family and members of your community better informed on the health issues that matter to you.

We have not yet received your completed questionnaire. To make sure HINTS provides accurate information, we need all the households invited to participate in this year's HINTS to complete the survey. If you did send back your survey and it crossed in the mail with this letter, thank you for the time you took to help make this study a success. In the event that your questionnaire was misplaced, an additional copy is enclosed.

**In order to make sure we get responses from a random sample of people, we ask that the adult in your household with the next birthday complete and return this questionnaire in the next two weeks.**

Additional information about HINTS is available at: [hints.cancer.gov](http://hints.cancer.gov). If you have any questions, or would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Thank you in advance for contributing to this important national study.

Sincerely,

A handwritten signature in blue ink that reads "Bradford W. Hesse".

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**



**ALL ADULT COVER LETTER – THIRD MAILING**

Dear {City} Resident:

We recently invited you to participate in an important national survey sponsored by the U.S. Department of Health and Human Services (HHS), the Health Information National Trends Survey (HINTS). By completing the HINTS survey, you will help us keep you and your family better informed on the health issues that matter to you.

If you already sent back your survey and it crossed in the mail with this letter, thank you for the time you took to help make this study a success.

If you have not yet had a chance to complete the HINTS survey, we understand that it is sometimes difficult to find the time to participate in a study like HINTS. In order to reduce the amount of time it will take, we have enclosed a shorter version of the survey. This shorter version focuses on issues that tell us how individuals like you find and use health information.

**We ask that each adult in this household complete a questionnaire and return it to us in the next two weeks.**

Additional information about HINTS is available at: [hints.cancer.gov](http://hints.cancer.gov). If you have any questions or would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Thank you in advance for contributing to this important national study.

Sincerely,

A handwritten signature in blue ink that reads "Bradford W. Hesse".

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.



**NEXT BIRTHDAY COVER LETTER – THIRD MAILING**

Dear {City} Resident:

We recently invited you to participate in an important national survey sponsored by the U.S. Department of Health and Human Services (HHS), the Health Information National Trends Survey (HINTS). By completing the HINTS survey, you will help us keep you and your family better informed on the health issues that matter to you.

If you already sent back your survey and it crossed in the mail with this letter, thank you for the time you took to help make this study a success.

If you have not yet had a chance to complete the HINTS survey, we understand that it is sometimes difficult to find the time to participate in a study like HINTS. In order to reduce the amount of time it will take, we have enclosed a shorter version of the survey. This shorter version focuses on issues that tell us how individuals like you find and use health information.

**In order to make sure we get responses from a random sample of people, we ask that the adult in your household with the next birthday complete and return this questionnaire in the next two weeks.**

Additional information about HINTS is available at: [hints.cancer.gov](http://hints.cancer.gov). If you have any questions or would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Thank you in advance for contributing to this important national study.

Sincerely,

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.



**ALL ADULT COVER LETTER – FOURTH MAILING**

Dear {City} Resident:

We recently invited you to participate in an important national survey sponsored by the U.S. Department of Health and Human Services (HHS), the Health Information National Trends Survey (HINTS). By completing the HINTS survey, you will help us keep you and your family better informed on the health issues that matter to you.

If you already sent back your survey and it crossed in the mail with this letter, thank you for the time you took to help make this study a success.

If you have not yet had a chance to complete the HINTS survey, we understand that it is sometimes difficult to find the time to participate in a study like HINTS. In order to reduce the amount of time it will take, we have enclosed a shorter version of the survey. This shorter version focuses on issues that tell us how individuals like you find and use health information.

**We ask that each adult in this household complete a questionnaire and return it to us in the next two weeks.**

Additional information about HINTS is available at: [hints.cancer.gov](http://hints.cancer.gov). If you have any questions or would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Thank you in advance for contributing to this important national study.

Sincerely,

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.



**NEXT BIRTHDAY COVER LETTER – FOURTH MAILING**

Dear {City} Resident:

We recently invited you to participate in an important national survey sponsored by the U.S. Department of Health and Human Services (HHS), the Health Information National Trends Survey (HINTS). By completing the HINTS survey, you will help us keep you and your family better informed on the health issues that matter to you.

If you already sent back your survey and it crossed in the mail with this letter, thank you for the time you took to help make this study a success.

If you have not yet had a chance to complete the HINTS survey, we understand that it is sometimes difficult to find the time to participate in a study like HINTS. In order to reduce the amount of time it will take, we have enclosed a shorter version of the survey. This shorter version focuses on issues that tell us how individuals like you find and use health information.

**In order to make sure we get responses from a random sample of people, we ask that the adult in your household with the next birthday complete and return this questionnaire in the next two weeks.**

Additional information about HINTS is available at: [hints.cancer.gov](http://hints.cancer.gov). If you have any questions or would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Thank you in advance for contributing to this important national study.

Sincerely,

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Si prefiere recibir la encuesta en español, por favor llame al 1-888-738-6812.**

The Health Information National Trends Survey is authorized under 42 USC, Section 285A.

**Appendix B:**  
**Spanish Cover Letters**



**NEXT BIRTHDAY COVER LETTER – SECOND MAILING**

Estimado residente de {City}:

Recientemente lo invitamos a participar en una importante encuesta nacional patrocinada por el Departamento de Salud y Servicios Humanos de Estados Unidos. El objetivo de la Encuesta Nacional de Tendencias de Información sobre la Salud (HINTS, por sus siglas en inglés) es averiguar cuál es la información sobre la salud que las personas quieren saber y dónde van a buscarla. Sus respuestas nos ayudarán a mantenerlo mejor informado a usted, a sus familiares y a los miembros de la comunidad sobre los temas de salud que les interesan.

Aún no hemos recibido su cuestionario completado. Para poder estar seguros de que HINTS provea información acertada, necesitamos que todos los hogares invitados a participar en la encuesta este año, la completen. Si usted ya nos envió de regreso su encuesta y se cruzó con esta carta en el correo, le agradecemos por el tiempo que se tomó para contribuir al éxito de este estudio. En caso que su cuestionario se haya extraviado, adjuntamos una copia adicional.

**Para asegurarnos de obtener respuestas que contengan un muestreo aleatorio de la población, le pedimos que el adulto en su hogar con el próximo cumpleaños, complete y devuelva este cuestionario en las próximas dos semanas.**

Usted podrá encontrar más información sobre HINTS en el sitio web [www.HINTS.cancer.gov](http://www.HINTS.cancer.gov). Si usted tiene preguntas o le gustaría completar esta encuesta en otro idioma distinto al inglés o español, llame a Westat al número libre de cargo, 1-888-738-6812.

Gracias de antemano por contribuir al éxito de este importante estudio nacional.

Atentamente,

Bradford W. Hesse, Ph. D.

Oficial del Proyecto HINTS  
Institutos Nacionales de la Salud  
Departamento de Salud y Servicios Humanos de EE.UU.



**ALL ADULT COVER LETTER (only sent by request)**

Estimado residente de {City}:

Recientemente lo invitamos a participar en una importante encuesta nacional patrocinada por el Departamento de Salud y Servicios Humanos de Estados Unidos. El objetivo de la Encuesta Nacional de Tendencias de Información sobre la Salud (HINTS, por sus siglas en inglés) es averiguar cuál es la información sobre la salud que las personas quieren saber y dónde van a buscarla. Sus respuestas nos ayudarán a mantenerlo mejor informado a usted, a sus familiares y a los miembros de la comunidad sobre los temas de salud que les interesan.

Aún no hemos recibido su cuestionario completado. Para poder estar seguros de que HINTS provea información acertada, necesitamos que todos los hogares invitados a participar en la encuesta este año, la completen. Si usted ya nos envió de regreso su encuesta y se cruzó con esta carta en el correo, le agradecemos por el tiempo que se tomó para contribuir al éxito de este estudio. En caso que su cuestionario se haya extraviado, adjuntamos una copia adicional.

**Pedimos que cada adulto en esta residencia complete un cuestionario y nos lo devuelva en las próximas dos semanas.**

Usted podrá encontrar más información sobre HINTS en el sitio web [hints.cancer.gov](https://hints.cancer.gov). Si usted tiene preguntas, le gustaría pedir copias adicionales de esta encuesta, o le gustaría completar esta encuesta en otro idioma distinto al inglés o español, llame a Westat al número libre de cargo, 1-888-738-6812.

Gracias de antemano por contribuir al éxito de este importante estudio nacional.

Atentamente,

Bradford W. Hesse, Ph. D.

Oficial del Proyecto HINTS  
Institutos Nacionales de la Salud  
Departamento de Salud y Servicios Humanos de EE.UU.



**NEXT BIRTHDAY COVER LETTER – THIRD MAILING**

Estimado residente de {City}:

Recientemente lo invitamos a participar en una importante encuesta nacional patrocinada por el Departamento de Salud y Servicios Humanos de Estados Unidos, la Encuesta Nacional de Tendencias de Información sobre la Salud (HINTS, por sus siglas en inglés). El completar esta encuesta nos ayudará a mantenerlos mejor informados en asuntos de salud que usted y su familia consideran importantes.

Si usted ya envió de regreso su encuesta y se cruzó con esta carta en el correo, le agradecemos por el tiempo que se tomó para contribuir al éxito de este estudio.

Si aún no ha tenido la oportunidad de completar la encuesta, comprendemos que a veces es difícil encontrar el tiempo para participar en un estudio como HINTS. Para reducir la cantidad de tiempo que va a tomarle, hemos incluido una versión más corta de la encuesta. Esta versión corta se concentra en asuntos que nos informan la manera como personas como usted buscan y utilizan información sobre la salud.

**Para asegurarnos de obtener respuestas que contengan un muestreo aleatorio de la población, le pedimos que el adulto en su hogar con el próximo cumpleaños, complete y devuelva este cuestionario en las próximas dos semanas.**

Usted podrá encontrar más información sobre HINTS en el sitio web [hints.cancer.gov](https://hints.cancer.gov). Si usted tiene preguntas o le gustaría completar esta encuesta en otro idioma distinto al inglés o español, llame a Westat al número libre de cargo, 1-888-738-6812

Gracias de antemano por contribuir con este importante estudio nacional.

Atentamente,

Bradford W. Hesse, PhD.

Oficial del Proyecto HINTS  
Institutos Nacionales de la Salud  
Departamento de Salud y Servicios Humanos de EE.UU.



**NEXT BIRTHDAY COVER LETTER – FOURTH MAILING**

Estimado residente de {City}:

Recientemente lo invitamos a participar en una importante encuesta nacional patrocinada por el Departamento de Salud y Servicios Humanos de Estados Unidos, la Encuesta Nacional de Tendencias de Información sobre la Salud (HINTS, por sus siglas en inglés). El completar esta encuesta nos ayudará a mantenerlos mejor informados en asuntos de salud que usted y su familia consideran importantes.

Si usted ya envió de regreso su encuesta y se cruzó con esta carta en el correo, le agradecemos por el tiempo que se tomó para contribuir al éxito de este estudio.

Si aún no ha tenido la oportunidad de completar la encuesta, comprendemos que a veces es difícil encontrar el tiempo para participar en un estudio como HINTS. Para reducir la cantidad de tiempo que va a tomarle, hemos incluido una versión más corta de la encuesta. Esta versión corta se concentra en asuntos que nos informan la manera como personas como usted buscan y utilizan información sobre la salud.

**Para asegurarnos de obtener respuestas que contengan un muestreo aleatorio de la población, le pedimos que el adulto en su hogar con el próximo cumpleaños, complete y devuelva este cuestionario en las próximas dos semanas.**

Usted podrá encontrar más información sobre HINTS en el sitio web [hints.cancer.gov](https://hints.cancer.gov). Si usted tiene preguntas o le gustaría completar esta encuesta en otro idioma distinto al inglés o español, llame a Westat al número libre de cargo, 1-888-738-6812

Gracias de antemano por contribuir con este importante estudio nacional.

Atentamente,

Bradford W. Hesse, PhD.

Oficial del Proyecto HINTS  
Institutos Nacionales de la Salud  
Departamento de Salud y Servicios Humanos de EE.UU.

**Appendix C:**  
**Frequently Asked Questions (FAQs)**  
**English and Spanish**

**Some Frequently Asked Questions  
about the  
Health Information National Trends Survey**

**Q: What is the study about? What kind of questions do you ask?**

A: The study concerns health and how people receive health information. For example, we will ask how you usually get information about how to stay healthy, the sources of information you most trust, and how you might like to get such information in the future. We will also ask about your beliefs on what contributes to good health, how best to prevent cancer, your participation in various health-related activities, and related topics.

**Q: How will the study results be used? What will be done with my information?**

A: Findings will help the U.S. Department of Health and Human Services promote good health and prevent disease by determining ways of better communicating accurate health information to Americans.

**Q: How did you get my address?**

A: Your address was randomly selected from among all of the known home addresses in the nation. It was selected using scientific sampling methods.

**Q: Why should I take part in this study? Do I have to do this?**

A: Your participation is voluntary, and you may refuse to answer any questions or withdraw from the study at any time. However, your answers are very important to the success of this study and will represent thousands of others. Getting an answer from all the households chosen for the study is the best way to make sure the study results reflect the thoughts and opinions of all Americans.

**Q: Will my answers to the survey be kept private?**

A: Yes. Your answers will be kept private under the Privacy Act. Your answers cannot be connected to your name or any other information that could identify you or your household, to the extent provided by law. The completed questionnaires will be stored in a separate file with restricted access. Both the paper and electronic versions of the information will be destroyed shortly after the research is finalized.

**Q: How long will it take to answer the questions?**

A: About 20 to 30 minutes.

**Q: Who is sponsoring the study? Is this study approved by the Federal Government?**

A: The study is sponsored by the U.S. Department of Health and Human Services. The study has been approved by the Office of Management and Budget (OMB), the office that reviews all federally-sponsored surveys. The OMB approval number assigned to this study is 0925-0538.

**Q: Who is Westat?**

A: Westat is a research company located in Rockville, Maryland. Westat is conducting this survey under contract to the U.S. Department of Health and Human Services.

## **Preguntas Frecuentes**

### **Encuesta Nacional de Tendencias de Información sobre la Salud**

**P: ¿De qué se trata el estudio? ¿Qué tipo de preguntas contiene?**

R: El estudio trata sobre la salud y la manera en que las personas reciben información sobre la salud. Por ejemplo, le preguntaremos cómo obtiene normalmente información sobre cómo mantenerse saludable, el tipo de información en la que más confía y cómo le gustaría obtener dicha información en el futuro. También le preguntaremos sobre lo que cree que contribuye a la buena salud, cómo prevenir mejor el cáncer y su participación en varias actividades afines.

**P: ¿Cómo se utilizarán los resultados del estudio? ¿Qué se hará con mi información?**

R: Los hallazgos ayudarán al Departamento de Salud y Servicios Humanos de EE.UU. a fomentar la buena salud y prevenir las enfermedades mediante la determinación de formas de comunicar mejor la información sobre la salud a los estadounidenses.

**P: ¿Cómo obtuvieron mi dirección?**

R: Su dirección fue seleccionada al azar entre todas las direcciones conocidas en la nación usando métodos científicos de muestreo.

**P: ¿Por qué debo participar en este estudio? ¿Es obligatorio hacerlo?**

R: Su participación es voluntaria y usted puede rehusarse a contestar cualquiera de las preguntas o retirarse del estudio en cualquier momento. Sin embargo, sus respuestas son muy importantes para el éxito de este estudio y representan a miles de personas. El obtener respuesta de todos los hogares escogidos para este estudio es la mejor manera de asegurar que éste refleje los pensamientos y opiniones de todos los estadounidenses.

**P: ¿Se mantendrá la privacidad de mis respuestas a la encuesta?**

R: Sí. Se mantendrá la privacidad de sus respuestas en virtud de la Ley de Privacidad. Sus respuestas no pueden asociarse a su nombre ni a ninguna otra información que podría identificarlo a usted o a su hogar en la medida de lo permisible por ley. Los cuestionarios completos se almacenarán en un archivo separado con acceso restringido. Las versiones impresas y electrónicas de la información se destruirán poco después de la finalización de la encuesta.

**P: ¿Cuánto tiempo tomará responder las preguntas?**

R: Cerca de 20 a 30 minutos.

**P: ¿Quién patrocina el estudio? ¿Está este estudio aprobado por el Gobierno Federal?**

R: El estudio es patrocinado por el Departamento de Salud y Servicios Humanos de EE.UU. El estudio ha sido aprobado por la Oficina de Administración y Presupuesto (OMB, por sus siglas en inglés). Esta oficina revisa todas las encuestas patrocinadas federalmente. El número de aprobación asignado por la OMB para este estudio es 0925-0538.

**P: ¿Quién es Westat?**

R: Westat es una compañía de estudios de investigación ubicada en Rockville, Maryland. Westat realiza esta encuesta en virtud de un contrato con el Departamento de Salud y Servicios Humanos de EE.UU.

**Appendix D:**  
**Targeted Cover Letter for All Adult Household**



Dear {City} Resident:

Recently, the U.S. Department of Health and Human Services mailed two questionnaires to your household for the Health Information National Trends Survey (HINTS). As of today we have received {one/two/three} completed questionnaire{s} from your household. Thank you for returning {that/those} questionnaire{s}.

The success of the survey depends on receiving a completed questionnaire from each adult 18 or older in the household. That is the best way to make sure our survey results reflect all adults in the U.S. We are writing to request that the adults who have not yet completed a questionnaire please do so in the next week.

According to the information provided on the completed questionnaire{s}, we still need a response from:

- Male/Female, age (xx)
- (additional as needed)

To understand the health topics important to people from communities across the U.S., HINTS needs to collect the current views and experiences of all the people we're trying to contact.

Westat, a research firm, is administering the survey. Your answers will be kept private under the Privacy Act. More information about HINTS can be found at: [hints.cancer.gov](http://hints.cancer.gov).

There are {one/two/three} questionnaires enclosed. If you or another adult completing one of these questionnaires has any questions or if you would like to complete this survey in a language other than English or Spanish, please call Westat toll free at 1-888-738-6805.

Sincerely,

Bradford W. Hesse, Ph.D.

HINTS Project Officer  
National Institutes of Health  
U.S. Dept of Health and Human Services

**Appendix E:**  
**Variable Values and Edits**

**Missing value definitions:**

Values identifying types of nonresponse or indeterminate responses:

- -1 = Valid skips or appropriately missing data following a dependent question (correctly skipped). Example: If SeekHealthInfo=2 'no' and WhereSeekHealthInfo was missing, WhereSeekHealthInfo was assigned the value -1.
- -2 = Question was answered, but respondent should not have answered the question. The question was answered in error by the respondent. Example: If SeekHealthInfo=2 'no' and WhereSeekHealthInfo was not missing, WhereSeekHealthInfo was assigned the value -2.
- -4 = Question was answered, but data was removed because the entry of the number or character could not be determined (e.g. unreadable or non-conforming numeric response).
- -5 = Respondent selected more response options than appropriate for the question. Example: If TrustRadio had values 3 'a little' and 2 'some', TrustRadio was assigned the value -5. In cases where both -2 and -5 values could be assigned, the -2 value was assigned.
- -6 = Missing data in variables following a missing filter question. Example: If filter question (e.g., SeekHealthInfo) was missing and variables up until the next question respondents who skipped were instructed to answer (e.g. ConfidentGetHealthInf) were missing (e.g., LookElsewhere= missing) variables with missing values were assigned the value -6.
- -9 = Missing data. Not ascertained. Question should have been answered, but no response was recorded. Example: If SeekHealthInfo was missing, it was assigned the value -9.

Variable	Editing Rule	Description of Rule
AdultsInHH	Recoding initial filter/skip question	The value of the following response, MailHHAdults, determined how missing responses to AdultsInHH were re-assigned. As an example, if AdultsInHH was missing and MailHHAdults initially had value 1 (adult in household) then AdultsInHH was assigned the value 2 'no' (indicating not more than 1 adult in the household) and MailHHAdults was assigned the 'missing value' -2 (answered inappropriately). If AdultsInHH was missing and MailHHAdults had value 2 (or greater) then AdultsInHH was assigned the value 1 'yes' (indicating more than 1 adult in the household) and the value for MailHHAdults was retained.
SeekHealthInfo	Recoding filter/skip questions	For these filter questions (questions containing a skip instruction associated with the particular response that was selected), response patterns following the question were examined if the filter question was not answered. The 'yes'
UseInternet		

Variable	Editing Rule	Description of Rule
FreqGoProvider		value (in the majority of cases where a 'yes' response instructed a respondent to continue answering the subsequent questions) was substituted for the missing filter question when any of the subsequent questions were answered. Similarly (when a 'no' response instructed a respondent to skip subsequent questions), the 'no' value was substituted for the missing filter question when all of the subsequent questions that a 'no' response would have directed the respondent to skip were left unanswered and the respondent answered the next applicable question all respondents were supposed to answer.
Smoke100		
GenderC		
EverHadCancer		
BuyOTCDrugs		
BuyRxDrugs		
VisitFDAWebsite		
FDASite_Other		
BornInUSA		
WhereSeekHealthInfo_IMP	Imputation for multiple responses	Imputation was carried out when multiple responses were selected, resulting in one unique response for these "mark only one" variables. Respondent's multiple answers were replaced with a single imputed answer that had the same distribution over the multiple answers as occurred in the single-answer responses. Missing values were not imputed for. Flags (indicated by suffix '_IFlag') indicate which values were imputed and which values were original.
StrongNeedHealthInfo_IMP		
WhoLookingFor	Edits for multiple responses	Multiple responses (e.g., 'myself', 'someone else') were recoded to the logically applicable third response option ('both myself and someone else').
InternetHealthInfoSite	Recoding filter question	This variable was treated differently than regular filter/skip items, in which subsequent responses determined if the filter question was re-assigned to have a 'yes' or 'no' value. For this variable, missing responses in the following InternetHealthInfoSite_OS 'other specify' field did not result in this variable being assigned the 'no' response, although non-missing responses in InternetHealthInfoSite_OS did result in this variable being assigned the 'yes' response.
Internet_DialUp	Recoding missing responses for items with forced-choice response formats	Respondents were asked to select 'yes' or 'no' to a series of sub-items, allowing them to select as many responses as would apply. These 'forced-choice' response formats sometimes result in respondents indicating which sub-items apply to them by selecting the 'yes' response option for some and leaving the others unanswered. To allow the data to reflect this practice, if respondents did not check a 'no' response option for any sub-item in the question, the sub-items that were missing a response were re-set to 'no.' However, if a respondent, in addition to leaving other sub-items unanswered, did select a 'no' response option for at least one sub-item, the unanswered sub-items were not
Internet_BroadBnd		
Internet_Cell		
Internet_WiFi		
IntRsn_InfQuitSmoking		
IntRsn_BuyMedicine		
IntRsn_SupportGroup		
IntRsn_TalkDoctor		
IntRsn_DietWebsite		

Variable	Editing Rule	Description of Rule
IntRsn_HCProviderSearch		assumed to be 'no' responses and instead remained missing.
IntRsn_PDADownload		
IntRsn_SocNetworkSite		
IntRsn_Blog		
IntRsn_TrackedPHR		
IntRsn_HealthInfSE		
HCCoverage_IHS		
HCCoverage_Insurance		
HCCoverage_Medicaid		
HCCoverage_Medicare		
HCCoverage_Private		
HCCoverage_Tricare		
HCCoverage_VA		
RxDrugRecall_StopTaking		
RxDrugRecall_KeepUsing		
RxDrugRecall_Ignore		
RxDrugRecall_OnGuard		
RxDrugRecall_VisitWebsite		
RxDrugRecall_ContactManuf		
RxDrugRecall_ContactDoctor		
RxDrugRecall_ContactPharm		
RxDrugRecall_Unsure		
MPRecall_StopUsing		
MPRecall_ContactDoctor		
MPRecall_ContactManuf		
MPRecall_Replace		
MPRecall_KeepUsing		
MPRecall_NoChange		
MPRecall_Unsure		
HCCoverage	Derived value	A derived variable was created using the various health insurance/health coverage responses (e.g., HCCoverage_Insurance, HCCoverage_Other) to indicate whether a respondent had any form of health insurance/health coverage.
HCCoverage_I	Imputation for	Missing values were imputed for variables that were used

Variable	Editing Rule	Description of Rule
GenderC_I	missing responses	in the process of assigning weights. Flags (indicated by suffix ‘_IFlag’) indicate which values were imputed and which values were original.
EverHadCancer_I		
Age_I		
MaritalStatus_I		
Education_I		
Hispanic_I		
Race_Cat_I		
TimesModerateExercise	Recoding filter/skip questions	For these filter questions (questions containing a skip instruction associated with the particular response that was selected), response patterns following the question were examined if the filter question was not answered. The 0 value (e.g., ‘none’ or ‘0 days’) was substituted for the missing filter question if all of the subsequent question that a 0 response directed the respondent to skip were left unanswered and the respondent answered the next applicable question all respondents were supposed to answer. However, missing values were not substituted with other values if the filter question was not answered but the follow-up question was answered.
DrinksPerWeek		
Height_Feet	Edits for implausible values	The rules that were applied minimized the number of out-of-range values by accounting for response measurements in incorrect boxes, responses using metric, responses using only one unit of measurement and other response errors. A list of the rules applied to clean the variables is included in ‘Appendix E – Data Editing Procedures.’
Height_Inches		
Cancer_cat	Summarized distribution of ‘mark all that apply’ responses	A variable was created to indicate each response selection a respondent made for these ‘mark all that apply’ variables. The derived variable with the suffix ‘_cat’ summarized the response selected or indicated that multiple responses were selected.
Race_cat		
Employed	Derived variables for multiple responses	For the variable OccupationStatus, derived variables were created to indicate each response selected, showing the unique response for respondents selecting one occupation, and showing each response for respondents who did not follow the ‘mark only one’ response instruction.
Unemployed		
Homemaker		
Student		
Retired		
Disabled		
OtherOCC		
Education	Edits for multiple responses	The highest order (e.g., education level or income range) was taken when multiple responses were selected.
IncomeRanges		

Variable	Editing Rule	Description of Rule
R_HHAdults	Derived variable	Responses to questions asking about household size as well as other information about the household (e.g., number of questionnaires returned) were compiled into a derived measure that best represented the number of adults in the household.
HHAdults_Num	Imputation for zero and missing responses	Missing values were imputed for the derived count of household adults when the derived variable had values of zero or missing. A flag (indicated by suffix '_IFlag') indicates which values were imputed and which values were original.
QDisp	Derived variable	A variable was created to indicate the proportion of items respondents answered in the first two sections. This was used to determine incompletely-filled out questionnaires.

**Appendix F:**  
**Data Editing Procedures**

### Rules Applied to Edit Height Variables:

1. If HEIGHT\_Feet was 0 or missing and HEIGHT\_Inches>48 and HEIGHT\_Inches<=60, then the first digit was taken as the feet value and the second digit was taken as the inches value (to correct for respondents expressing both feet and inches in the inches box).
2. If HEIGHT\_Feet was 0 or missing and HEIGHT\_Inches>61 and HEIGHT\_Inches<=83, then the inches value was converted to its feet-and-inches equivalent (to correct for respondents expressing height in inches, resulting in heights from 5'1" to 6'11").
3. If HEIGHT\_Feet was 1 and HEIGHT\_Inches>=3 and HEIGHT\_Inches<=9 (or HEIGHT\_Inches>=30 and HEIGHT\_Inches<=90) then this metric value was converted to feet-and-inches (to correct for respondents using meters and tenths and hundredths of a meter to express height).
4. If HEIGHT\_Feet>3 and HEIGHT\_Feet<7 and HEIGHT\_Inches = 20, 30, etc. thru 90 then the trailing 0 was removed
5. If HEIGHT\_Feet>3 and HEIGHT\_Feet<7 and HEIGHT\_Inches = 15, 25, etc. thru 95 then the trailing 5 was removed (to correct for respondents expressing values in tenths of an inch).
6. If HEIGHT\_Feet>3 and HEIGHT\_Feet<7 and HEIGHT\_Inches = 12, 23, 34, 45 etc. thru 89 then the first digit was taken (to correct for respondents giving an inch value as a range, e.g., 1-2 or 8-9 inches).
7. If HEIGHT\_Feet>3 and HEIGHT\_Feet<7 and HEIGHT\_Inches = a two digit value whereby the first digit equaled the feet value the second digit was taken as the inches value (to correct for respondents expressing the height in inches as well as in feet, e.g., 5'58" resulted in value 5'8")
8. If HEIGHT\_Feet>6 and HEIGHT\_Feet<12 and HEIGHT\_Inches>3 and HEIGHT\_Inches<7, then the values were switched (to correct for respondents putting measurements in the wrong boxes, resulting in edited values from 4'7" to <7 feet)

### Deriving and Imputing Measure of Household Adults:

A derived measure of the number of household adults was created for households that returned questionnaires. This measure involved recoding questionnaire responses and making comparisons with the operational information collected on the households (i.e., the number of questionnaires returned).

#### Editing Steps:

1. Create composite from two variables (the 'raw' and 'edited' versions of **MailHHAdults**) to give count of adults in the household (named **RS\_HHAdults**):
  - a. First, take **MailHHAdults** values as initially entered (e.g., if R says 'no' there are not >1 adults in the household but initially entered something for the 'raw' value of **MailHHAdults**, take the initial **MailHHAdults** entry; if R says 'yes' there are >1 adults in the household, but **MailHHAdults**=1, retain 1)
  - b. Second, if **MailHHAdults** = -1/skipped appropriately, then set **RS\_HHAdults** = 1
2. Create another indicator of household adults (named **Demo\_HHAdults**) from non-missing demographic variables. **Demo\_HHAdults** = (**TotalHousehold** - **ChildrenInHH**).

3. Create edited measure (named **DemoRS\_HHAdults**) by substituting **RS\_HHAdults** value for any **Demo\_HHAdults** value that is negative, 0, or 11+, if the **RS\_HHAdults** value is between 1 and 10 inclusive.
4. Create count of number of **HHAdultAge[#]** values between 18 and 115 inclusive that are entered in the roster: (named **Roster\_HHAdults**)
5. Create final household adult size measure (**R\_HHAdults**)
  - a. If **DemoRS\_HHAdults** is still negative or 0 after the step 3 substitution, further substitute **Roster\_HHAdults** values if they are positive
  - b. If **Roster\_HHAdults** value is larger than **DemoRS\_HHAdults** value, substitute the **Roster\_HHAdults** value
6. Reconcile differences in **R\_HHAdults** values within households:
  - a. Use the number of returns from the household as the **R\_HHAdults** value if it is greater than **R\_HHAdults**
  - b. If another measure of the number of adults (e.g. **RS\_HHAdults**) is consistent among household respondents, substitute that value for **R\_HHAdults**
  - c. If all interim measures of adults within a household differ (e.g. **Demo\_HHAdults**, **RS\_HHAdults**) take the largest **R\_HHAdults** value and apply it to all in the household
7. Reconcile differences between **R\_HHAdults** and disposition codes, keeping **R\_HHAdults** value unless:
  - a. The **R\_HHAdults** value is logically inconsistent with the number of questionnaires returned from the household (e.g., there number of returned questionnaires was greater than the **R\_HHAdults** value)
  - b. The **R\_HHAdults** value was edited after review of questionnaire data identified likely respondent/scanning error (e.g., entry of no children in the household but someone under 18 entered on the roster, an entry of 1 person in the household scanned as 10)
8. Edit non-zero values of **R\_HHAdults** that are inconsistent with the number of questionnaires returned from the household (e.g., **R\_HHAdults**=2 but 3 questionnaires received from the household)

Imputation for the remaining values of zero or missing for **R\_HHAdults** involved replacing these values with the average number of adults in responding households with non-zero or non-missing values of **R\_HHAdults**, resulting in the variable **HHAdults\_Num**.