DESIGN AND METHODOLOGY OF THE

ARIZONA HEALTH SURVEY

OCTOBER 2008

This report describes how data were collected for the Arizona Health Survey. It was a telephone survey of adults in households with landline telephone numbers using a random digit dialing (RDD) sample. The sample was geographically stratified to represent Maricopa County and the remainder of Arizona. In Maricopa County, children and adolescents were also sampled when present in a household. All data were collected using a computer-assisted telephone interviewing (CATI) system, with interviewing in English and Spanish. The data were weighted to represent the Arizona household population.

This report was prepared for the Arizona Health Survey by Sherman Edwards, Ismael Flores-Cervantes, and Susan Fraser of Westat.

The Arizona Health Survey is a project of St. Luke's Health Initiatives. We gratefully acknowledge financial support from the Virginia G. Piper Charitable Trust, the Arizona Community Foundation and Valley of the Sun United Way, and the intellectual contributions of numerous members of the nonprofit, government and research communities.

Suggested citation:

Arizona Health Survey. *Design and Methodology of the Arizona Health Survey*. Phoenix AZ: St. Luke's Health Initiatives, 2008.

CONTENTS

Contents (continued)

Contents (continued)

List of Tables

List of Tables (continued)

List of Figures

List of Exhibits

AHS DESIGN AND METHODOLOGY – EXECUTIVE SUMMARY

The Arizona Health Survey (AHS), sponsored by St. Luke's Health Initiatives (SLHI), is a population-based random-digit dial telephone survey of Arizona's population conducted in the first half of 2008. It was designed to collect data on individual indicators of health status, health care access, health-related behaviors and various demographic and social/environmental factors related to health. Results will be used to inform and improve public policy and community health/health care program planning decisions at the local, regional and state levels. In addition, it was designed to enable service providers and funders to:

- plan resource allocation and target intervention activities to increase access to care for high-risk, underserved and uninsured populations;
- determine community strengths, resources, barriers and needs;
- increase understanding of attitudes toward prevention and utilization; and,
- establish a mechanism by which to evaluate efforts to improve community health and quality of life.

The AHS sample is representative of Arizona's non-institutionalized population living in households with landline telephones.

This report describes the AHS sample design, data collection and processing procedures, and weighting and variance estimation. These details are summarized here.

To achieve its objectives, AHS employed a multi-stage sample design, described in Chapter 1. Landline residential telephone numbers were selected within two geographic strata – Maricopa County and the remainder of Arizona. Within each household, one adult (age 18 and over) respondent was randomly selected. In those households in Maricopa County with adolescents (ages 12-17) and/or children (under age 12), one adolescent and one child were randomly selected; the adolescent was interviewed directly, and the adult most knowledgeable about the child's health completed the child interview. The sample was selected so as to complete 3,000 interviews with adults in Maricopa County and 1,000 adult interviews in the remainder of Arizona. Table ES-1 shows the number of completed interviews by type and stratum (as reported by the respondent).

Table ES-1. Number of completed AHS interviews by self-reported location and instrument

	Adult	Child	Adolescent
Total	4,196	635	203
Maricopa County	3,130	635	203
Remainder of Arizona	1,066	N/A	N/A

Table ES-2 at the end of this summary shows the major topic areas for each of the three survey instruments (adult, child, and adolescent). Chapter 2 describes the structure and content of the survey instruments. The average adult interview took about 30 minutes to complete. The average child and adolescent interviews took about 15 minutes and 19 minutes, respectively.

Westat, a private firm that specializes in statistical research and large-scale sample surveys, conducted AHS data collection under contract with SLHI. Interviews were conducted in English and Spanish using Westat's computer-assisted telephone interviewing (CATI) system. Interviewer recruitment and training are described in Chapter 3, and Chapter 4 presents details of data collection procedures and results.

Adults who completed at least approximately 80 percent of the questionnaire (i.e., through Section K (on employment, income, poverty status, and food security), after all follow-up attempts were exhausted to complete the full questionnaire, were counted as "complete." Proxy interviews were allowed for frail and ill persons over the age of 65 who were unable to complete the extended adult interview. For 35 elderly adults, a proxy interview was completed by either a spouse/partner or adult child. A reduced questionnaire, with questions identified as appropriate for a proxy respondent, was administered. (Note: questions not administered in proxy interviews are given a value of "-2" in the data files.)

The overall AHS response rate is a composite of the screener completion rate (i.e., success in introducing the survey to a household and randomly selecting an adult to be interviewed) and the extended interview completion rate (i.e., success in getting one or more selected persons to complete the extended interview). To maximize the response rate, especially at the screener stage, an advance letter in English and Spanish was mailed to all sampled telephone numbers for which an address could be obtained from directory services.

The AHS screener completion rate was 36.6 percent, and was higher for households that were sent the advance letter. The extended interview completion rate varied across the adult, child and adolescent interviews. The adult extended completion rate was 53.7 percent, resulting in an overall adult response rate of 19.2 percent for adults. The rates were slightly higher in the remainder of Arizona than in Maricopa County.

Westat conducted a variety of data preparation, coding, and cleaning operations to enhance the quality and utility of the survey data. These included resolving problems identified during data collection, coding text strings, including race and ethnicity, geocoding, and converting the survey responses to SAS and Excel files. These operations are described in Chapter 5.

To produce population estimates from AHS data, weights are applied to the sample data to compensate for the probability of selection and a variety of other factors, some directly resulting from the design and administration of the survey. The sample is weighted to represent the non-institutionalized population for each sampling stratum and statewide. AHS weighting procedures accomplish the following objectives:

- Compensate for differential probabilities of selection for households and persons;
- Reduce biases occurring because nonrespondents may have different characteristics than respondents;
- Adjust, to the extent possible, for undercoverage in the sampling frame and in the conduct of the survey; and
- Reduce the variance of the estimates by using auxiliary information.

As part of the weighting process, a household weight was created for all households that completed the screener interview. This household weight is the product of the "base weight" (the inverse of the probability of selection of the telephone number) and a variety of adjustment factors. The household weight is used to compute a person-level weight, which includes adjustments for the within-household sampling of persons and nonresponse. The final step is to adjust the person-level weight using a raking method so that the AHS estimates are consistent with population control totals. Raking is an iterative procedure that forces the AHS weights to sum to known totals from an independent data source. The sources used were 2007 Arizona Department of Commerce Population Estimates, 2008 Arizona Department of Commerce Projections (State of Arizona, Department of Commerce, 2006, 2006b), and the 2006 American

Community Survey estimates for Arizona. The procedure requires iteration to make sure all the control totals, or raking dimensions, are simultaneously satisfied within a specified tolerance.

Missing values in the AHS data files were replaced for a handful of variables used in the weighting process, using random allocation and hot-deck methods.

AHS weighting procedures are described in detail in Chapter 6.

Table ES-2. AHS survey topic areas by instrument

Health status	**Adult**	**Teen**	**Child**
General health status	✓	✓	✓
Height and weight		✓	✓
Days missed from school due to health problems		✓	✓
Health conditions	**Adult**	**Teen**	**Child**
Asthma	✓	✓	✓
Diabetes	✓	✓	
Heart disease, high blood pressure	✓		
Gastrointestinal disorders	✓		
Bi-polar disorder, anxiety disorder, depression	✓		
Falls (elderly)	✓		
Developmental disorders			✓
Mental health	**Adult**	**Teen**	**Child**
Mental health status	✓	✓	✓
Emotional functioning	✓	✓	✓
Resiliency, coping	✓		
Perceived need, use of mental health services	✓	✓	
Reasons for not seeking treatment	✓		
Intimate partner violence	✓	✓	
Childhood experiences	✓		
Parent's behavioral and developmental concerns			✓
Behavioral and developmental concerns of school, doctor			✓
Health behaviors	**Adult**	**Teen**	**Child**
Dietary intake	✓	✓	✓
Physical activity and exercise	✓	✓	✓
Sedentary time		✓	
Parental influence over diet and exercise			✓
Use of nonprescription drugs and dietary supplements	✓		
Flu shot	✓		
Alcohol and tobacco use	✓	✓	
Illegal drug use		✓	
Sexual behavior, birth control practices		✓	
Dental health	**Adult**	**Teen**	**Child**
Last dental visit		✓	✓
Not getting needed care		✓	✓
Days missed from school due to dental problems		✓	✓
Usual source of dental care		✓	✓

Table ES-2. AHS survey topic areas by instrument (Continued)

Access to and use of health care	**Adult**	**Teen**	**Child**
Usual source of care, visits to medical doctor	✓	✓	✓
Emergency room use	✓		✓
Unmet needs for care or prescriptions	✓	✓	✓
Racial/ethnic discrimination in health care	✓		
Communication with doctor	✓	✓	✓
Health insurance	**Adult**	**Teen**	**Child**
Current health insurance coverage and source	✓	✓	✓
Spouse's health insurance coverage	✓	✓	✓
Coverage of prescription drugs, dental and behavioral health services	✓	✓	✓
Coverage over past 12 months	✓	✓	✓
Medical debt	✓		
Housing and Neighborhood	**Adult**	**Teen**	**Child**
Type of housing and tenure	✓		
Neighborhood safety	✓	✓	✓
Characteristics of neighbors	✓		
Use of parks		✓	
Volunteer service	✓	✓	
Parental involvement	**Adult**	**Teen**	**Child**
Parental presence after school		✓	
Marital status of parents		✓	
Support for parenting	✓		
Child's activities with family			✓
Child care and school attendance	**Adult**	**Teen**	**Child**
Current child care arrangements			✓
Preschool/school attendance, name of school		✓	✓
Employment	**Adult**	**Teen**	**Child**
Employment status, spouse's employment status	✓		
Work in last week	✓		
Hours worked at all jobs	✓		

Table ES-2. AHS survey topic areas by instrument (Continued)

Income	Adult	Teen	Child
Respondent's and spouse's earnings last month before taxes	✓		
Household income (annual before taxes)	✓		
Number of persons supported by household income	✓		
Receipt of Social Security disability, SSI	✓		
Participation in TANF	✓		
Participation in food stamps, WIC			✓
Respondent characteristics	**Adult**	**Teen**	**Child**
Age, gender, height, weight	✓	✓	✓
Race and ethnicity	✓	✓	✓
Marital status	✓		
Education	✓		
First language, English proficiency	✓		
Languages spoken at home			✓
Country of birth	✓	✓	✓

1. SAMPLE DESIGN

This chapter describes the sampling methods used in the 2008 Arizona Health Survey (AHS). AHS consisted of a telephone random-digit-dialing (RDD) sample drawn using a list-assisted approach. The first section describes list-assisted RDD sampling and the procedures implemented in AHS to save costs by reducing the number of calls to ineligible telephone numbers. The second section describes two noncoverage problems that affect telephone surveys in general and how these were addressed in AHS. The remaining sections review the AHS sample design and describe the sampling procedures used to select households and persons within households.

1.1 List-Assisted Random-Digit-Dial Sampling

List-assisted RDD sampling is currently the standard method of choice for telephone surveys. This method results in an unclustered sample that has good operational features (Tucker, Lepkowski, and Piekarski, 2002). In list-assisted sampling, the set of all telephone numbers in operating telephone prefixes is composed of 100-banks, each containing the 100 telephone numbers with the same first eight digits. All 100-banks with at least one residential number listed in a published telephone directory are used to create the sampling frame. A simple random or a systematic sample of telephone numbers is selected from this frame.

1.1.1 Methods to Increase the Efficiency of Data Collection

When using a list-assisted approach, special data collection procedures are often implemented before data collection to reduce costs and to increase efficiency. Three of these methods were implemented in AHS. The first method is the use of tritone purges and directory matching for removal of unproductive numbers (i.e., business and nonworking numbers). The procedure used in AHS, called Comprehensive Screening Service (CSS), is offered by Marketing Systems Group (MSG), which also provided the sampling frames.

The CSS process matches to White and Yellow Pages to identify nonresidential business numbers and dials all numbers to identify those that are nonworking. The method also identifies cellular telephone numbers, which were excluded in this survey. Table 1-1 shows the

CSS result codes as well as the distribution of the sampled telephone numbers. Approximately 56 percent of the sampled numbers (result codes CP, LB, FM, NR NW, and some UB) were excluded from dialing. This percentage was larger than those seen in other surveys.

Table 1-1. CSSR result codes and their distribution in the AHS sample

CSSR result code	Description	Number of telephones	Percentage
CP	Agent dispositioned cell phone	32	0.0
DK	Undetermined	36,410	28.2
FM	Fax/modem	4,168	3.2
LA	Language barrier	476	0.4
LB	Listed business	4,275	3.3
NR	No-ring back	1,554	1.2
NW	Nonworking	56,850	44.0
PM	Privacy manager	4,176	3.2
RS	Residence	14,745	11.4
UB	Unlisted business	6,338	4.9
WR	Pro-t-s wireless detection	226	0.2
Total		129,250	100.0

The second technique used to increase the efficiency of data collection was stratifying the telephone numbers by mailable status and subsampling the strata at different rates (Brick, Judkins, Montaquila, and Morganstein, 2002). Telephone numbers with a mailing address (or "mailable") are more likely to be residential so the cost of finding a residence is much lower in the substratum of mailable numbers. In addition, households with mailable telephone numbers are more likely to cooperate with most surveys[1]. The subsampling rate was determined using the principles of optimal allocation to balance both data collection costs and the variances of the estimates; 75 percent of non-mailable numbers were selected.

The third technique used to reduce costs while improving the sample efficiency in AHS was subsampling of refusals for refusal conversion (Brick et al., 2005). In this procedure, a larger sample of telephone numbers than would otherwise be selected is drawn in the first phase. Each number in the first-phase sample is randomly assigned to one of two conditions in the second phase: subsampled for refusal conversion or not subsampled for refusal conversion. When refusals are encountered at the screening stage of data collection, only numbers in the subsample are eligible for refusal conversion followup (at the screener level). The numbers subsampled for refusal followup are generally fielded first so that refusal cases can be worked completely (i.e., all

[1] The subsampling increases the percentage of respondents but not the response rates since the response rates are weighted to account for the subsampling.

of the appropriate scheduling procedures including holding periods for refusal cases can be fully implemented). For AHS, the initial subsampling rate was 60 percent. In practice, because not all sampled numbers were fielded and those designated for conversion were fielded first, the rate wound up at about 65 percent for each stratum.

One disadvantage of subsampling of telephone numbers without a mailing address and refusals for refusal conversion is that a weighting adjustment is needed to account for the subsampling, so that those cases that refuse and are subsampled are weighted to represent themselves and the cases that refuse and are not subsampled. This weighting decreases the precision of the survey estimates, but only very slightly. These weighting adjustments are described in more detail in Chapter 6.

1.1.2 Noncoverage Issues in Telephone Surveys

As in most RDD surveys, households with no access to landline telephones, including those in households with only cellular telephones and households with no telephone service, were not sampled for AHS. For estimates correlated with socioeconomic measures such as health insurance coverage and poverty, this coverage loss could introduce bias. The bias is related to the percentage of households with no landline telephones and the difference in characteristics of persons in households with and without a landline telephone. The proportion of nonlandline households has increased in recent years due to an increase in households with only cellular telephones (see Blumberg et al., 2006). By 2007, about 15 percent of households nationally had only cellular telephones. To reduce potential biases that result from the exclusion of households with no landline telephones, special weighting procedures were used in AHS with the use of socioeconomic variables such as household tenure in addition to age and race/ethnicity that are correlated to cell phone use.

Another source of coverage error in telephone surveys arises when persons who do not speak English are sampled but are not interviewed because of language limitations. These cases are treated as nonrespondents, but could easily be thought of as a coverage problem since none of the persons speaking languages other than those included in the survey protocol are interviewed. In AHS, an effort was made to limit this potential bias by interviewing in Spanish. This effort should eliminate a large source of the bias that might result from conducting interviews in English only.

1.2 Sample Design

In this section, we describe the steps used in selecting the sample of telephone numbers for AHS. After describing the population of interest, we describe the stratification of the frame, the selection of the sample of telephone numbers after adjusting for expected losses due to nonresponse, and subsampling the numbers based on mailable status and refusal status to improve the efficiency of the sample.

1.2.1 Population of Interest

The AHS sample was intended to represent the adult (age 18 and older) residential population of Arizona, as well as adolescents (age 12-17) and children (age 11 and under) in Maricopa County. Eligible residential households included houses, apartments, and mobile homes occupied by individuals, families, multiple families, extended families or multiple unrelated persons, if the number of unrelated persons was less than nine. Persons living temporarily away from home were eligible and enumerated at their usual residences. These include college students in dormitories, patients in hospitals, vacationers, business travelers, and so on. The survey excluded group quarters – any unit occupied by nine or more unrelated persons (e.g., communes, convents, shelters, halfway houses, or dormitories). Institutionalized persons (e.g., those living in prisons, jails, juvenile detention facilities, psychiatric hospitals and residential treatment programs, and nursing homes for the disabled and aged), the homeless, persons in transient or temporary arrangements, and those in military barracks were also excluded. As described in the previous section, some individuals who were part of the residential population did not have a chance of selection, including those living in households without landline telephones (either without any telephone service or with cellular telephone service only), and children and adolescents living in a household without a parent or legal guardian.

1.2.2 Sampling Households

The AHS RDD sample had an initial goal of completing 4,000 adult interviews statewide with 3,000 of those interviews from adults residing in Maricopa County. To meet this goal, AHS used a stratified sample, where telephone numbers were divided into two non-overlapping strata, one corresponding to Maricopa County and the other for the remainder of the

state. In this design, the stratum that includes Maricopa County was oversampled at a higher rate than the remainder of the state.

Since the geographic information required to stratify the sampling frame is only available at the exchange level[2], the 100-banks of telephone numbers were stratified indirectly by mapping the exchanges to the strata. Although the majority exchanges could be assigned to a stratum in this way, some telephone exchanges crossed stratum boundaries (i.e., they serviced households inside and outside Maricopa County). To allocate these exchanges we used the coverage report for Maricopa County with the list of exchanges that provided service in the county. For each exchange, we examined the total number of listed households in the exchange and the proportion of listed households that are within the county. We allocated the exchanges to Maricopa County if 50 percent or more of the serviced households in the exchange were located in the county.

1.2.3 Sample Selection

The number of telephone numbers selected in any RDD survey has to be greater than the targeted number of completed interviews to account for a variety of factors. For example, a substantial percentage of the sampled telephone numbers are not residential. We increased the initial sample draw to deal with losses due to the following sources:

- Nonworking, nonresidential, and never-answered numbers;
- Subsampling of telephone numbers without a mailing address;
- Subsampling for refusal conversion;
- Nonresponse to the screening interview; and
- Nonresponse to the extended interview.

The first, fourth, and fifth sources noted above are typical of all RDD surveys. To deal with these losses we used information from the coverage reports to estimate the percentage of the telephone numbers that would not be residential. We used the results from other surveys to estimate the percentage that would not respond to the screener and extended interviews, and increased the sample size accordingly. Taking all of these factors into consideration,

[2]A telephone exchange consists of 10,000 consecutive telephone numbers with the same first six digits including area code. An exchange is a set of area codes and prefixes serving the same geographic area.

129,250 telephone numbers were drawn for AHS. Not all the telephone numbers were selected at the same time, as assumptions of the sample design were modified during the field period. Notably, the rate of nonresidential and nonworking numbers proved higher than expected After each selection, duplicate telephone numbers were removed from the samples. Table 1-2 summarizes the number of telephone numbers drawn in each stratum.

Table 1-2. Number of telephone numbers drawn by sampling stratum

Sampling stratum	Number of telephone numbers drawn
Maricopa County	103,835
Remainder of Arizona	25,415
Total	129,250

1.3 Within-Household Sampling

Once the sample of telephone numbers was selected, interviewers called the numbers, and conducted interviews with sampled persons within the household. The AHS design called for the random selection of one adult from all the adults in each sampled household. In addition, in those households in Maricopa County with adolescents (ages 12-17) and/or children (under age 12), one adolescent and one child were sampled and interviewed (a parent of the child was interviewed about the child).

To reduce the burden on the respondents and increase the utility of the data, children and adolescents were selected using a *linked* sampling approach. In this approach, children and/or adolescents for whom a sampled adult was a blood or adoptive parent or a legal guardian were considered as linked to or "associated" with that adult. Persons were sampled in two phases. In the first phase, an adult was randomly sampled from all the adults in the household. In the second phase, a child was sampled from all the children associated with the sampled adult. The probability of sampling the child was the product of the probability of sampling the adult (phase one) and the probability of sampling the child from all children associated with that adult (phase two). Adolescents were sampled in the same way, that is, one adolescent was sampled from all adolescents associated with the adult sampled in the first phase.

To use the linked sampling method, data are needed to link children and adolescents in a household to the sampled adult and his/her spouse/partner. (Children or adolescents linked to

both the sampled adult and spouse/partner could be selected if either adult was sampled.) These data were collected during the screener interview or the adult interview.

1.3.1 Child First Procedure

To increase the number of child and adolescent interviews, they could be conducted prior to the adult interview under certain conditions. This process was an operational method (not a sampling method) called the "child-first" procedure, which involved enumerating and sampling children and adolescents at the end of the screener rather than during the adult interview. Thus, if the sampled adult did not complete the extended interview, we could still obtain child and/or adolescent interviews.

The child-first procedure was used only when the screener respondent was the spouse or partner of the sampled adult, there were children in the household, and the sampled adult was not available at the time of the interview. If these conditions were met, a child and or adolescent could be sampled and the appropriate interview conducted without waiting for the completion of the adult interview. The child first procedure was used in 445 households (25 percent) out of 1,799 households with children identified during the screener.

1.3.2 Adult Sampling

An adult was defined as any person 18 years or older residing in the household. The procedure to select adults in AHS is called the Rizzo method (see Rizzo et al., 2004, for a complete discussion of the method and its implementation). The principal advantage of this method is that the enumeration of adult household members is bypassed in most households, so it is less intrusive while still resulting in a valid probability sample. In this method, all sampled adults have an equal probability of selection.

A sampled adult was selected using the following steps:

- Ask the screener respondent (who must be an adult living in the household) how many adults are in the household (i.e., N); The respondent answers N = 1, 2, 3, ...;

- If there is only one adult in the household (i.e., N = 1), then that adult is selected;
- If there are two adults in the household (i.e., N = 2), then the CATI system accesses a pre-generated uniform random number between 0 and 1;
  - If the random number is less than or equal to 0.5 then the screener respondent is selected;
  - If the random number is greater than 0.5 then the other adult is selected;
- If there are more than two adults in the household (i.e., N > 2), then the CATI system accesses a pre-generated uniform random number between 0 and 1;
  - If the random number is less than or equal to 1/N (i.e., the inverse of the number of adults in the household) then the screener respondent is selected;
  - If the random number is greater than 1/N, then the screener respondent is asked which of the other adults is the next to have a birthday;
  - If the screener respondent knows which of the other adults is next to have a birthday, then the adult with the next birthday is selected; and
  - If the screener respondent does not know which of the other adults is next to have a birthday then the screener respondent is asked to list the adults in the household (excluding himself/herself) and the CATI system randomly chooses one of the adults from this roster.

If the number of adults in the household was unknown then the screener respondent was asked to list the adults in the household (including himself/herself) and the CATI system randomly chose one of the adults from this roster. No other sampling steps were necessary.

1.3.3 Child Sampling

The method used to select one child, in Maricopa County only, was developed to increase the number of interviews for younger children (0 to 5 years old) while reducing the number of interviews for older children (6 to 11 years old). If there were only younger or older children in the sampled households, a child was selected with equal probability of selection. In households with both younger and older children, younger children were assigned a greater probability of selection with respect to the older children. The probability assigned to children *i* in the household *h*, p_{hi} , was assigned as

$$p_{hi} = \begin{cases} \dfrac{2NC_{1h}}{2NC_{1h} + NC_{2h}} & \text{If age of child } i \text{ in household } h \text{ is between 0 and 5 years old (younger child)} \\ \dfrac{NC_{2h}}{2NC_{1h} + NC_{2h}} & \text{If age of child } i \text{ in household } h \text{ is between 6 and 11 years old (older child)} \end{cases}$$

where NC_{1h} is the number of younger children and NC_{2h} is the number of older children in the household *h*. For example, in a household with one young child and one older child, the young child was twice as likely to be selected as the older child. The disadvantage of this approach was that the number of interviews about older children was reduced and there was a slight increase in the design effect for estimates for all children due to the disproportionate sampling.

1.3.4 Adolescent Sampling

In AHS, an adolescent was sampled from among all eligible adolescents in a household associated with the sampled adult with equal probability. Adolescents were enumerated and sampled at the end of the screener interview if the child-first procedure was used or in section G of the adult extended interview. Since adolescents could be sampled and interviewed before the adult interview, there were some households with a completed adolescent interview where adult and/or child interviews were not completed.

1.4 Achieved Sample Sizes

The goals for AHS were stated in terms of the total number of completed adults by geographic area obtained at the end of the data collection period. The actual number of completed interviews was a function of the number of telephone numbers sampled, the within-household person sampling, and different reasons for attrition.

Table 1-3 shows the number of completed adult interviews by two methods of classifying the geographic area in which the sampled adult resides. The first column of completed interviews in the table uses the data on the county that was available at the time of sampling (and during the data collection period). As noted in Section 1.2, each telephone number was assigned to exactly one stratum for sampling purposes, but the number may actually be for a household in a different stratum. The third column in the table uses the self-reported residence county of the adult respondent. This classification was based on the geocoded location of the adult's residence

derived from data collected on the county, ZIP Code, address, and street intersection in the adult interview. This classification is the most appropriate for analysis of AHS data. This table shows that the target goals for the number of completed adult interviews were met in AHS.

Table 1-3. Number of completed adult interviews by sampling and self-reported stratum*

Stratum	Sampling location		Self-reported location	
	Completed interviews	Percentage	Completed interviews	Percentage
State Total	4,196	100.0%	4,196	100.0%
Maricopa County	3,139	74.8%	3,130	74.6%
Remainder of Arizona	1,057	25.2%	1,066	25.4%

*Partially completed interviews (completed through at least Section K) are counted as complete

Table 1-4 shows the number of completed child and adolescent interviews by sampling and self-reported stratum. There were 8 children and 1 adolescent reported to be outside Maricopa County. These cases are ineligible for the survey.

Table 1-4. Number of completed child and adolescent interviews by sampling and self-reported stratum

Stratum	Child Interview		Adolescent Interview	
	Sampling location	Self-reported location	Sampling location	Self-reported location
Maricopa County	643	635	204	203
Remainder of Arizona	0	8	0	1

2. SURVEY QUESTIONNAIRES

The AHS interview structure was modeled on that of the 2007 California Health Interview Survey (CHIS). AHS interviews could include, for a given household in Maricopa County, up to three substantive questionnaire sections: the adult, child, and adolescent extended questionnaires. For households in the remainder of Arizona, only the adult interview was conducted. In addition to the substantive survey content, the CATI instruments performed sampling and administrative functions, including identifying eligible individuals and selecting sample members from among them, identifying appropriate respondents for the various questionnaires, and sequencing the activities within a household. All of these functions were programmed into the CATI instrument and are described in this chapter.

2.1 Screening Interview

The AHS sample was composed of telephone numbers selected as described in Chapter 1. On first contact with a sampled landline RDD telephone number, interviewers needed to:

- Identify a household member 18 years of age or older to act as informant (i.e., screener respondent);
- Determine whether the telephone number was associated with a residence; and
- Ask how many persons 18 or older lived in the household and select one for the extended interview.

These basic elements were scripted into the initial screening interview. For Maricopa County "child-first" cases, once an adult was sampled, the screening interview included enumeration and sampling of children and adolescents The following elements were also included in the screener to assist in developing survey weights:

- The number of children under 12 years of age living in the household;
- The number of adolescents between 12 and 17 years of age living in the household; and
- The number and use (home, business) of telephone numbers ringing into the household.

2.2 Overall Structure of AHS Interviews

Given the number of different instruments and the rules for who could respond to each, one household could potentially have several individuals acting as respondents, including:

- The screener respondent;
- A sampled adult;
- An adult who could give permission for the adolescent interview, who except in rare instances was the sampled adult or the screener respondent;
- A sampled adolescent; and
- A "most knowledgeable adult" (MKA) for the child extended interview.

In practice, one adult usually filled multiple roles in households with adolescents and/or children. However, the possibilities of multiple respondents required rules for the order of instruments and of the various administrative activities (e.g., selecting sample persons, identifying and contacting respondents), and CATI tools for navigating through the administrative and questionnaire screens. The default sequence of questionnaire and navigation sections for Maricopa County cases is presented in Figure 2-1. The sequence was much simpler for cases in the remainder of Arizona, since there were no child or adolescent interviews. A basic principle of the interview flow is that once the sampled adult is on the telephone, the interviewer should attempt to complete as many different parts of the interview as possible with that person. Once that has happened, the system goes to the HHSELECT screen. If there are remaining parts of the interview, the interviewer selects another individual (e.g., the MKA for the Child Questionnaire), and so on.

The AHS protocol allowed sampling of children and adolescents as part of the screening interview under prescribed circumstances. If the screener respondent who was the sampled adult's spouse was determined to be the MKA, the child interview could be completed immediately or at another time before the adult questionnaire. The adolescent interview could also be completed before the adult interview in child-first cases.

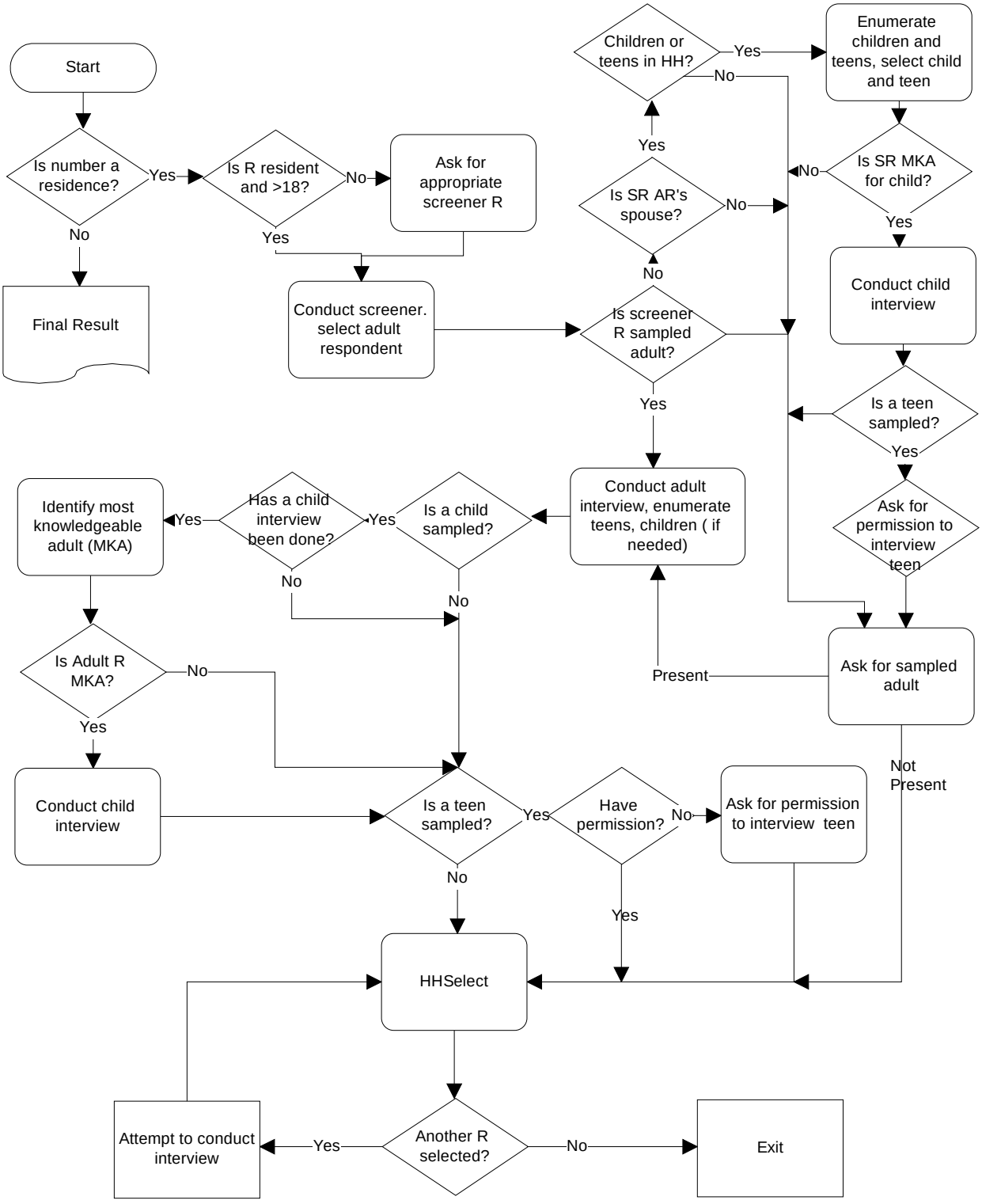

Figure 2-1. AHS Interview Flow, Maricopa County Sample

For cases other than those meeting the child-first criteria, the screening interview for Maricopa County cases resumed in the middle of Section G of the Adult Extended Questionnaire, with the following items:

- Identification of adult respondent's spouse if living in the household;
- Enumeration of adolescents and children in the household; and
- Determining for which adolescents and children the adult respondent and/or spouse is the parent or legal guardian.

This information was used by the CATI program to select one adolescent and one child among those for whom the sampled adult was the parent or legal guardian. Adolescents or children who did not have a parent or legal guardian in the household were not eligible for selection. Although child and adolescent extended questionnaires were not completed for cases in the remainder of Arizona, children and adolescents were also enumerated in these adult interviews, and an eligible child and/or adolescent sampled for the relevant health insurance sections.

Because sampling children and adolescents was part of the adult interview except for child-first cases, the adult interview had to be completed first. Other basic principles of the CATI system flow, once the adult interview is completed, included attempting to complete as many components as possible with the current respondent before asking for someone else, and attempting the child interview before asking permission for the adolescent interview. After an adult interview was completed for non-child-first cases, if an adolescent and/or child was selected the sampled adult was asked to identify the respondent for the child interview and to give permission for the selected adolescent to be interviewed.

Once all possible components were attempted with the current respondent, the CATI program displayed a master navigation screen called HHSELECT. A sample HHSELECT screen is presented as Exhibit 2-1. HHSELECT displayed all interviews scheduled for a household, the name of the respondent, and whether the interview had been completed. The interviewer selected one of the outstanding interviews from HHSELECT, and was routed to the appropriate introductory screens for that interview. HHSELECT reappeared after each component was completed, or attempted but not completed. It also appeared when an interviewer first entered a case started by another interviewer.

Exhibit 2-1. AHS HHSELECT CATI screen

```
0.0020  HHSELECT              900009990201 – (301) 215-1500 – 08:26

     [ASK FOR PEOPLE WITH RESULT THAT IS NOT FINAL. ENTER NUMBER FOR
CHOSEN
     PERSON. ENTER 0 TO LEAVE THIS CASE.]

                                    (  )
                                                AT
                                                THIS            APPOINTMENT
#      RESPONDENT      TYPE    SUBJECT          PHONE   RSLT    DATE/TIME
1      MARY/30/F       ADLT                       Y     CA

2-SR   ALFRED/32/M     CHLD    WILL/8/M           Y
```

2.3 Extended Interviews

AHS included three separate extended interviews: adult, child, and adolescent. This section describes Westat's involvement in the development of these questionnaires, the content of each, pretesting of the questionnaires, translation of the questionnaires from English into Spanish, changes in the questionnaires during data collection, and how proxy interviews were conducted.

2.3.1 Questionnaire Development Process

The AHS questionnaire design was driven by the research needs of SLHI and its partners, as well as by concerns about respondent burden, response rates, and costs. The target was an adult questionnaire that would not exceed 25 minutes in mean administration time, and child and adolescent questionnaires that would not exceed 15 minutes each.

In early 2008, SLHI began collaboration with Westat staff for drafts of the adult, adolescent, and child questionnaires. These drafts were developed by SLHI and its partners to cover a wide variety of health-related research topics. Westat reviewed the drafts and provided

comments on the selection of question items, wording and sequence, and on the estimated length of the draft instruments. There were several iterations of draft instruments before complete instruments of reasonable length were ready for testing.

AHS used the CHIS 2007 CATI questionnaires as a base. Once the content for testing was agreed upon, Westat staff began modifying the specifications and CATI programs developed for CHIS. Westat's automated management system for CATI specifications tracked question text, sequencing, response categories, the appropriate use of "fills" within questions based upon previously recorded information, and range and logic checks. The CATI specification document, published both in PDF and Microsoft Word format, provided the guide for project staff and programmers as to what the CATI instrument should include. The system tracked each change to the specifications and the reason for that change, whether it originated from SLHI, Westat project staff, or the programming team.

Once programming commenced, all sections of the questionnaires were tested to make sure that the CATI instrument was working according to the specifications. Testing also covered the database used to store the captured responses. This testing included review by project staff, TRC staff (including interviewers), data preparation staff, and by SLHI. After the pilot test and then again during the first few weeks of the statewide field period, the data preparation and programming staffs reviewed frequency counts from each instrument to make sure that the CATI program was performing correctly and all responses and administrative data were being stored in the appropriate variable fields.

2.3.2 Questionnaire Content

The adult extended questionnaire is divided into 11 sections:

A. **Demographics** – Age, gender, race, ethnicity, marital status.

B. **Health Conditions** – General health, asthma, diabetes, high blood pressure, heart disease, gastrointestinal disorders, bi-polar disorder, anxiety disorder, depression, flu shot, falls among the elderly.

C. **Health Behaviors** – Physical activity, dietary intake, tobacco and alcohol use, use of nonprescription drugs and supplements.

F. **Mental Health** – Mental health status, effects of mental health problems, resiliency and coping, use and sources of treatment, reasons for not seeking treatment

G. **Demographics, Part II** – Country of birth, first language, English proficiency, household composition, education, employment status of self and spouse, and receipt of Social Security disability or SSDI.

H. **Health Care and Health Insurance –** Usual source of care, emergency room use, current coverage by public or private plans, source of coverage, spouse's coverage, coverage of prescription drugs, dental care, and behavioral health services, duration of coverage, whether any uncovered period in past year, and medical debt.

I. **Adolescent and Child Health Insurance** – For sampled adolescent and child, current coverage by public or private plans, source of coverage, coverage of prescription drugs, dental care, and behavioral health services, duration of coverage, and whether any uncovered period in past year.

J. **Health Care Utilization and Access and Violence** – Height and weight, doctor visits in past year, communication with doctor, delays in getting care, interpersonal violence from intimate partner or acquaintance, family life in childhood.

K. **Employment, Income, Poverty Status** – Employment status, earnings for self and spouse, household annual income, participation in TANF.

M. **Housing, Parks, Transportation** –. Type of housing and tenure, feeling safe in neighborhood, characteristics of neighborhood and its residents, volunteer service.

N. **Final Demographics** – County of residence, address, willingness to participate in follow-up study.

The child extended questionnaire comprises eight sections:

A. **Demographics and Health Conditions** – Age, height, and weight, school attendance, asthma, developmental disorders.

B. **Dental Health** – Most recent visit to a dentist, usual source of dental care, dental insurance, unmet needs, loss of school time.

C. **Diet and Physical Activity –** Types of food eaten, getting to school, name of school, physical activity.

D. **Access to and Use of Health Care Services** – Usual source of care, most recent physician visit, communication with doctor, emergency room visits, delays in care.

E. **Public Programs** – Participation in TANF, Food Stamps, and WIC, parental influence.

F. **Parental Involvement, Concerns, Mental Health –** Parental involvement with child, developmental and behavioral concerns of parent, school, and doctor, mental health and development, parenting support.

G. **Child Care** – Types of child care used, difficulty finding care

H. **Demographics, Part II** – Race and ethnicity, respondent's English proficiency, and level of education of respondent and primary caretaker of child.

For child-first cases, some completed child interviews do not have completed adult interviews in the same household. The following topics from the adult questionnaire were administered to the MKA as part of the child questionnaire for child-first cases so that these children would have essential household-level and insurance information for analysis and weighting:

- Adult respondent's (AR's) employment status and age;
- Health insurance coverage for AR, spouse, the sampled child, and the sampled adolescent (if there is one);
- Household income;
- Own/rent home; and
- Address information.

Finally, the adolescent extended questionnaire comprises twelve sections, presented in the order they appear in the interview:

A. **Demographics** – Age, gender, school attendance.

B. **Health Status, and Health Conditions** – Self-reported health status, height and weight, days missed from school, asthma, diabetes.

D. **Diet, Nutrition, and Food Environment –** Dietary intake, sources of meals**.**

E. **Physical Activity and Sedentary Time –** Exercise, transportation to school, sports team participation, physical education in school, sedentary time on weekdays and weekends, park or playground availability and use, volunteering.

F. **Tobacco, Alcohol, and Drug Use –** Smoking habits, drinking, use of recreational drugs.

G. **Emotional Functioning** – Mental health over past 30 days, past week.

H. **Sexual Behaviors** – Sexual activity, contraception use.

L. **Interpersonal Violence** – Fights, intimate partner violence

I. **Health Care Utilization and Access** – Usual source of care, most recent doctor visit, recall of provider advice, emotional or psychological counseling, unmet needs.

M. **Dental or Oral Health** – Most recent dental visit, usual source of care, unmet needs, loss of school time because of dental problems.

J. **Adult Supervision** – Marital status of parents, adult presence after school.

K. **Demographics, Part II** – Race and ethnicity, country of birth, languages spoken at home, follow-up information.

2.3.3 Translation of Questionnaires and Other Materials

The AHS was administered in both English and Spanish. The questions were developed in English, and then translated into Spanish. Where questions were taken from CHIS, the Spanish translations were adopted as well. For other items, Westat arranged for translation by an outside service. The translations were then reviewed internally by bilingual Westat staff and sent to SLHI for an independent review. The Westat and SLHI reviewers compared notes and agreed upon the final translations.

Other materials, including letters to respondents and scripts for frequently asked questions, were also translated into Spanish. The translations were drafted by Westat staff and reviewed by SLHI. Prenotification letters were printed in English on one side and Spanish on the other.

2.3.4 Pilot Test

Westat conducted a small pilot test of the AHS CATI adult, child, and adolescent interviews the week of March 17, 2008. The purpose of this test was to estimate the time to administer the instruments and to assess the interview flow and wording. Respondents were recruited by SLHI. Westat interviewers in the Merced, California, Telephone Research Center (TRC) conducted 16 adult interviews, 2 adolescent interviews, and 1 child interview. All pretest

interviews were conducted by experienced interviewers and monitored by Westat staff; SLHI staff also monitored many of the interviews. Results from the pretest informed subsequent decisions about dropping or revising questions.

3. INTERVIEWER RECRUITING AND TRAINING

3.1 Organization of the Telephone Research Center

Westat conducted AHS at four of its Telephone Research Centers (TRCs) – in Rockville and Frederick, Maryland, and Citrus Heights and Merced, California – in addition to utilizing interviewers working from their homes nationwide. Overall direction of telephone survey operations was from the TRC central office at the Rockville headquarters.

Westat's TRC has applied new technologies to expand the multi-site call centers to include interviewers working throughout the US. Westat's computing systems and telephony capabilities enable the networked combination of geographically diverse interviewer locations to operate as a single and secure "virtual" TRC managed from the home office location at Rockville. All interviewing and supervisory stations at all locations are interconnected on a high-speed data communications network that provides a single integrated database and a single call scheduling and reporting capability. Integrated voice and data monitoring is available for supervisors at all locations and at a central facility at the Rockville home office. Each center, including the home-based interviewers, has an administrative director and a group of supervisors who schedule and supervise the center's interviewing staff.

3.2 Training for English-language Interviewing

The AHS interviewing force consisted entirely of interviewers who had previously worked on CHIS 2007. Due to the similarity of AHAS to CHIS, interviewers were able to make this transition smoothly after a two-hour self-paced tutorial. After completion of this web-based training, 97 interviewers worked on the study.

The two hours of project-specific training involved interviewers completing a web-based distance learning session. This training included completion of a full interview in an interactive format. A program was used which simulates the administration of an actual interview, complete with respondent answers to ensure all trainees followed the identical path. Other materials to be reviewed in this self-paced training included the AHS advance letter, a pronunciation guide, answers to commonly asked questions, the toll-free numbers and web site for reference and a memo with details of the study.

Trainings began on March 20, 2008 and continued until the required number of interviewers were working on AHS. After all interviewers started production, they received supplemental training on specific questionnaire issues that arose after training. Interviewers already trained in refusal conversion for CHIS 2007 were recruited to work on refusal conversion cases for AHS.

3.3 Training for Spanish-language Interviewing

All Spanish bilingual interviewers were trained according to the protocol described in Section 3.2.1. Spanish interviewing was conducted at all TRCs and also by bilingual Spanish speakers working from home. After completing the English-language AHS project-specific training, Spanish bilingual interviewers initially worked in English. Once the Spanish-language instrument was ready, bilingual interviewers were given practice using it before proceeding to live interviewing in Spanish. The training was monitored by Spanish-speaking supervisors. Since the English and Spanish instruments were so similar, there were few substantive or operational issues to work through during training.

Once the interviewers began interviewing at the TRCs in Spanish, they were monitored closely by Spanish-speaking supervisors. The first priority in CATI for Spanish bilingual interviewers were cases from the work class identified as speaking Spanish. Bilingual Spanish interviewers worked primarily in the Spanish work class for the rest of the field period but also made the initial follow-up calls to households that English speaking interviewers categorized as OTHER LANGUAGE (not Spanish). The expectation was that some of these would turn out to be Spanish speaking households not identified by a non-bilingual interviewer.

3.4 Training for Proxy Interviewing

For cases where a sampled adult was 65 or older and unable to be interviewed for physical or mental health reasons, the interviewer attempted to identify an appropriate proxy respondent. The proxy had to be an adult member of the household who knew about the sampled adult's health and health care. The CATI questionnaire was modified to accommodate proxy interviews.

A group of selected interviewers were trained to conduct the proxy interviews. Training comprised a review of how to contact households identified as candidates for proxy interviews, determining whether a proxy would be appropriate, identifying a respondent, and a review of the changes to the questionnaire for proxy interviews. Cases identified as eligible for proxy interviews were grouped in a separate work class and delivered by the CATI system only to interviewers trained for proxy interviewing.

3.5 Quality Control and Post-training Support for Interviewers

3.5.1 Evaluating Interviewer Performance

Interviewer performance was evaluated through examination of cooperation rate reports and monitoring of live interviewing. Ten percent of interviewing time was monitored throughout the data collection period. Supervisors monitored interviewers for a minimum of ten minutes at a time. The monitoring was followed by a one-on-one coaching session to reinforce exemplified skills or provide feedback for improvement. Skills addressed in monitoring included use of a conversational style; reading fluency; ability to answer respondent questions quickly, accurately, and completely; ability to gain respondent cooperation; reading screens verbatim; and using neutral probes. Interviewers whose performance fell below acceptable levels attended additional coaching sessions with an emphasis on gaining respondent cooperation and answering respondent questions.

3.5.2 Triage

Interviewing during all hours of TRC operation is supported by a specially trained "triage" team leader. The triage team leader was called whenever a problem interfered with the ability to conduct CATI interviewing. When the triage team leader received a problem report, he or she diagnosed the problem and called the appropriate personnel. Hardware, software, and project-specific support were always available via home telephones or beeper numbers. The appropriate support personnel were able to respond to problems within minutes of a problem report, regardless of the time.

3.5.3 Communication with Interviewers

Multiple communication tools were in place to ensure that timely support was available for immediate data collection concerns, that interviewers were all provided with current information for interviewing, and that all questions were answered.

During work shifts, all interviewers logged onto a business instant messaging system. A global message provided the name of the contact person for all problems occurring during the current work time. Interviewers sent instant messages with pertinent questions or concerns to this person and were either instructed on the steps to take for remediation or provided an answer. Additionally, a person monitoring an interview could send an instant message to a interviewer to provide instantaneous guidance and support or post-interview feedback.

A website provided instructions interviewers were instructed to access before each work shift. This site provided basic information on the background of the study, the time frame for data collection, study goals and commonly asked questions and answers. Frequent updates were made to apprise interviewers of the progression of completed interviews. Memos were posted on the website. Guidance was provided with specifics on how to code certain questions based on problem cases, interviewer questions and monitoring notes.

Each interviewer had a personalized Learning Management System (LMS) which supplied and tracked completion of training materials. In addition to memos being posted on the web site, these were added to each person's LMS requiring an acknowledgment of having reviewed each item.

Interviewers were invited to weekly scheduled conference calls as an open forum across studies for sharing experiences and concerns. The sessions were hosted by senior TRC staff with knowledge of specific study procedures and general techniques for successful data collection. These sessions provided an opportunity to interact with other interviewers and to ask any questions about AHS or general procedures.

4. DATA COLLECTION

4.1 Scheduling and Release of Work

This section describes activities related to initiating data collection, including preparation and release of sampled telephone numbers, how the sample was organized in the CATI system, purging the sample of nonworking and business numbers, mailing advance letters, and handling inbound calls to Westat's AHS toll-free number. Data collection began March 24, 2008, and ended June 9, 2008.

4.1.1 Preparation and Release of Sampled Numbers

As described in Chapter 1, the initial AHS sample draw was 129,250 telephone numbers. About 56 percent of these were purged as nonworking or business numbers. The remainder were matched with directory services to find mailing addresses. Numbers without mailing addresses were subsampled. After the purge and subsampling, 55,495 numbers were available to be called in CATI. These were divided into release groups, with all of the numbers designated for refusal conversion fielded first. Ultimately, 8,103 numbers from the "no conversion" stratum were never fielded, leaving 47,392 numbers that were called.

An advance letter signed by the AHS Director at SLHI was sent for all sampled telephone numbers for which an address was available. The advance letter was printed in on SLHI letterhead in English and Spanish. A different letter, also signed by the AHS Director, was sent after initial refusals for the screening interview (for cases designated as "conversion"), adult interview, or permission to interview a selected adolescent, if an address had been obtained for the sampled number. These letters were also printed in English and Spanish. Appendix 1 includes copies of the survey letters.

4.1.2 Sample Management

All of the numbers released to CATI were classified by whether they were designated for refusal conversion at the screener stage or not. Numbers designated for conversion were fielded before those that were not. Sample numbers were divided into "release groups," or

random subsets of the overall samples. They were fielded in such a way that the pre-notification letters would be received within a few days of the initial telephone contact attempt. Both cases with and without addresses were given the same priority within the CATI scheduler.

Within the CATI system, active and completed cases were allocated into work classes, which are divisions of the sample that are to be worked by interviewers with special training or skills. Westat's CATI scheduler treats each work class as an independent sample. Work classes were given priority order for delivery of work to qualified interviewers. For example, a refusal converter would always be delivered a refusal work class case if one was available before being given a case from the default work class. The AHS work classes were defined as follows:

- **Default—**All cases on initial release, and continuing cases that had not been moved to another work class; available to all interviewers;
- **Refusal—**Any sample case that encountered a refusal at any point in the interview process, whether at the screener or any extended interview level; available only to interviewers selected to work and trained as refusal converters. There were five different refusal work classes: screener initial refusal, extended refusal (other than adolescent and adolescent permission), adolescent refusal, adolescent permission refusal, and second refusals of any type;
- **Hearing/Speech—**Any case in which a respondent was determined to have difficulty communicating because of hearing or speech impairment;
- **Language (Spanish)—**Any case determined or suspected to require a Spanish bilingual interviewer to re-contact; available only to the appropriate bilingual interviewers; and
- **Proxy Interviews—**For sampled adults 65 or older who could not complete the interview because of poor health or physical limitations, selected interviewers attempted to complete an interview with a proxy respondent in the household.

Westat data collection and statistical staff monitored the yield (number of completed interviews) by stratum to assess sample performance. Some cases in each stratum were held in reserve; because the yield was lower than initially expected, additional sample was released for calling. The monitoring process was repeated several times, re-calibrating the fielded sample as more information on progress to date became available.

4.1.3 Inbound Toll-Free Calls

Westat maintained a toll-free number for respondents to call with questions about the survey. Interviewers provided the number throughout the data collection period to respondents who requested additional information. The toll-free line was staffed weekdays from 9 a.m. to midnight Eastern Time, Saturdays from 10 a.m. – 6 p.m. Eastern Time, and Sundays from 2 p.m. – 10 p.m. Eastern Time. In the event an operator was not available to answer the call or for calls made outside of the above time frames, the caller was directed to a voicemail message specific to AHS.

Between the start of data collection in March 2008 and the end in June 2008, 88 calls were made to Westat's toll-free number. Most were simply to verify the legitimacy of the study or ask general questions with no further action required.

SLHI also maintained a separate toll-free number during the field period, which was included on all respondent letters and was also available on the AHS web site. Westat interviewers provided the SLHI number to respondents who specifically wanted to talk with someone at SLHI. There was continual back-and-forth contact between SLHI and Westat in response to these calls, usually to remove a person from the sample who indicated to SLHI that they did not want to be called.

4.2 Data Collection Strategies

4.2.1 Answering Machines

Studies indicate that leaving a message on an answering machine seems to increase cooperation rates (e.g., Xu et al., 1993). Apparently the message acts as an advance letter in that it legitimizes the study, allows the respondent time to make an informed decision, and distinguishes the "survey telephone call" from telemarketing calls. Because of this finding in the literature, the message below was left the first time an answering machine was encountered at a dialed telephone number.

> Hello, I'm {INTERVIEW NAME} from the Arizona Health Survey. This is a scientific study about health care that may help improve the services in your community. I'm not asking for money. I do hope that you can spare a few minutes to take the survey. We will call back

within the next few days, or you may call us toll-free at 1-888-290-7033 to schedule an appointment. Thank you for your help.

Overall, about 45 percent of all cases attempted at the screener level had at least one call reach an answering machine. For Maricopa County, the proportion was about 47 percent, and it was about 40 percent for the remainder of Arizona. Most of these cases wound up with a screener result other than "NM" – about 22 percent of the cases with at least one answering machine contact had no human contact over 14 calls. At the adult extended interview, about 37 percent of all cases had at least one answering machine contact.

4.2.2 Time Slice Strategy

If the initial call attempt resulted in "no answer," a busy signal, or an answering machine, the call scheduler would automatically place the telephone number into time slice queues so that additional calls would be made over several days at several different times of day. The goal was to find a time when someone would answer the telephone.

The time slices were defined as: (1a) early weekdays, 9 a.m. to 2 p.m.; (1b) late weekdays, 2 p.m. to 6 p.m.; (2) early evening, 6 p.m. to 7:30 p.m.; (3) late evening, 7:30 p.m. to 9 p.m.; (4) Saturday, 10 a.m. to 6 p.m.; (5) Sunday, 2 p.m. to 9 p.m. The initial strategy consisted of a total of 14 calls if there was no contact with a person:

- four calls consisting of an early or late day, early evening, late evening, and weekend (either Saturday or Sunday), in any order;
- 1 week wait;
- three calls consisting of an early evening, late evening, and the weekend day not called in the preceding four calls, in any order;
- 1 week wait;
- four calls consisting of a an early or late day (whichever was not called in the first set), early evening, late evening, and weekend (either Saturday or Sunday), in any order;
- 1 week wait; and
- three calls consisting of an early evening, late evening, and the weekend day not called in the preceding 4 calls, in any order.

The waiting times were compressed later in the field period.

If, after these 14 calls, there was still no contact, the telephone number was retired by coding it NA (all no answer or busy) or NM (at least one answering machine, but no "live" contact). Telephone numbers with no contact (including no answering machine contact) after the first 4 calls were sent to a vendor for further calling. This vendor used a predictive dialer, so that no operator (interviewer) was necessary for calls that were not answered by a live person. The vendor used the time slice strategy described above for the final 10 calls, and the same set of result codes to record the outcome. If a call was answered by a live person, an operator would come on the line and ask whether the number was for business or household use. Numbers with answered calls were returned to Westat for further follow-up. The operator's script did not mention AHS specifically.

4.2.3 Maximum Call Limits

When a person answered the telephone, the telephone number was removed from the time slice strategy described above. Once contact was made, all subsequent calls were based upon the respondent's assessment of the best time to call or it was left to the interviewer to suggest the best time. This was generally in terms of an exact appointment or a general "best time" to call (e.g., day, evening, or weekend). The maximum call counter for these cases for both the screener and the extended interview was set at 23 each. This limit was set to allow enough calls for two refusal conversion efforts and calls in Spanish. As a result, only 2 percent of the sample telephone numbers that were called ended as "maximum calls" (MC or LM) at the screener level (see Table 4-1). Some cases active until the end of the field period received maximum call codes without reaching the call limit.

At the adult extended level, about 10 percent of cases (Table 4-2) received one of the "maximum call" codes—MC, LM/ML (maximum calls where the number was coded a language problem at some point), MR (maximum calls where a refusal was encountered at some point), and MT (maximum calls where we were given a different telephone number to reach the adult respondent). About 13.5 percent of child interviews (Table 4-3) and 8.5 percent of adolescent interviews (Table 4-4) were in these categories.

4.2.4 Language Strategy

An important capability for AHS was conducting interviews in Spanish. Section 2.3.3 of this report describes the process by which the questionnaires were translated, and Section 3.3 describes the recruitment and training of interviewers bilingual in English and Spanish. This section describes how the Spanish-language interviews were managed in the CATI system.

Sampled telephone numbers with names and addresses from directory services were matched against a list of Hispanic surnames. A total of 2,719 matching numbers were loaded into the Spanish work class and worked only by bilingual interviewers. About 24 percent of these cases resulted in completed screeners, with a 35 percent cooperation rate, about 2 points lower than the overall screener cooperation rate. Sixty percent of the screeners completed with this group were done in Spanish. Of those that completed the adult interview, 85 percent reported being of Hispanic or Latino origin, and they represented 44 percent of the total number of AHS adult respondents reporting Hispanic/Latino ethnicity.

The remainder of the sample numbers were loaded into the default CATI work class, which meant that they were available to any interviewer. Cases determined to require a Spanish bilingual interviewer were put into the Spanish-language work class, and became available to bilingual interviewers. In all of the completed adult interviews from this group, the respondent reported being Hispanic/Latino; these represented 22 percent of the total in AHS. The remaining 34 percent of the self-reported Hispanics/Latinos were from cases worked only by English-speaking interviewers, and they were about 5 percent of the total number of adult interviews completed outside of the Spanish-language work class.

Overall, some 275 adult interviews were conducted in Spanish, about 6.5 percent of the total. A slightly higher percentage were conducted in Spanish in the remainder of Arizona than in Maricopa County.

The mean interview times differed by language. The adult interview was about 29 minutes in English, 42 minutes in Spanish. The child interview averaged 14 minutes in English, 19 minutes in Spanish, and the adolescent 18 minutes in English and almost 25 minutes in Spanish.

4.2.5 Refusal Conversion

At each stage of the interview process, Westat interviewers made extensive conversion efforts for refusals that were not judged to be hostile or abusive. When a refusal was first encountered, the interviewer would attempt to be persuasive but not pushy, and would document the nature of the refusal. Initial refusals were moved to the refusal work class and generally not called again for 2 weeks. An exception for screener refusals was that telephone numbers designated as "no conversion" were considered final after the initial refusal. Initial refusals that were considered hostile or abusive also received a final result code. Interviewers selected and trained for refusal conversion as described in Chapter 3 would call back all but those categories. Second conversion attempts were made for screener refusals.

4.2.6 Proxy Interviews

SLHI decided to allow proxy reporting for sample persons over 65 who were unable to respond for themselves because of physical, mental, or emotional limitations. Proxy respondents had to be adult members of the household knowledgeable about the sampled adult's health. Some 65 candidates for proxy interviews were identified based upon interviewers' notes; of these, 35 interviews were completed with proxies, and another 5 were completed with the sampled adults themselves.

Interviewers who conducted the proxy interviews were trained to substitute the name of the sampled adult or an appropriate pronoun wherever "you" appeared in the question text; in cases where "you" referred specifically to the respondent (e.g., "You said earlier . . ."), the word "you" was highlighted for the proxy interviews.

4.3 Detailed Results by Outcome

This section will present detailed tables of outcomes at each interview level – screener, adult, child, adolescent permission, and adolescent interview. Interviewers assign a result code to each attempt to reach a sampled telephone number. The telephone result codes are divided into interim (numeric) and final (alpha) codes. During data collection, each case is tracked according to its most recent result code. Cases with interim codes are typically managed automatically by the scheduler according to preset parameters, such as how to work through

"time slices" (see Section 4.3.2) and how long to wait before re-contacting an initial refusal. Problem cases require manual intervention before they are re-fielded.

At the end of the field period, all remaining interim cases were assigned final result codes according to their call history. Many cases for which some contact had been made received codes beginning with "M" (maximum calls), with the actual designation depending on what else had happened during their call history.

Tables 4-1, 4-2, 4-3, and 4-4 present the complete final result code dispositions, by sample, for the screener, adult, child, and adolescent interviews, respectively. The following sections discuss these results by instrument.

4.3.1 Screening Interview

Table 4-1 shows the screening interview outcomes, by stratum, for all sampled numbers that were called. Some proved to be ineligible or out of scope. Eligibility criteria were quite limited; only 22 cases were determined to be ineligible during the screener (e.g., no one 18 or older in the household). Overall, 20 percent of the numbers called were determined to be out of scope – either businesses or non-working.

Completed screeners accounted for almost 16 percent of numbers called in the Maricopa County sample, and 17 percent in the remainder Arizona. Refusals accounted for about 28 percent of Maricopa County numbers and 26 percent in the remainder of Arizona. The raw cooperation rate, or completed screeners divided by completed screeners plus refusals, was about 36 percent for Maricopa County and 40 percent for the remainder of Arizona. Sixty-five percent of the numbers called were designated for refusal conversion. Among those, the cooperation rates (not shown in the table) were 42 percent for Maricopa County and 46 percent for the remainder of Arizona. Among the no conversion cases, the cooperation rates were 25 percent and 27 percent, respectively.

Other final screener results include noncontacts (with and without answering machines) and "other nonresponse." Noncontacts accounted for more than 18 percent of numbers called in Maricopa County and 14 percent in the remainder of Arizona. Other nonresponse, including language problems and "maximum call" cases (those with at least one contact but no definitive outcome) accounted for just over 2 percent of all cases worked.

Table 4-1. Detailed results of AHS data collection, screening interview, by stratum

	MARICOPA COUNTY			REMAINDER OF ARIZONA			TOTAL		
	Number	Percentage		Number	Percentage		Number	Percentage	
		Within category	of Total		Within category	of Total		Within category	of Total
CS – COMPLETED SCREENER (C)	5,760		15.6%	1,780		17.1%	7,540		15.9%
INELIGIBLE	20		0.05%	2		0.02%	22		0.05%
Out of Scope									
NR – NON-RESIDENTIAL PHONE NUMBER	2,321	31.7%		499	22.7%		2,820	29.6%	
NW – NON-WORKING PHONE NUMBER	4,994	68.3%		1,702	77.3%		6,696	70.4%	
Total Out of Scope	7,315		19.8%	2,201		21.2%	9,516		20.1%
Noncontact									
NA – NO CONTACT MADE AFTER TIME SLICES FILLED	3,555	52.3%		740	49.4%		4,295	51.8%	
NM – NO CONTACT – REACHED ANSWERING MACHINE	3,238	47.7%		757	50.6%		3,995	48.2%	
Total Noncontact	6,793		18.4%	1,497		14.4%	8,290		17.5%
Refusal (R)									
R1: NO SCREENER REFUSAL CONVERSION	4,154	40.7%		1,098	40.8%		5,252	40.7%	
R3 – FINAL REFUSAL – RECEIVED 3 OR MORE 2S	1,128	11.0%		306	11.4%		1,434	11.1%	
RB – FINAL REFUSAL	3,487	34.1%		828	30.7%		4,315	33.4%	
RM – REFUSAL REACHED MAXIMUM CALL LIMIT	1,265	12.4%		417	15.5%		1,682	13.0%	
RX – RE-RELEASED RB REACHED MAX CALL LIMIT	183	1.8%		45	1.7%		228	1.8%	
Total Refusal	10,217		27.6%	2,694		26.0%	12,911		27.2%
Other Nonresponse									
LH – HEARING AND SPEECH PROBLEM	15	1.9%		4	1.7%		19	1.9%	
LM – LANGUAGE PROBLEM REACHED MAX CALLS	76	9.7%		13	5.5%		89	8.7%	
LP – FINAL LANGUAGE PROBLEM	70	8.9%		12	5.0%		82	8.0%	
MC – MAXIMUM CALLS	595	76.0%		204	85.7%		799	78.3%	
NO – OTHER NON-RESPONSE	27	3.4%		5	2.1%		32	3.1%	
Total Other Nonresponse	783		2.1%	238		2.3%	1,021		2.2%
TOTAL	37,011		100.0%	10,381		100.0%	47,392		100.0%
ELIGIBILITY RATE (C / (C+I))			99.95%			99.98%			99.95%
COOPERATION RATE ((C+I) / (C+I+R))			36.1%			39.8%			36.9%

4.3.2 Adult Extended Interview

The number of completed screeners becomes the total number of cases available for the adult extended interview. The results of data collection efforts for the adult extended interview are shown in Table 4-2. Adult extended interviews were completed for about 56 percent of sampled adults, with the completion rate about 5 points higher for the remainder of Arizona than for Maricopa County. The AHS team decided that it would use data from partially completed adult interviews, so long as the interview went at least through Section K. Fewer than 1 percent of all adult interviews counted as complete were only partially done (CP).

Refusals accounted for almost 30 percent of sampled adults, and other nonresponse almost 14 percent. The adult interview cooperation rate was 65 percent, and was about 4 points higher in the remainder of Arizona than in Maricopa County.

4.3.3 Child Extended Interview

Child interviews were attempted only for Maricopa County. In some cases children were sampled in the screener, and the child interview could be completed before the adult interview. In other cases, children were sampled in the adult interview. The child interview was done with the parent in the household who knew the most about the child's health and health care – the "most knowledgeable adult." As shown in Table 4-3, interviews were completed for 74 percent of sampled children, with about 11 percent of cases resulting in refusal and 15 percent in other nonresponse. The child interview cooperation rate was 87.5 percent.

Forty-seven percent of all child interviews were completed using the "child first" procedure, that is, the child was sampled before the adult interview. For 60 percent of these, and for 28 percent of child interviews overall, no adult interview was completed in the household.

Table 4-2. Detailed results of AHS data collection, adult extended interview, by stratum

	MARICOPA COUNTY			REMAINDER OF ARIZONA			TOTAL		
	Number	Percentage		Number	Percentage		Number	Percentage	
		Within category	of Total		Within category	of Total		Within category	of Total
Completed Interviews									
CA – COMPLETED ADULT EXTENDED	3,125	99.6%		1050	99.3%		4,175	99.5%	
CP – ADULT PARTIAL COMPLETE – FINISHED	14	0.4%		7	0.7%		21	0.5%	
Total Completed Interviews	3,139		54.5%	1057		59.4%	4,196		55.6%
Ineligible									
IA – INELIGIBLE AGE FOR ADULT EXTENDED	2	28.6%		0			2	28.6%	
IO: INELIGIBLE OUT OF STATE	5	71.4%		0			5	71.4%	
Total Ineligible	7		0.1%	0		0.0%	7		0.1%
Out of Scope									
OE – OUT OF SCOPE ENUMERATION ERROR	42		0.7%	14		0.8%	56		0.7%
Refusal									
R3 – FINAL REFUSAL RECEIVED 3 OR MORE 2S	5	0.3%		1	0.2%		6	0.3%	
RB – FINAL REFUSAL	1,202	68.3%		362	73.4%		1,564	69.4%	
RM – REFUSAL REACHED MAXIMUM CALL LIMIT	554	31.5%		130	26.4%		684	30.3%	
Total Refusal	1,761		30.6%	493		27.7%	2,254		29.9%
Other Nonresponse									
LH: LANGUAGE PROBLEM HEARING/SPEECH	23	2.8%		9	4.2%		32	3.1%	
LM: LANGUAGE PROBLEM REACHED MAX CALLS	19	2.3%		7	3.2%		26	2.5%	
LP: FINAL LANGUAGE PROBLEM	10	1.2%		6	2.8%		16	1.6%	
MC – MAXIMUM CALLS	344	42.4%		80	37.0%		424	41.3%	
ML – MAXIMUM CALLS – SCRNRSLT PROB IN HH	59	7.3%		24	11.1%		83	8.1%	
MR – MAXIMUM CALLS – REFUSAL IN HH	180	22.2%		39	18.1%		219	21.3%	
MT – MAXIMUM NUMBER OF CALL ATTEMPTS	0	0.0%		1	0.5%		1	0.1%	
ND – RESPONDENT DECEASED	2	0.2%		1	0.5%		3	0.3%	
NF -- NOT AVAILABLE IN FIELD PERIOD	17	2.1%		7	3.2%		24	2.3%	
NL: NOT LOCATABLE THROUGH TRACING	116	14.3%		25	11.6%		141	13.7%	
NS – SUBJECT SICK/INCAPACITATED	40	4.9%		17	7.9%		57	5.6%	
NW -- NON-WORKING PHONE NUMBER	1	0.1%		0	0.0%		1	0.1%	
Total Other Nonresponse	811		14.1%	216		12.1%	1,027		13.6%
TOTAL	5,760		100.0%	1,780		100.0 %	7,540		100.0%
ELIGIBILITY RATE			99.9%			100.0 %			99.9%
COOPERATION RATE			64.1%			68.2%			65.1%

Table 4-3. Detailed results of AHS data collection, child extended interview, by stratum

	MARICOPA COUNTY		
		Percentage	
	Number	Within category	of Total
Completed Interviews			
CC – COMPLETED CHILD EXTENDED	643		74.2%
Ineligible			
IC – INELIGIBLE AGE FOR CHILD EXTENDED	3		0.3%
Out of Scope			
OE – OUT OF SCOPE ENUMERATION ERROR	1		0.1%
Refusal			
R3 – FINAL REFUSAL, RECEIVED 3 OR MORE 2S	1	1.1%	
RB – FINAL REFUSAL	55	59.8%	
RM – REFUSAL REACHED MAXIMUM CALL LIMIT	36	39.1%	
Total Refusal	92		10.6%
Other Nonresponse			
MC – MAXIMUM CALLS	53	41.4%	
ML – MAXIMUM CALLS – LANGUAGE PROB IN HH	18	14.1%	
MR – MAXIMUM CALLS – REFUSAL IN HH	45	35.2%	
MT – MAXIMUM NUMBER OF CALL ATTEMPTS	1	0.8%	
NF – RESPONDENT NOT FOUND AT CALL BACK	2	1.6%	
NL – NOT LOCATABLE THROUGH TRACING	9	7.0%	
Total Other Nonresponse	128		14.8%
TOTAL	867		100.0%
COOPERATION RATE			87.5%

4.3.4 Adolescent Extended Interview

As with child interviews, those with adolescents were only attempted in Maricopa County, and adolescents could be sampled either in the screener or in the adult interview. An additional step for adolescents was to get permission from a parent before attempting to contact them. As shown in Table 4-4, 500 adolescents were sampled, and permission was obtained for about 61 percent. Interviews were completed with 67 percent of those. Another 20 percent refused the interview, and 13 percent were classified as other nonresponse. The cooperation rate among adolescents was 77 percent.

Table 4-4. Detailed results of AHS data collection, child extended interview, by stratum

	MARICOPA COUNTY		
		Percentage	
	Number	Within category	of Total
Completed Interviews			
CT – COMPLETED ADOLESCENT EXTENDED	204		67.1%
Ineligible			
IT – INELIGIBLE AGE FOR ADOLESCENT EXTENDED	1		0.3%
Refusal			
RB – FINAL REFUSAL	50	83.3%	
RM – REFUSAL REACHED MAXIMUM CALL LIMIT	10	16.7%	
Total Refusal	60		19.7%
Other Nonresponse			
MC – MAXIMUM CALLS	9	23.1%	
ML – MAXIMUM CALLS – LANGUAGE PROB IN HH	5	12.8%	
MR – MAXIMUM CALLS – REFUSAL IN HH	12	30.8%	
NF – NOT AVAILABLE IN FIELD PERIOD	3	7.7%	
NL – NOT LOCATABLE THROUGH TRACING	7	17.9%	
NS -- SUBJECT SICK/INCAPACITATED	3	7.7%	
Total Other Nonresponse	39		12.8%
TOTAL	304		100.0%
COOPERATION RATE			77.3%
ADOLESCENTS SAMPLED	500		
PERMISSION NOT RECEIVED	196		39.2%
COMBINED COMPLETION RATE			40.8%

4.4 Response Rates

4.4.1 Defining Response Rates

The term "response rate" is used in many different ways across surveys and organizations so its careful definition is important. Two organizations that describe response rates in a relatively consistent manner are the Council of American Survey Research Organizations (CASRO, 1982) and the American Association for Public Opinion Research (AAPOR, 2006). The AAPOR report is periodically updated and is available on the organization's website (http://www.aapor.org).

We use definitions described in the AAPOR report, which includes several different ones. Among them are the RR4 and RR3 definitions that are most commonly accepted in the current survey research field. The only difference between the two is that RR3 does not include partial completes while RR4 does. This report uses AAPOR's RR4. Since telephone numbers were sampled with different selection probabilities in the two AHS strata, we use the weighted number of telephone numbers rather than the unweighted number in the response rate computation.

Both AAPOR and CASRO recommend that a survey response rate be defined as the ratio of completed interviews to eligible reporting units (i.e., residential households). This recommendation is more difficult to apply than it may appear, especially in RDD surveys, as determining the eligibility of some sampled numbers is problematic; because some telephone numbers, even after being called multiple times over a range of days and times of day, are never answered or are picked up only by answering machines. The eligibility of these numbers cannot be determined directly, adding ambiguity to the definition of a response rate.

The proportion of unanswered telephone numbers that are eligible (or residential) is denoted as '*e*' in the AAPOR RR4 equation. Once the eligibility proportion is established, the response rate can be computed as the weighted ratio of the responding telephone numbers to the total of known and estimated eligible numbers. One of the first approaches used for estimating *e* was suggested in CASRO (1982). CASRO estimates *e* as the proportion of the resolved telephone numbers that are residential, which is the approach taken here.

Further complications in calculating the response rates arise from the subsampling of non-mailable cases and of screener refusals for conversion described in Chapter 1. These are handled in the calculation through weighting adjustments. The adjustment for refusal subsampling allocates the weights of all households assigned to the "no conversion" treatment to those selected for conversion.

4.4.2 Screener Response Rate

A screener response rate is calculated for each sample stratum and for the state as a whole. The formula for the screener response rate (rr_S) in a sample stratum is

$$rr_s = \frac{\sum_{i \in S_{resp}} w_i}{\sum_{i \in S_{resid}} w_i}, \quad (1)$$

where w_i is the weight for household i in the stratum after adjusting for differential sampling rates, refusal conversion subsampling, and the assignment of households with unknown residential status; S_{resp} is the set of households in the stratum that responded to the screening interview; and S_{resid} is the set of households in the stratum that were residential.

The screener response rate for the state is computed in exactly the same way, except the sum is over the whole state rather than in the specific stratum. Thus, the state screener response rate is a weighted average of the two stratum screener response rates where weights are equal to the population in the strata. As a result, the state response rate differs from what would be obtained from the unweighted average of the response rates of the strata.

Table 4-5 presents the screener response rates by stratum, and by whether a mailable address was obtained for the sampled telephone number. The rate is about 9 points higher overall for cases with mailable addresses. There are two primary reasons for this difference: first, "mailable" cases are much more likely to be contacted and prove to be households; and second, there is very likely some benefit in increased cooperation from the letters mailed to those cases where it was possible. We also see about a 3 point higher response rate in the remainder of Arizona than in Maricopa County. Typically, large urban areas are the most difficult in which to complete survey interviews.

Table 4-5. AHS Screening Interview response rates by stratum and whether a mailing address was obtained

	Response Rate		
Stratum	All Cases	With Mailing Address	Without Mailing Address
Maricopa County	35.3%	39.8%	30.0%
Remainder of Arizona	38.7%	42.0%	34.1%
State Total	36.6%	40.7%	31.5%

4.4.3 Adult Extended Interview Response Rate

The extended response rate for the adult interview is the weighted percentage of the adults sampled in the screener who completed the adult extended interview. The weight in this case is the inverse of the probability of selecting the adult within the household. Because of this weighting, adults sampled from households with more than one adult have a larger effect on the response rate than those in households with only one adult. The extended adult response rate (rr_a) is

$$rr_a = \frac{\sum_{i \in A_{resp}} w_i'}{\sum_{i \in A_{eligsamp}} w_i'}, \tag{2}$$

where the numerator is summed over all adult respondents, and the denominator is summed over all eligible sampled adults. The weight being summed in this case, w', is the adult weight that accounts for selecting the adult within the household. The adult response rate is conditioned on the completion of the screener interview.

Table 4-6 shows the conditional response rates for the adult extended interview, with an overall rate of 52.3 percent. As with the screener, the rates are higher in the remainder of Arizona than in Maricopa County (about 5 percentage points), and higher for households with mailing addresses than for those without. The difference by mailing status, at about 3 points, is much less than for the screener.

Table 4-6. AHS Adult Interview response rates by stratum and whether a mailing address was obtained

Stratum	Response Rate		
	All Cases	With Mailing Address	Without Mailing Address
Maricopa County	50.1%	51.3%	48.3%
Remainder of Arizona	55.4%	57.1%	52.7%
State Total	52.3%	53.7%	50.1%

The adult response rate defined above is conditional in that it depends on the household participating in the screener. We calculate overall response rates to eliminate the conditioning. Since the adult response rate is conditioned only on the completion of the screener,

the product of the screener and adult response rate is an unconditional or overall adult response rate. Thus, the overall adult response is

$$orr_a = rr_s \cdot rr_a \tag{3}$$

Table 4-7 includes the overall response rates for the adult interview, with the rate for all cases being 19.2 percent. The patterns by stratum and mailing status are similar to those for the screener and conditional adult rates.

Table 4-7. AHS overall adult response rates by stratum and whether a mailing address was obtained

Stratum	Response Rate		
	All Cases	With Mailing Address	Without Mailing Address
Maricopa County	17.7	20.4	14.5
Remainder of Arizona	21.4	24.0	18.0
State Total	19.2	21.8	15.8

4.4.4 Child Extended Interview Response Rate

The extended response rate computation for children is similar to the adult procedure; however, the child-first procedure adds some complexity. Persons under 18 years of age could be enumerated either at the end of the screener (the child-first procedure) or during the adult extended interview. In the latter case, the child and adolescent extended response rates include only those households in which the adult extended interview is completed, and the child response rate is conditional on the adult interview. If the child first procedure was implemented, then the child response rate is conditioned only on the screener. The extended child response rate (rr_C) is

$$rr_c = \frac{\sum_{i \in C_{resp}} w_i''}{\sum_{i \in C_{eligsamp}} w_i''}, \tag{4}$$

where the numerator is summed over all child respondents, and the denominator is summed over all eligible sampled children. The weight being summed in this case, w'', is the inverse of the probability of selecting the child within the household. To discriminate between the different

sampling situations we add a subscript *K* to identify the procedure; $rr_{c,K}$ is the child extended interview response rate for children who were interviewed using the child-first procedure, and $rr_{c,\bar{K}}$ is the child extended interview response rate for children who were interviewed without using the child-first procedure.

The child response rate is conditioned on the screener being completed and either the child interview being completed for households using the child-first procedure or the adult interviews being completed for others. The overall response rate for the child, orr_c, is defined as

$$orr_c = rr_s \cdot \left(p_K \cdot rr_{c,K} + p_{\bar{K}} \cdot rr_{ac,\bar{K}} \cdot rr_{c,\bar{K}} \right) \tag{5}$$

where $rr_{ac,\bar{K}}$ is the extended adult interview response rate for adults in households with children who were sampled without using the child-first procedures, and p_{Kc} and $p_{\bar{K}c}$ are the proportions of households with children in which the child-first procedures were used or not, respectively (i.e., $p_{Kc} + p_{\bar{K}c} = 1$).

Table 4-8 shows the conditional and overall response rates for the child extended interview, for Maricopa County only. The difference by mailing status for the conditional rate, at about 15 points, is substantially larger than in the adult interview.

Table 4-8. AHS Child Interview response rate and combined response rate by whether a mailing address was obtained (Maricopa County only)

	Response Rate		
	All Cases	With Mailing Address	Without Mailing Address
Child Interview rate	73.9%	81.1%	66.2%
Overall rate	26.1%	32.3%	19.9%

4.4.5 Adolescent Extended Interview Response Rate

The adolescent response rates are calculated in exactly the same way as those for the child interview. However, an important source of nonresponse for the adolescent interview is the parent refusing to provide permission to conduct the interview with the adolescent. As we saw in Table 4-5, about 40 percent (unweighted) of sampled adolescents did not have permission granted.

Table 4-9 shows the conditional (36.1 percent) and overall (12.8 percent) response rates for the adolescent extended interview, incorporating nonresponse from both the adolescent him- or herself and from the permission-giving parent or guardian. The difference by mailing status (3 points for the conditional rate) is about at the same level as for the adult interview.

Table 4-9. AHS Adolescent Interview response rate and combined response rate by whether a mailing address was obtained (Maricopa County only)

	Response Rate		
	Overall	With Mailing Address	Without Mailing Address
Child Interview rate	36.1%	37.4%	34.4%
Combined rate	12.8%	14.9%	10.3%

5. DATA PROCESSING

5.1 Data Editing Procedures

Survey data for AHS were collected using a computer-assisted telephone interview (CATI) system. In a CATI environment, the data collection and interview process is controlled using a series of computer programs designed to ensure consistency and quality. The CATI system programming determines which questions are asked based on household composition, respondent characteristics or preceding answers, and the order in which the questions are presented to interviewers. The system also presents the response options that are available for recording respondents' answers.

CATI range and logic edits do much to help ensure the integrity of the data during collection. This editing at the time of the interview greatly reduces the need to recontact respondents to verify responses and allows questionable entries to be reviewed in real time with the respondent as part of the collection process. Although the CATI system virtually eliminates out-of-range responses and many other anomalies, some consistency and edit issues may arise. For example, interviewers may note concerns or problems that must be handled by data preparation staff after the interview is complete. Updating activities require that both manual and machine editing procedures be developed to correct interviewer, respondent, and CATI program errors and to check that updates made by data preparation staff were input correctly. Because data editing resulted in changes to the survey data, specific quality control procedures were implemented. AHS survey data were carefully examined and edited before delivering final data files to SLHI. Quality control procedures involved limiting the number of staff who made updates, using the CATI specifications to resolve issues in complex questionnaire sections, carefully checking updates, and performing computer runs to identify inconsistencies or illogical patterns in the data.

The data editing procedures for AHS consisted of four main tasks: (1) managing and resolving problem cases, (2) reading and using interviewer comments to make data updates, (3) coding questions with text strings (i.e., "other specify" responses), and (4) verifying data editing updates. The final step was to convert the edited data from the CATI system to the SAS data delivery files. The sections below describe each of these processes in turn.

5.1.1 Resolving Problem Cases

The data preparation staff, as well as project staff and staff from the Telephone Research Centers (TRCs), worked collectively to resolve problem cases. In this section, the method interviewers used to communicate problems is described, along with the system used by data editing and preparation staff to update or modify the data.

An interviewer who experienced a problem while working a case during data collection could alert the project team in one of two ways. One method was to fill out a problem sheet for the case. Problem sheets from all the TRCs were sent via e-mail to a single staff member who distributed them to the appropriate department or project staff person. Data preparation staff often used these problem sheets as a guide to review cases and to make certain that any required updates were made accurately.

The second method of communicating problems was to assign a specific result code to cases within the CATI system, obviating the need for a problem sheet. The problem result code category had three sub-categories for special queues to which these problem cases could be assigned for review. These sub-categories were used to indicate the person responsible for investigating the case further—TRC staff, project staff, or data processing staff. Problem cases were reviewed electronically by a TRC supervisor and either re-fielded to the interviewing staff or distributed to the appropriate TRC, data processing, or project staff.

Database updates were unnecessary for some problems, and these cases could simply be released for general interviewing accompanied by an appropriate message. If, for example, an adult extended interview was stopped during the middle of Section C, the interviewer would enter a detailed comment explaining why the case could not proceed (e.g., "Respondent wanted to change several answers. I was unable to back up properly"). The solution for these types of cases was to re-field the interview with a message stating, "Case will restart in Section C. Re-ask beginning with screen AE24."

Most restart cases were made available to the general interviewing staff. For unusual or complex problems, the case could be assigned to a specific interviewer with experience in handling these types of problems. Some examples of cases reviewed by project staff were those in which an error was made in enumerating a household member or when a change in the person named as most knowledgeable about the sampled child was needed.

5.1.2 Interviewer Comments

Another important data editing task was reading and using interviewer comments. Comments are text phrases interviewers enter in CATI when they want to record respondents' statements but are unable to enter as a standard response in the instrument. Sometimes these phrases were merely an elaboration of a previously recorded response, an expression of opinion, or comments unrelated to the survey, which did not necessarily require modifying or updating survey responses. Other times, comments were substantive to data quality and indicated that an update was needed.

Comments were also used to identify specific responses that interviewers felt could not be coded using the existing response option set. Data preparation staff resolved all of these situations for AHS without the need to add any new response codes.

5.1.3 Coding with Text Strings

Most items in AHS had only closed-ended response options, so coding of open-ended responses was not needed. The survey had several other-specify questions, however, that required coding of narrative text strings recorded by interviewers. Other-specify questions had specific response categories but also allowed for text or values to be typed into an "other" category. AHS questions with an "other" category from the adult extended interview included;

- racial identification (AA5)
- country of birth (AH33),
- first language (AH36),
- place visited for health care (AH3),
- health insurance coverage items (AI15, AI45), and
- child/adolescent health insurance coverage items (CF2A, CF18, IA2A, IA18).

Westat data preparation staff reviewed these responses and up-coded them to the existing categories whenever possible. For two of the items (AA5 and AH3), some "other" responses were standardized. At AA5, responses such as Mexican, Latino, and Hispanic, were updated to HISPANIC-LATINO. (See Section 5.3 for more on coding race and ethnicity.) At AH43, a few

responses such as chiropractor, naturopath, and herbologist were updated to "Alternative/Complementary Medicine."

AHS did not collect open-ended responses that required a specially developed coding scheme or structure. Some survey items, however, collected amounts or values such as the respondent's age or income. For such items, the CATI system utilized "soft-" and "hard-range" edits, which are documented in the CATI specifications.

Soft-range edits were activated during the interview when the respondent gave an unlikely response (a value outside the specified range). The CATI system responded by placing a message on the screen and required the interviewer to re-enter the response. This system feature gave interviewers an opportunity to verify that the response was recorded accurately or re-ask the question to be certain the respondent understood what was being asked as needed. Hard-range edits prevented recording unacceptable values. For example, for a question on how many times in the past week the respondent ate fruit, the soft range was 0-35, the hard range 0-70. In practice, 3 responses fell outside the soft range for this item and were verified.

In circumstances when the respondent insisted on giving a response that violated the hard-edit specifications, interviewers recorded the respondent's answer in the comment field and data preparation staff reviewed and updated the case as needed.

The AHS child and adolescent interviews included an item that collected the name of the school attended by the selected child or adolescent (CB22 and TA4B, respectively). Interviewers recorded the respondent's answers as a verbatim text entry. There were a number of spelling problems with these entries. In many cases, an English-speaking adult respondent had difficulty reporting a Spanish school name. Hispanic respondents whose first language was not English had similar difficulties in accurately reporting or spelling school names. Westat data preparation staff used the Arizona Public School Directory in conjunction with the respondent's Zip code to update school names.

5.1.4 Verifying Data Updates

Updates to the original interview data were required due to a variety of circumstances as described above. Generally speaking, data updates and corrections were made to account for these situations including interviewer and respondent error, information captured in

comments and "other-specify" fields, and problem sheets, so that the final survey data reflected the most accurate information possible.

A series of techniques verified that survey updates were made accurately. First, the intended updates were recorded on a hard-copy printout or on an associated problem sheet. The CATI case identification number was also recorded to ensure that updates were associated with the appropriate case. This printout was checked for accuracy and for logical effects on any other questions or skip patterns in the questionnaire. Next, the updates were entered into the computer and verified again – matching the resulting information against the print-out. For more complicated circumstances, the data preparation staff carefully reviewed interviewer comments, messages, and problem descriptions to verify data updates.

An entry in an electronic transaction journal was created for each data update. Transaction journal entries maintained information such as the CATI case identification number, initial data value(s), the updated value(s), and the date that the update was made. The editing and verification process was performed throughout the data collection period; 2,753 database values were updated and verified for AHS, with school name the most frequently updated item.

Cases with similar problems were reviewed together and then updated at one time in manageable batches. This process ensured consistency in the handling of discrete data problems. Following the series of updates, a program checked for the full set of errors that had been identified to date to ensure that data editing had not created any new errors. Frequency distributions and cross-tabulations of survey variables were used extensively by data preparation staff to verify data updates.

Structural edits designed to assess the integrity of the CATI database (i.e., verifying that all database records that should exist actually do exist, and those that should not exist do not), and, as necessary, edits that evaluated complex skip patterns, were run periodically during data collection.

5.2 Data Conversion and Delivery

The final survey data were delivered to SLHI formatted as SAS data sets, and also as Excel files. The SAS data sets were created by converting the CATI database using a series of SAS macro programs. Initially, the CATI survey data were stored in a hierarchical database to

improve data efficiency and enhance performance while interviewing. This conversion was accomplished using Westat's CATISAS macro program that extracts information stored in the CATI data dictionary (e.g., variable names, variable labels, allowable values, and formats) and then converts each of the CATI database segments into a "flat" SAS data set. Using the CATI data dictionary to define the SAS data set variables is advantageous because variables are stored in questionnaire order, allowing for meaningful presentation of the variables in frequency output and file listings without additional programming. SAS data sets created by the macro were later combined to facilitate processing and file delivery. After the survey data were converted from the CATI system, all further processing relied on the SAS system.

During the conversion process from CATI database elements to SAS files, diagnostic edit checks were run on the entire database. Frequencies for categorical data were also run and examined. These reviews were made to ensure that errors had not been inadvertently introduced into the data (i.e., no data were lost, no unexpected shifts in variable distributions occurred). In going from the CATI to SAS file organization, for example, frequency runs from the CATI database and the post-CATI SAS files were compared.

5.3 Race and Ethnicity Coding

The AHS survey items about Hispanic ethnicity (AA4) and race (AA5A) were consistent with those in the 2000 Decennial Census. This section describes how we handled situations when the respondent reported a race that was not classified into one of the pre-existing categories. These responses were recorded in the "other specify" category as a text string.

Item AA5A asked respondents for their race: "Please tell me which one or more of the following you would use to describe yourself. Would you describe yourself as Native Hawaiian, Other Pacific Islander, American Indian, Alaska Native, Asian, Black, African American, or White?" The race question allowed the respondent to indicate that they belonged to any or all of the coded races and also to say "other" race. The "other specify" race was recorded in text (AA5AOS). The AHS procedures for coding these responses were consistent with the ones used to code the 2000 Census data documented in *Census 2000 Redistricting Data (Public Law 94-171) Summary File – Technical Documentation* (U.S. Census Bureau, 2001) available at http://www.census.gov/prod/www/abs/pl94-171.pdf. The specific sections of interest are in Appendix B, pages B-2 and B-3. When we refer to the Census procedures, we mean our interpretation of the information in this document.

An initial review of cases showed that the largest group of cases with "other race" categories were ones in which the respondent identified as being Hispanic or Latino and did not identify with any pre-coded race categories. The most common responses to the "other race" were those such as "Hispanic," "Latino," or "Mexican." Following the Census procedures, the person was left in the "other race" category. The "other specify" text was updated to "Hispanic/Latino." Other procedures included the following:

- If the "other specify" text clearly should have been included in an existing code (following the Census procedures), then it was up-coded and removed from the "other" category. For example, if the respondent was coded only as other race and the "other specify" was "Irish," then the code for "white" was upcoded to "yes," other race was revised to "no" and the other specify text eliminated.

- If the "other specify" text did not fit into an existing code (following the Census procedures), then it was left in the "other" category with the existing text in the "other specify." For example, if the "other specify" text for race was "Indian" and no other race category was identified, then no changes were made in the responses.

- If the respondent reported being Hispanic or Latino at AA4, we never revised this code based upon information in the other specify comments of the other variables. For example, if the person was coded as "Hispanic" and the specific Hispanic origin item was only coded as "other" with the text "Jewish," then the Hispanic code was not altered.

- If the respondent was not coded as being Hispanic or Latino in AA4 but the text in the "other specify" field for race indicated they were Hispanic or Latino, then the Hispanic or Latino coding was revised to "yes."

- If the "other race" text was similar to "none of above," we left the response as it was.

- If the "other race" text was similar to "human race," we coded this as a refusal. The value for race was then imputed along with other cases that were more direct refusals.

The Census procedures clearly state that persons who say they have European, Middle Eastern, or North African origin are to be classified as "White" race. This rule has many implications. For example, suppose a person says they are not Hispanic and only identify the "other race" as being Spain. We would upcode Hispanic origin to "yes" (to be consistent with the Census procedures for Hispanic origin) and then upcode "race" to "White" (since the person is of European origin).

5.4 Geographic Coding

The AHS adult extended interview asked all respondents the name of the county where they lived: “To be sure we are covering the entire state, what county do you live in?” (AH42). In addition, for cases in which an address had been matched to the sampled telephone number[3], interviewers verified the street address and Zip code with the adult respondent (AO1) and then collected the name of a nearby cross-street (AM9). These same questions were asked of adults who completed the child interview under the “child first” protocol.

If there was no matched address for a given case, respondents were asked to provide their Zip code (AM7), their street address (AO2) and then the name of a nearby cross-street (AM9). Adult respondents who refused to provide a complete street address with house number were asked just for the name of the street they lived on (AM8) and the nearest cross street.

Westat conducted geocoding of the address information available from the directory service and the interview. The geocoding software automatically matched the input addresses to a spatial database of roads, which returned the addresses' latitude/longitude, state fips, county fips, tract id, block group id, and block id. If the software was unable to match to the street address, it automatically matched to the geographic zip centroid. In such cases, the latitude/longitude, state fips, county fips, tract id, block group id, and block id of the zip centroid were provided. A flag in the geocoding dataset indicated which match occurred.

The latitude and longitude produced by this geocoding does not give the "roof-top" coordinates of individual households, but rather an interpolation based on the street segment in the road database. For example, if a segment of road called Main St has a range of possible addresses starting at 100 and ending at 200, an input address of 130 Main St will result in a placement 30% along the road segment. Even though the returned latitude and longitude do not represent the exact roof-top location of 130 Main St., in practice the location is often close enough to be considered information that would identify individual households at the full six-digit code for latitude and longitude, particularly in rural areas. Therefore, Westat provided to SLHI coordinates at the five-digit level for Maricopa County and the four-digit level for the remainder of Arizona.

[3] The verification was not done if the telephone number was unlisted or if the sample vendor indicated that the number was on the “do not call” list.

6. WEIGHTING AND ESTIMATION

This chapter introduces the concept of weighting and provides some background on the weights developed for analyzing AHS survey data. It then describes the steps involved in creating the household and person weights, including the raking procedure and the development of control totals. The last two sections describe imputation methods and a review of how to estimate the variance for estimates produced using weighted AHS data.

6.1 Weighting Approach

In an ideal survey, all the units in the inference population are eligible to be selected into the sample and all those in the sample participate in the survey. In practice, neither of these conditions occurs. Some units are not eligible for the sample (undercoverage) and some of the sampled units do not respond (nonresponse). If undercoverage and nonresponse are not addressed, then estimates from the survey will be biased. Weighting is a process that attempts to make the estimates from the survey respondents representative of the total population that was sampled by accounting for the chances of selecting units into the sample and making adjustments for imperfections in the sample.

The philosophy used in AHS weighting is a classical design-based approach with the base weights constructed from the inverse of the probabilities of selection. In the perfect data collection, this scheme produces unbiased estimates and does not require any model assumptions. However, these weights must be modified because of imperfections such as undercoverage (some households in the target population are not covered in the standard RDD sampling frame) and the fact that some sampled units do not respond. If undercoverage and nonresponse are not addressed, then the estimates from the survey will be biased.

The weighting procedure used for AHS accomplishes the following objectives:

- Compensates for differential probabilities of selection for households and persons;
- Reduces biases occurring because nonrespondents may have different characteristics than respondents;
- Adjusts, to the extent possible, for undercoverage in the sampling frames and in the conduct of the survey; and

- Reduces the variance of the estimates by using auxiliary information.

6.2 Household Weighting

6.2.1 Base Weights

Each telephone number in the AHS sample is assigned a base weight, computed as the inverse of the probability of selection of the telephone number. The sample was drawn using a list-assisted approach from a stratified frame of 100 banks[4] with at least one listed telephone number in the state of Arizona. The strata were created allocating telephone exchanges that are geographical1y located in the areas to sample. In this approach, a bank is drawn from the frame and two digits are randomly generated to complete the sampled telephone number. The base weight ($HHBSW_{ki}$) can be expressed as

$$HHBSW_{ki} = \frac{100 * NBANKS_k}{n_k}$$

where NBANKS is the number of 1+ banks in the stratum k and n_k is the number of telephone numbers sampled in stratum n_k. In AHS, we sampled 129,250 telephone numbers from a total of 51,176 working banks.

6.2.2 Subsampling Cases with a Mailable Address

As mentioned in Chapter 1, the efficiency of data collection was improved by stratifying the telephone numbers by mailable status and subsampling the strata at different rates. The stratified samples were then subsampled at rates determined using the principles of optimal allocation to balance both data collection costs and the variances of the estimates.

The substrata were created using the original strata and the information on working status (residential, business, or nonworking telephone number), the listed status (telephone number listed or not), and mail status (the telephone number has a "mailable" address or not). Table 6-1 shows the sample size and observed subsampling rates by sampling stratum for AHS.

[4] A bank is defined as 100 consecutive telephone numbers with the same first eight digits including area code.

Table 6-1. Sample size and number of telephone numbers and subsampling rates by sampling stratum for AHS

Stratum	Substratum	Sample size	Retained sample	Subsampling rate
Maricopa County	With a mailing address	18,201	18,201	100.0%
	Without a mailing address	26,210	18,810	71.8%
	Purged	59,424	59,424	100.0%
	Total	103,835	96,435	
Remainder of Arizona	With a mailing address	5,481	5,481	100.0%
	Without a mailing address	6,750	4,900	72.6%
	Purged	13,184	13,184	100.0%
	Total	25,415	23,565	
Statewide Total		129,250	120,000	

The adjusted household weight, $HHAlW_i$, that accounts for this subsampling is computed as

$$HHAlW_i = HHAlF_c \cdot HHBSW_i ,$$

where $HHAlF_c$ is the mailing status subsampling factor adjustment computed as $HHAlF_c = 1/RATE_c$, where $RATE_c$ is the subsampling rate shown in Table 6-1.

6.2.3 Refusal Subsampling Adjustment

In the third step of weighting, the mailing-status-adjusted weights were adjusted to reflect the differential refusal conversion efforts made during data collection. In order to adjust the weights for screener interview refusal subsampling, telephone numbers were classified into screener refusal groups using their refusal status (i.e., whether the respondent ever refused) and the value of the refusal conversion flag as shown in Table 6-2.

Table 6-2. Screener refusal groups for RDD sample

Screener refusal group	Respondent ever refused screener interview?	Refusal Subsampling Flag	Description
NRef	No	N/A	Households where respondent did not refuse the screener interview (includes complete and noncomplete interviews)
RefC	Yes	Yes	Households where respondent refused the screener interview and refusal conversion procedures were used
RefNC	Yes	No	Households where respondent refused the screener interview and refusal conversion procedures were not used

The refusal subsampling adjusted weight, $HHA1W_i$, is:

$$HHA2W_i = HHA2F_i * HHA1W_i$$

where $HHA2F_i$ is the refusal subsampling adjustment factor computed as:

$$HHA2F_c = \begin{cases} \dfrac{\sum_{i \in RefC, RefNC} \delta_i(c)}{\sum_{i \in RefC} \delta_i(c)} & \text{If } i \in RefC \\ 0 & \text{If } i \in RefNC \\ 1 & \text{If } i \in NRef \end{cases},$$

where the groups *RefC*, *RefNC*, and *NRef* are defined in Table 6-2, $HHBSW_i$ is the base weight, and $\delta_i(c)$ is 1 if the number is in sampling stratum *c* and is zero otherwise. The factor $HHA1F_i$ is the inverse of observed proportion of cases that are flagged for conversion.

6.2.4 Multiple Telephone Adjustment

At the end of the screener interview, information about the existence of additional telephone numbers and their use in the household was collected. If more than one telephone number is used for residential purposes (not solely for business, fax or computer use, etc.), the

household had a greater probability of selection because it could have been selected through any of the additional telephone numbers in the household. In such cases, the household weight is adjusted to reflect the increased probability of selection. The multiple telephone adjusted household weight, $HHA3W_i$, is computed as:

$$HHA3W_i = HHA3F_i * HHA2W_i,$$

where $HHA3F_i$ is the multiple telephone adjustment factor computed as:

$$HHA3F_i = \begin{cases} 0.5 & \text{If household } i \text{ has more than one residential telephone number} \\ 1 & \text{Otherwise} \end{cases}.$$

In this adjustment, we assumed that there was at most one additional residential-use telephone number in the household.

6.3 Person Weights

A person weight was created for each adult, child and adolescent who completed the extended interview.[5] The initial person weight is the product of the household weight and the reciprocal of the probability of selecting the person in the household. The creation of the person weight is different for adult, child, and adolescent interviews.

6.3.1 Adult Weight

As described in Chapter 1, one adult was sampled with equal probability from all adults in the household using the Rizzo method (see Rizzo et. al., 2004). The initial adult weight is the product of the final household weight and the inverse of the probability of selection of the adult. The expression for the adult initial weight, $ADA0W_j$, is

$$ADA0W_j = ADCNT_i \cdot HHA3W_i,$$

where $ADCNT_i$ is the total number of adults in household i, and $HHA3W_i$ is the adjusted household weight.

[5] Adult extended interviews are considered complete provided the adult completed through Section K.

6.3.2 Child Weight

The child weight is the product of the adjusted household weight and the probability of sampling the child within the household. As described in Chapter 1, the selection of the child was done in two steps. In the first step, one adult was randomly selected among all adults in the household. In the second step, one child was randomly selected among all the children associated with the sampled adult (i.e., the sampled adult is the parent or legal guardian of the child). If the sampled adult did not have an associated child, then no child was sampled even if there were children present in the household.

Since the child sampling depended on the relationships among children and adults within the household, these relationships were defined before sampling children. The probability of selection reflects the fact that the sampled child could have been selected through the spouse/partner of the sampled adult if both are the parents or legal guardians[6] of the sampled child. Accordingly, the initial child weight, $CHA0W_j$, is

$$CHA0W_j = \frac{1}{CHPROB_j} HHA3W_i \ ,$$

where $HHA3W_i$ is the household adjusted weight, $CHPROB_j$ is the probability of selecting the j^{th} child associated with the i^{th} sampled adult. If the sampled adult does not have a spouse/partner living in the household or if the spouse/partner of the sampled adult is not the parent or legal guardian of the sample child, then

$$CHPROB_j = \frac{1}{ADLTCNT} \cdot \frac{SACHMOS_j}{\sum_j SACHMOS_j} \ ,$$

where $ADLTCNT$ is the number of adults in the household and $SACHMOS_j$ is the measure of size of child *j*. Within the same household children age 0 to 5 years have a measure of size twice that of children age 6 to 11 years. If the sample adult has a spouse/partner living in the household and the spouse/partner of the sample adult is the parent or legal guardian of the sample child then

[6] If the spouse/partner of the sampled adult is living in the household.

$$CHPROB_j = \frac{1}{ADLTCNT}\left(\frac{SACHMOS_j}{\sum_j SACHMOS_j} + \frac{SACHMOS_j}{\sum_k SPCHMOS_k}\right),$$

where $ADLTCNT$ and $SACHMOS_j$ are defined as before and $SPCHMOS_k$ the measure of size child *k* associated with the spouse/partner of the sample adult.

6.3.3 Adolescent Weights

In AHS, adolescents were sampled and responded to the interview for themselves after parental permission to conduct the interview. As with the child weighting, the initial weights for the adolescents incorporated the probability of sampling the adult and the probability of sampling an adolescent among the adolescents associated with the sampled adult. The initial weight, $TNA0W_i$, was

$$TNA0W_i = \frac{1}{TNPROB_j} \cdot HHA3W_i$$

where $HHA3W_i$ was the household weight, and $TNPROB_i$ was computed the same as $CHPROB_i$. However, the measure of size was unity for all adolescents regardless of their age.

6.3.4 Trimmed Weight

Before raking the person weights, we examined the distribution of the weights to determine if there were very large weights that could have a large effect on either the estimates or the variances of the estimates. When observations with large weights were found, the weights for these cases were reduced in a process called trimming.

We computed three statistics to identify influential weights that were candidates for trimming (Liu et al., 2004). The first statistic is a function of spacing of the weights, and the second evaluated the distance between a weight and the next largest weight relative to the size of the weight. The third statistic evaluated the difference of the relative distance from the median weight of the weight to be trimmed and the next smaller weight. These three statistics for the largest 20 weights in each stratum were examined separately. When all three statistics were greater than 1

then the case was a primary candidate for trimming. The final decision on trimming involved the inspection of the weight distribution within sampling stratum.

The trimmed person weight $TRMW_i$ is

$$TRMW_i = TFACT_i * PSW_i ,$$

where PSW_i is the person weight (i.e., $AD0W_i$, $CH0W_i$, or $TN0W_i$) and $TFACT_i$ is the trimming factor for the sampled adult *i* given by

$$TFACT_i = \begin{cases} 1 & \textit{if the weight i is not trimmed} \\ t_i & \textit{otherwise} \end{cases}$$

where $0 < t_i < 1$.

Two adult and three child weights were trimmed[7]. The trimming factor, t_i, was determined as the ratio of the largest weight not to be trimmed to the weight being trimmed rounded up to the nearest hundredth. This factor ranged from 0.54 to 0.90.

6.3.5 Raked Weight

In the last step, the initial person weights were raked to known control totals. This step adjusted the weights for nonresponse and undercoverage. In a landline RDD survey, the undercoverage includes persons in households without a landline telephone.

The primary objective of raking is to reduce bias, but it also helps to reduce response errors, nonresponse bias, and sampling errors Raking may be thought of as a multidimensional poststratification procedure because the weights are basically poststratified to one set (a dimension) of control totals, then these adjusted weights are poststratified to another dimension. After all dimensions are adjusted, the process is iterated until the control totals for all the dimensions are simultaneously satisfied within a specified tolerance.

---

[7] The trimming was done prior to the raking adjustment; however, it was an iterative process. After the trimming and raking, the distribution of the weights was re-examined, and new decisions were made about trimming. This might have changed the decision about which weights should be trimmed or the magnitude of the trimming factor. If the decision was made, the trimmed and raked weights were discarded and new trimming and raking were undertaken. The number of trimmed weights reported here is at the completion of the overall process.

The raking-adjusted estimator is design-unbiased in large samples and is very efficient in reducing the variance of the estimates. Brackstone and Rao (1979), Deville and Särndal (1992), and Kalton and Flores Cervantes (2003) are some references for this procedure.

The raked weight, $RAKEDW_i$, can be expressed as

$$RAKEDW_i = TRMW_i \cdot \prod_{k=1}^{K} RAKEDF_{k_l} ,$$

where $RAKEDF_{k_l}$ is the raking factor for dimension k, level l which adult i is in. For example, if the 4th dimension (k =4) is sex with two levels (l=1 for male and l=2 for female), then the raking factor for this dimension is $RAKEDF_{4_1}$ for the adult male. The raking factors are derived so the following relationship holds for every raking dimension k, and level l,

$$CNT_{k_l} = \sum_i \delta(k_l)_i \cdot RAKEDW_i ,$$

where CNT_{k_l} is the control total, and $\delta(k_l)_i = 1$ if the adult i is in level l of dimension k and zero otherwise.

6.4 Raking and Control Totals

6.4.1 Raking Dimensions

The six raking dimensions used in AHS are shown in Table 6-3. The dimensions were created by combining demographic variables (age, sex, race, and ethnicity) and geographic areas (Maricopa County and the remainder of the state). The dimensions include variables such as household tenure (rent, own) and education that are correlated with characteristics of households without a landline telephone.

Before raking, dimensions with levels or cells with fewer than 50 respondents were collapsed with adjacent cells. When collapsing the cells, we ensured that none crossed stratum boundaries.

Table 6-3. Definitions of the dimensions used in raking

Dimension	Level	Description	Categories	
1	Stratum (collapsed where necessary)	Sex (2) x Age groups (4)	11	Under 18 years, male
			12	18 to 34 years, male
			13	35 to 64 years, male
			14	65 or older, male
			21	Under 18 years, female
			22	18 to 34 years, female
			23	35 to 64 years, female
			24	65 or older, female
2	Stratum	Age groups (8) (collapsed where necessary)	1	Under 11 years
			2	12 to 17 years
			3	18 to 24 years
			4	25 to 29 years
			5	30 to 39 years
			6	40 to 49 years
			7	50 to 64 years
			8	65 years or older
3	Stratum	Race-ethnicity (3) x Age groups (2)	11	Hispanic, under 18 years
			21	Non-Hispanic White, under 18 years
			31	Other, under 18 years
			21	Hispanic, 18 years or older
			22	Non-Hispanic White, 18 years or older
			23	Other, under 18 years or older
4	Stratum	Household tenure (2) x Age Race-ethnicity (3) (collapsed where necessary)	111	Hispanic, Under 18 years, Male
			121	Hispanic, 18 to 34 years, Male,
			131	Hispanic, 31 to 64 years, Male
			112	Hispanic, Under 18 years, Female
			122	Hispanic, 18 to 34 years, Female,
			132	Hispanic, 31 to 64 years, Female
			211	Non-Hispanic White, Under 18 years, Male
			221	Non-Hispanic White, 18 to 34 years, Male,
			231	Non-Hispanic White, 31 to 64 years, Male
			212	Non-Hispanic White, Under 18 years, Female
			222	Non-Hispanic White, 18 to 34 years, Female,
			232	Non-Hispanic White, 31 to 64 years, Female
			311	Other, Under 18 years, Male
			321	Other, 18 to 34 years, Male,
			331	Other, 31 to 64 years, Male
			312	Other, Under 18 years, Female
			322	Other, 18 to 34 years, Female,
			332	Other, 31 to 64 years, Female

Table 6-3. Definitions of the dimensions used in raking (Continued)

Dimension	Level	Description	Categories
5	Stratum	Household tenure (2) x Age Race-ethnicity (3) (collapsed where necessary)	111 Own, Hispanic Under 18 years 112 Own, Hispanic , 18 to 34 years 113 Own, Hispanic, 31 to 64 years 114 Own, Hispanic, 65 years or older 121 Own, Non-Hispanic White Under 18 years 122 Own, Non-Hispanic White, 18 to 34 years 123 Own, Non-Hispanic White, 31 to 64 years 124 Own, Non-Hispanic White, 65 years or older 121 Own, Other, Under 18 years 122 Own, Other, 18 to 34 years 123 Own, Other, 31 to 64 years 124 Own, Other, 65 years or older 111 Rent, Hispanic Under 18 years 112 Rent, Hispanic , 18 to 34 years 113 Rent, Hispanic, 31 to 64 years 114 Rent, Hispanic, 65 years or older 121 Rent, Non-Hispanic White Under 18 years 122 Rent, Non-Hispanic White, 18 to 34 years 123 Rent, Non-Hispanic White, 31 to 64 years 124 Rent, Non-Hispanic White, 65 years or older 121 Rent, Other, Under 18 years 122 Rent, Other, 18 to 34 years 123 Rent, Other, 31 to 64 years 124 Rent, Other, 65 years or older
6	Stratum	Education attainment (4) X age (collapsed where necessary)	1 Not applicable (under 18 years) 2 Less than High School, 18 years or older 3 High School grad or more, 18 years or older 4 At least some college, 18 years or older

6.4.2 Raking Factors

Table 6-4 shows the overall adjustment factors for the adult, child, and adolescent weights. The overall adjustment factors were computed as the ratio of the control total to the sum of weights before raking. The factors were an approximation of the bias correction of estimates of totals due to both nonresponse and undercoverage. Since the weights were adjusted for nonresponse and undercoverage at the same time, the raking factor could not be used as an indirect measure of these because the nonresponse and undercoverage are confounded. Nevertheless, they may be used as an indicator of which groups were harder to reach, or did not complete the interview.

Table 6-4. Overall adjustment raking factors for adult, child, and adolescent interviews by sample characteristics

Characteristic	Adult	Child	Adolescent
Total	9.68	7.30	12.46
Self-reported stratum			
Maricopa	10.11	7.30	12.46
Remainder of Arizona	9.12		
Sex			
Male	12.16	12.16	12.66
Female	8.08	8.08	12.23
Age			
6 – 11 years		7.26	
12 – 17 years			12.46
18-24 years	18.08		
25-29 years	24.12		
30-39 years	13.71		
40-49 years	11.11		
50-64 years	7.21		
65 years and over	5.98		
Race/Ethnicity			
Latino	13.66	13.66	13.66
Non-Latino			
White alone	8.47	8.47	8.47
Other	12.42	12.42	12.42
Educational Attainment			
N/A (age < 18 years)		7.30	12.46
Less than High School	14.96		
HS grad or GED	10.70		
Some college or more	8.40		
Household Tenure [a]			
Owner	8.65	6.81	11.83
Renter	13.66	8.32	14.47
Number of adults in HH			
One	7.76	7.56	12.35
Two	9.23	7.15	12.50
Three or more	12.38	7.56	12.45

Table 6-4 shows that adolescents have the largest adjustment, which is not surprising given the two levels of nonresponse (parent and adolescent). Children have the lowest adjustment due in part to the child first procedures. As expected in RDD surveys, Maricopa, an urban area, has a larger adjustment than the remainder of the state. In addition, as expected, females have a lower adjustment than males. Young adults (25 to 29 years old) have a large adjustment factor because they are less likely than other age groups to respond to surveys.

6.4.3 Creation of Control Totals

In AHS, the control totals are mainly derived from the 2007 Arizona Department of Commerce Population Estimates, 2008 Arizona Department of Commerce Projections (State of Arizona, Department of Commerce, 2006, 2006b), and the 2006 American Community Survey estimates for Arizona (U.S. Census Bureau, 2003).

Population control totals of the number of persons by age group and sex at the stratum level for AHS were created from the Arizona 2008 Population Projections. Since these were projections made several years ago, we used the Arizona 2007 Population Estimates to adjust the population totals. We were not able to use the 2007 estimates directly because these were not available by age group and gender.

Some limitations of using Arizona Department of Commerce data were the inclusion of about 1.95 percent of the population of Arizona who resided in "group quarters" (i.e., persons living with nine or more unrelated persons), and the lack of population estimates for any variable other than sex and age.

To overcome these limitations we developed a procedure to remove the population of group quarters and produce population totals by variables such as race/ethnicity, education and household tenure. This procedure used information from the 2006 American Community Survey (ACS) estimates for Maricopa County and Arizona to estimate the percentage of persons living in group quarters by age group, race/ethnicity, and household tenure within stratum. We also computed the proportion of the population by sampling stratum by race/ethnicity × age group × gender, household tenure × race/ethnicity × age group, education attainment × age group, and race/ethnicity × age group. Using these distributions, we adjusted the population total using raking so that the sum of the population matched the Arizona Department of Commerce total by age and gender and the proportions of the population from the ACS. The result was a micro-level file with

population totals by sampling stratum, group quarters, age group, sex, education, household tenure and race/ethnicity. This file was summarized to produce the control totals use in raking, excluding the population in group quarters and the child and adolescent totals outside of Maricopa County.

6.5 Item Imputation

As in most surveys, responses to some data items in AHS were not obtained for all interviews. For the items that were needed for weighting, such as race, ethnicity, age, sex, education, and tenure (own/rent), the items were imputed.

Table 6-5 lists the variables needed for weighting that were imputed. The level of missing data is relatively small. When the amount of missing data is small, and assuming that the data are missing at random (i.e., the missing data have the same distribution as those with complete data within groups defined for imputation), then the bias of the estimates due to the missing data should be relatively small. The imputations may also increase the variance of the estimates, but this effect should be negligible.

Table 6-5. Description of imputed variables

Variable name	Description	Extended interview items	Variable type
SRAGE	Self-reported age	AA2, CA3, TA2, KAA2	Demographic
SRSEX	Self-reported sex	AA3, CA1, TA3, KAA3	Demographic
SRTENR	Self-reported household tenure	AK25, KAK25	Socio-economic
SREDUC	Self-reported educational attainment	AH47, KAK47	Socio-economic
SRH	Self-reported Latino	AA4, CH1, TI1	Ethnicity
SRW	Self-reported white	AA5A_6, CH3_6, TI2_6	Race
SRAA	Self-reported African American	AA5A_5, CH3_5, TI2_5	Race
SRAS	Self-reported Asian	AA5A_4, CH3_4, TI2_4	Race
SRAI	Self-reported American Indian/ Alaska Native	AA5A_3, CH3_3, TI2_3	Race
SRPI	Self-reported Native Hawaiian and Other Pacific Islander	AA5A_1, AA5A_2, CH3_1, CH3_2, TI2_1, TI2_2	Race
SRO	Self-reported Other race	AA5A_7, CH3_7, TI2_7	Race

In AHS, random allocation and hot-deck imputation were used to fill in the missing responses. Random selection from the observed distribution was used only to impute age (there were no missing values for sex), which had 5 missing values after considering information from the screener.

The second technique, hot-deck imputation, was used to impute race, ethnicity, household tenure, and educational attainment. The hot-deck approach is probably the most commonly used to assign values for missing responses in large-scale household surveys (Sande, 1983; Ford, 1983). With a hot deck, a value reported by a respondent for a particular item is assigned or donated to a "similar" person who did not respond to that item. In hot-deck imputation, the respondents to an item form a pool of donors while the nonrespondents are a group of recipients. A recipient is matched to the subset pool of donors with the same characteristics. The recipient is then assigned a randomly imputed value from one of the donors in the pool. Once a donor is used, it is removed from the donor pool.

6.5.1 Self-Reported Age

Among the 13 records with missing age (0.1 percent), 12 were adults and one was a child. The imputed value of age for the child was randomly selected from the observed age distribution of children in the sample. A hierarchical process was followed to impute the 12 missing self-reported age values for adults. The process used the values for self-reported age (question AA2 on the adult interview), the self-reported adult age range (question AA2A on the adult interview) asked when the adult refused to provide a specific age, the proxy-reported adult age collected during the child-first interview (question KAA2) if available, and the adult age collected during the screener interview (question ADULTAGE on the screener interview). The procedure used any available information and checked that the assigned age was consistent with the age range if provided. If there were inconsistencies or an age could not be assigned, age was imputed using the observed age distribution of the adult sample.

6.5.2 Household Tenure and Educational Attainment

Household tenure and the adult respondent's educational attainment were used to create raking dimensions 5 and 6. As shown in Table 6-6, household tenure had 225 missing responses (5.1 percent) and educational attainment had 19 missing (0.45 percent).

Table 6-6. Counts and percentages of imputed self-reported education attainment and household tenure

	Adult interviews	
	Count	Percentage
Self-reported Education Attainment		
Under 18 years of age	NA	
Less than HS, 18 years of age or older	3	0.07%
High School (or equivalent), 18 years of age or older	3	0.14%
Some college, 18 years of age or older	3	0.12%
BS and above, 18 years of age or older	5	0.12%
Total	19	0.45%
Self-reported Household Tenure		
Owner	182	4.15%
Renter	43	0.98%
Total	225	5.13%

Hot-deck imputation was used to impute missing values for these two variables. The search algorithm CHAID (Kass, 1980) was used to create hot-deck cells using the variables available for both donors and recipients that were found to be good predictors. A donor was then randomly drawn from the cell and its value for the variable being imputed was assigned to the recipient. The variables considered in CHAID to create the hot-deck cells were age race, gender, ethnicity, number of adults in the household, presence of children or adolescents, poverty, and the percentages of college graduates, home owners, African Americans and Hispanics in the telephone exchange.

6.5.3 Imputation of Single Self-Reported Race and Ethnicity

A hot-deck imputation was also developed to deal with race and ethnicity within households. Table 6-7 shows the number and percentage of cases with imputed values by type of extended interview (adult, child, and adolescent).

Table 6-7. Number and percentage of imputed interviews with missing self-reported race and/or ethnicity

	Imputed race*		Imputed ethnicity	
Type of interview	Count	%	Count	%
Adult	101	2.4%	11	0.26%
Child	50	7.8%	2	0.31%
Adolescent	19	9.3%	3	1.47%
Total	170	3.4%	16	0.32%

* At least one value of race was imputed.

The hot-deck imputations were done separately by the completed extended interview structure of the household. In general, the imputation procedure was done at the household level and handled households with the fewest missing values first and then moved to the cases with more missing values. The simplest household structure is where only an adult was interviewed (versus a household with an adult and an adolescent and/or a child). A household with only one adult with missing ethnicity was imputed before a household with only an adult that had both missing race and ethnicity.

The patterns of missing data for race and ethnicity varied by the structure of the household. For the simple case where only an adult was interviewed, the donors were selected from other adult-only households. If the adult was missing both race and ethnicity, both values were imputed from the same donor. If the adult had a reported race but was missing ethnicity, then a donor with the same race (all six race values were placed into a vector and only adults with exactly the same values could be donors) was randomly selected. For an adult with reported ethnicity and missing race, the same procedure was used; only adults in adult-only households with the same value of ethnicity could be donors. Whenever possible, the donors and the recipients were from the same sampling stratum. For cases where the pool formed in this way had too few donors, sampling strata were combined. Once a donor was used, it was removed from the pool for all future hot deck runs.

The same principles were used for more complex household structures. In these cases, some households had missing race and ethnicity for all sampled persons, while in others one or more of the sampled persons might have a reported race and ethnicity. Various combinations, such as a reported ethnicity but not race, were also encountered. Separate hot deck runs were made to accommodate all of these situations. As an illustration, consider households where an adult and a child are interviewed. Assume the adult reported non-Latino ethnicity and Asian race and the child only reported non-Latino ethnicity but no race. The pool of donors for imputing the child's race consists of households where only an adult and a child were interviewed and where the adult reported non-Latino ethnicity and Asian race and the child reported non-Latino ethnicity. The households with other combinations of persons with missing race and/or ethnicity were imputed in an similar way. Table 6-8 shows the counts and percentages of imputed values by self-reported race and ethnicity and type of extended interview (adult, child, and adolescent).

Table 6-8. Counts and percentages of imputed interviews with missing self-reported race and ethnicity by type of extended interview

	Total		Extended interview type					
			Adult		Child		Adolescent	
	Count	%	Count	%	Count	%	Count	%
Self-reported race								
White alone	83	1.65	55	1.31	23	3.58	5	2.45
African American alone	2	0.04	1	0.02	1	0.16	0	0.00
Asian alone	1	0.02	1	0.02	0	0.00	0	0.00
American Indian/ Alaska Native alone	7	0.14	4	0.10	1	0.16	2	0.98
Pacific Islander alone	3	0.06	1	0.02	0	0.00	2	0.98
Other race alone	70	1.39	37	0.88	23	3.58	10	4.90
Two or more races	4	0.08	2	0.05	2	0.31	0	0.00
Total	170	3.37	101	2.41	50	7.78	19	9.31
Self Reported Ethnicity								
Hispanic	1	0.02	0	0.00	0	0.00	1	0.49
Non-Hispanic	6	0.12	2	0.05	2	0.31	2	0.98
Total	7	0.14	2	0.05	2	0.31	3	1.47
Completed interviews	5043	100	4196	100	643	100	204	100

6.5.4 Self-Reported County and Self-Reported Stratum

In AHS, geographic location variables such as self-reported stratum (i.e., inside/outside Maricopa County) were assigned after geocoding (see Chapter 5) Table 6-9 shows the variables used in the geocoding process.

Table 6-9. Variables used in geocoding

Variable	Description	Source
AH42	County of residence (self report)	Adult questionnaire
AO1ADDR	Confirmed/corrected street address	Adult questionnaire
AO1CITY	Confirmed/corrected city	Adult questionnaire
AO1ZIP	Confirmed/corrected ZIP Code	Adult questionnaire
AM7	ZIP Code (self report)	Adult questionnaire
AO2ADDR	Street address (self report)	Adult questionnaire
AO2CITY	City (self report)	Adult questionnaire
AM8	Street name of residence (self report)	Adult questionnaire
AM9	Street name of nearest cross street (self report)	Adult questionnaire
M_ADDR	Street address (matched to phone number prior to interview)	Directory service
M_CITY	City (matched to phone number prior to interview)	Directory service
M_ZIP	ZIP Code (matched to phone number prior to interview)	Directory service
S_ZIP	ZIP Code (provided by sample vendor for every phone)	Sample vendor

The derived location variable SRSTRATA (self-reported stratum), was a household-level variable assigned to all adult, child and adolescent records within the same household before creating the raking dimensions.

6.6 Variance Estimation

This section addresses computing sampling errors for AHS data. The first section describes the methodology for producing estimates of sampling variability. The second section provides a general review of the two main methods of computing sampling errors or variances of estimates from surveys with complex sample designs like AHS. The following section describes a replication method of variance estimation that can be used with the data; and the last section describes the commercially available software that can be used to compute sampling errors for AHS estimates.

6.6.1 Methods for Variance Estimation

Variance estimation procedures have been developed to account for complex sample designs. Using these procedures, factors such stratification, multistage sampling, and the use of differential sampling rates to oversample a targeted subpopulation can be appropriately reflected in estimates of sampling error. The two main methods are replication and linearization or the Taylor series approximation. Wolter (1985) is a useful reference on the theory and applications of these methods. Shao (1996) is a more recent review paper that compares these methods. The rest of this section briefly reviews these methods.

The basic idea behind replication is to draw subsamples from the sample, compute the estimate from each of the subsamples, and estimate the variance of the original sample using the variability of the subsample estimates. Specifically, subsamples of the original "full" sample are selected to calculate subsample estimates of a parameter for which a "full-sample" estimate of interest has been generated. The variability of these subsample estimates about the estimate for the full sample can then be assessed. The subsamples are called replicates, and the estimates from the subsamples are called replicate estimates.

Replicate weights are created to produce the corresponding replicate estimate. Each replicate weight is computed using the same estimation steps as the full sample weight but using

only the subsample of cases comprising each replicate. The variance estimate takes the following form:

$$v\left(\hat{\theta}\right) = c \sum_{k=1}^{G} \left(\hat{\theta}_{(k)} - \hat{\theta}\right)^2 \quad (1)$$

where

θ is an arbitrary parameter of interest;
$\hat{\theta}$ is the estimate of θ based on the full sample;
$\hat{\theta}_{(k)}$ is the k^{th} estimate of θ based on the observations included in the k^{th} replicate;
G is the total number of replicates formed;
c is a constant that depends on the replication method; and
$v\left(\hat{\theta}\right)$ is the estimated variance of θ.

The other widely used method for variance estimation for complex sample surveys is called linearization and is based on the Taylor series approximation. In this method, the Taylor series linearization of a statistic is formed and then substituted into the formula for calculating the variance of a linear estimate appropriate for the sample design. Linearization relies on the simplicity associated with estimating the variance for a linear statistic even with a complex sample design.

6.6.2 Design of Replicates

Two major reasons for using replication to estimate variances for AHS are operational convenience and the ability to reflect all components of the design and estimation in the estimates of variability. With respect to operational convenience, once replicate weights are constructed, it is very simple to compute estimates of sampling errors. No special care is needed for subgroups of interest, and no knowledge of the sample design is required. If an estimator is needed that was not previously considered, replication methods can be easily used to develop an appropriate estimate of variance. In such a case, variance estimates using a Taylor series approach would require additional work. The variance estimation stratum and unit must also be included in the file for the Taylor series method.

The second reason for using replication is probably more important. The raking adjustment made in developing the AHS analysis weights affects the sampling errors of the

estimates produced from the survey. The replicate weights prepared for AHS reflect this aspect of weighting. Currently existing software for using the Taylor series method for variance estimation cannot reflect these weighting adjustments. In some Taylor series software poststratification can be taken into account, but only in specific situations.

In AHS, a paired unit jackknife method (JK2[8]), a form of jackknife replication, was selected for computing variances. In the JK2 replication method, adjacent pairs of sampled telephone numbers are treated as having been sampled from the same stratum. Each pair of sampled telephone numbers is treated as an implicit stratum, where each such stratum is defined by the sort order used in the sample selection. In this method, the constant, *c*, in equation (1) equals 1. This approach has been used in other RDD studies such as all cycles of the California Health Interview Survey.

The first step in designing the replicate structure is to determine the number of variance estimation strata. In the JK2 method, the number of replicates is equal to the number of variance estimation strata. The choice of the number of variance estimation strata is based on the desire to obtain an adequate number of degrees of freedom to ensure stable estimates of variance while not having so many as to make the cost of computing variance estimates unnecessarily high. For AHS, we elected to create 80 variance estimation strata, even though many more could have been created. Once the variance strata are created, the replicate weights can be created. The full replicate weights are constructed by first modifying the full sample base weights. The same sequence of weighting adjustments used in the full sample weight is then applied to the replicate base weights to create the final replicate weights. Thus, all of the different components of the weighting process are fully reflected in the replicate weights, ranging from household adjustments (nonresponse, adjustment for household noncoverage, and adjustment to control totals) to person adjustments (nonresponse and raking).

6.6.3 Software for Computing Variances

In the past, most standard statistical software packages assumed a simple random sample when computing estimates of variance. As a result, estimates of variance from these packages had the potential to seriously understate the true variability of the survey estimates. However, in recent years, specialized commercial software has been developed to analyze data

[8] This method is denoted as JK2 in the software program, WesVar, which was used to compute all the sampling errors in this report.

from complex surveys (Lepkowski and Bowles, 1996). In this section, we describe the elements needed to compute estimates for AHS using some of these programs.

WesVar Version 4.2 (Westat, 2000) is a software package developed and distributed by Westa**t.** WesVar uses replication methods to compute variance estimates. WesVar is an interactive program with a graphical interface that makes it simple to specify the estimates for sampling errors for estimates of interest.

SUDAAN® (Research Triangle Institute, 2005) is a package developed by Research Triangle Institute to analyze data from complex sample surveys. SUDAAN is available as a standalone package or it can be called using SAS. SUDAAN and WesVar produce the same point estimates. The difference between the two packages is in the method used to compute the variances. While WesVar uses replication exclusively, SUDAAN can use either a first-order Taylor series expansion approximation (linearization) or replication to compute variances of the estimates.

SAS has also introduced procedures to analyze survey data. SAS® Version 9 (SAS Institute, 2005) has two procedures for analyzing survey data: PROC SURVEYMEANS and PROC SURVEYREG. Both use the linearization approach to estimate standard errors and the new release will use replication. These procedures are relatively new in SAS and do not contain as many features as most of the other packages. At the current time, the SAS procedures are the most limited of all the packages discussed here.

Another software package that can be used to analyze survey data is STATA (version 9 is the latest version as of this writing) (STATA Corporation, 2005). STATA is a command driven, fully programmable statistical package used for managing, analyzing, and graphing data. STATA was developed by StataCorp and is available for a variety of platforms, including DOS, Windows, Macintosh, and UNIX. STATA's statistical, graphical, and data management capabilities are fully expandable through programming. STATA has a family of *svy*- commands to analyze data from sample surveys. The set of analytic methods in STATA is more exhaustive than any other package. Like SUDAAN, STATA can use linearization (**linear** variance type option) or replication (**jack** variance type option) to estimate variances. \

When using linearization theory to estimate variances the software packages referred to above require auxiliary variables that provide information about the sample design. Two variables have been defined and included in the data files (TSVARSTR and TSVRUNIT). TSVARSTR is required for all analyses, but TSVRUNIT is required only when analyses are

performed using the combined data. In other words, when separate analyses are done by adults, children or teens the variable TSVRUNIT is not required. The definitions of TSVARSTR and TSVRUNIT are

- TSVARSTR (Taylor's series variance stratum). The variable TSVARSTR indicates the variance stratum to be used for software that computes estimates of variance using the Taylor series method. The variable TSVARSTR was created by sequentially numbering the sampling strata.

- TSVRUNIT (Taylor's series unit). The variable TSVRUNIT indicates the primary sampling unit (PSU). In this case the PSU is the sampled household. TSVRUNIT was created by sequentially numbering the PSU's within the sampling strata.

The same variables, TSVARSTR and TSVRUNIT, can be used for linearization variance estimation in SUDAAN, SAS, and STATA.

REFERENCES

Anderson, J.E., Nelson, D.E., and Wilson, R.W. (1998). Telephone coverage and measurement of health risk indicators: Data from the national health interview survey. *American Journal of Public Health*, 88, 1392-1395.

Blumberg, S. J., Luke, J. V., Cynamon, M. L. (2006). Telephone coverage and health survey estimates: Evaluating the need for concern about wireless substitution. *American Journal of Public Health*, 96(5), 926-931.

Brackstone, G.J., and Rao, J.N.K. (1979). An investigation of raking ratio estimation. *Sankhya,* 41, 97-114.

Brick, J. M., Montaquila, J., Hagedorn, M. C., Roth, S. B., and Chapman, C. (2005). Implications for RDD design from an incentive experiment, *Journal of Official Statistics*, 21(4), 571-589.

Brick, J.M., and Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, 5, 215-238.

Brick, J.M., and Waksberg, J. (1991). Avoiding sequential sampling with RDD. *Survey Methodology*, 17(1), 27-41.

Brick, J.M., Judkins, D., Montaquila, J., and Morganstein, D. (2002). Two-phase list-assisted RDD sampling. *Journal of Official Statistics*, 18(2), 203-215.

Brick, J.M., Waksberg, J., Kulp, D., and Starer, A. (1995). Bias in list-assisted telephone surveys. *Public Opinion Quarterly*, 59(2), 218-235.

Deville, J.C., and Särndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association,* 87, 376-382.

Ford, B. M. (1983). *An overview of hot-deck procedures*: *Vol. 2. Incomplete data in sample surveys*. New York: Academic Press.

Ford, E.S. (1998). Characteristics of survey participants with and without a telephone: Findings from the third national health and nutrition examination survey. *Journal of Clinical Epidemiology*, 51, 55-60.

Giesbrecht, L.H., Kulp, D.W., and Starer, A.W. (1996). Estimating coverage bias in RDD samples with current population survey data. *Proceedings of the Survey Research Methods Section,* 503-508. American Statistical Association.

Groves, R.M. (1989). *Survey errors and survey costs*. New York: John Wiley and Sons.

Hansen, M.H., and Hurwitz, W.N. (1946). The problem of nonresponse in sample surveys. *Journal of the American Statistical Association*, 41, 517–529.

Hartely, H.O. (1974). Multiple frame methodology and selected applications. *Sankhya, the Indian Journal of Statistics,* Series C, 36, 99-118.

Holt, D., and Smith, T.M.F. (1979). Post Stratification. *Journal of the Royal Statistical Society*. Series A (General), 142(1), 33-46.

Kalton, G., and Anderson, D.W. (1986). Sampling rare populations. *Journal of the Royal Statistical Society*, A, 149, 65-82.

Kalton, G., and Flores Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19(2), 81-97.

Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied statistics*, 29, 119-127.

Kish, L. (1949). A procedure for objective respondent selection within the household. *Journal of the American Statistical Association,* 44, 380-87.

Kish, L. (1965). *Survey sampling*. New York: John Wiley and Sons.

Kish, L. (1992). Weighting for unequal P_i. *Journal of Official Statistics*, 8, 183-200.

Lee, S., Kurata, J., Nguyen, H.A., and Jawad, M. (2006). *Shifting Focus from Nonresponse to Coverage Error: Linguistic Minorities in Surveys.* Presented at the annual meeting of the American Association for Public Opinion Research, Montréal, Canada.

Lepkowski, J., and Bowles, J. (1996). Sampling error software for personal computers. *The Survey Statistician*, 35, 10-16.

Liu, B., Ferraro, D., Wilson, E., and Brick, J.M. (2004). Trimming extreme weights in household surveys. *Proceedings of the Survey Research Methods Section of the American Statistical Association* [CD-ROM]. Alexandria, VA: American Statistical Association.

Office of Management and Budget (1997). *Revisions to the standards for the classification of federal data on race and ethnicity*, 62(210).

Oh, H.L., and Scheuren, F.J. (1983). *Weighting adjustments for unit nonresponse*. Incomplete data in sample surveys: *Vol. 2.* W.G. Madow, I. Olkin, and D.B. Rubin (Eds.). Academic Press,
143-184.

Research Triangle Institute (2005). *SUDAAN user's manual*, *Release 9.0*. Research Triangle Park, NC: Research Triangle Institute.

Rizzo, L, Brick, J.M., and Park, I. (2004). A minimally intrusive method for sampling persons in Random Digit Dialing Surveys. *Public Opinion Quarterly,* 68, 267-274.

Rust, K.F., and Rao, J.N.K. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5, 282-310.

Sande, I. G. (1983). *Hot-deck imputation procedures*, *incomplete data in sample surveys*: *Vol. 3*. New York: Academic Press.

SAS Institute (2005). *SAS OnlineDoc® version 8,* Cary, NC: SAS Institute, Inc.

Shao, J. (1996). Resampling methods in sample surveys (with discussion). *Statistics*, 27, 203-254.

Shapiro, G., Battaglia, M., Camburn, D., Massey, J., and Tompkins, L. (1995). Calling local telephone company business offices to determine the residential status of a wide class of unresolved telephone numbers in a random-digit-dialing sample. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 975-980.

Skinner, C.J., Holt D., and Smith, T. M. F. (1989). *Analysis of complex surveys*. New York: John Wiley and Sons.

Stata Corporation (2005). *STATA survey data reference manual*, *Release 9*. College Station, TX: Stata Corporation.

Sudman, S., Sirken, M.G., and Cowan, C.D. (1988). Sampling rare and elusive populations. *Science*, 240, 991-9.

Tucker, C., Brick, J.M., Meekins, B., and Morganstein, D. (2004*).* Household telephone service and usage patterns in the U.S. in 2004. *Proceedings of the Survey Methods Section of the American Statistical Association* [CD-ROM].

Tucker, C., Lepkowski, J., and Piekarski, L. (2002). The current efficiency of list-assisted telephone sampling designs. *Public Opinion Quarterly,* 66, 321-338.

U.S. Census Bureau (2003). *2002 American Community Survey PUMS file* (Prepared by the U.S. Census Bureau). California.

Verma, V., Scott, C., and O'Muircheartaigh, C. (1980). Sample designs and sampling errors for the world fertility survey. *Journal of the Royal Statistical Society*, A 143, 431-473.

Westat (2000). *WesVar ™ 4.0 user's guide.* Westat, Rockville, MD.

Wolter, K. (1985). *Introduction to variance estimation*. Springer-Verlag.