DESIGN AND METHODOLOGY OF THE

2010 ARIZONA HEALTH SURVEY

SEPTEMBER 2010

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 Ismael Flores-Cervantes, Sherman Edwards, Jonathan Wivagg, and Susan Fraser of Westat.

Suggested citation:

Arizona Health Survey. *Design and Methodology of the Arizona Health Survey*. Phoenix AZ: St. Luke's Health Initiatives, 2010.

Table of Contents

Tables

Tables (continued) Page

Figures

Executive Summary ES

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 2010. 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 six geographic strata defined by counties or groups of counties. Within each household, one adult (age 18 and over) respondent was randomly selected. In those households with children (under age 6), one child was randomly selected, and the adult most knowledgeable about the child's health completed the child interview. For operational efficiency, the RDD sample was supplemented by listed telephone numbers for households expected to have eligible children. No adults were sampled in these households; if the household did not have an eligible child, it was considered ineligible. The samples were selected so as to complete about 8,100 interviews with adults and 2,000 interviews about children, with separate

targets for each stratum. 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

Stratum Number, Counties Included	Adult	Child
Total	4,196	635
1. Mohave, Coconino, Navajo, Apache, Yavapai		
2. Yuma, La Paz		
3. Graham, Greenlee, Cochise, Santa Cruz		
4. Pinal, Gila		
5. Pima	3,130	635
6. Maricopa	1,066	N/A

Table ES-2 at the end of this summary shows the major topic areas for each of the survey instruments (adult and child). Chapter 2 describes the structure and content of the survey instruments. The average adult interview took about minutes to complete, and the average child interview in a household with an adult interview took about XX minutes. In households where an adult interview had not been completed at the time the child interview was completed, selected questions from the adult interview were included in the child interview. The average length of these child interviews was about XX minutes.

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."

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).

The overall AHS screener response rate was 46 percent for the adult sample and 39 percent for the child sample. The adult extended response rate was 42 percent, resulting in an overall adult response rate of 19 percent. The rates were slightly higher in the remainder of Arizona than in Maricopa County. The child extended response rate was 42 percent, resulting in an overall adult response rate of 19 percent.

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 and converting the survey responses to SPSS. 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, 2010 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 and functioning	Adult	Child
General health status	✓	✓
Height and weight	✓	✓
Limitations of activity	✓	✓
Health conditions	**Adult**	**Child**
Asthma	✓	✓
Diabetes	✓	
Heart disease, high blood pressure	✓	
Arthritis, gout, lupus, fibromyalgia	✓	
Gastrointestinal disorders	✓	
Bi-polar disorder, anxiety disorder, depression	✓	
Developmental disorders		✓
Mental health	**Adult**	**Child**
Mental health status	✓	
Psychological distress	✓	
Experiences of stress	✓	
Emotional functioning	✓	✓
Interpersonal relationships	✓	
Perceived need, use of behavioral health services	✓	
Reasons for not seeking treatment	✓	
Parent's behavioral and developmental concerns		✓
Behavioral and developmental concerns of school, doctor		✓
Health behaviors	**Adult**	**Child**
Dietary intake	✓	✓
Physical activity and exercise	✓	✓
Alcohol and tobacco use	✓	
Illegal drug use	✓	
Abuse of prescription drugs, steroids	✓	
Dental health	**Adult**	**Child**
Last dental visit	✓	✓
Not getting needed care		✓
Days missed from school due to dental problems		✓
Unmet needs		✓
Usual source of dental care		✓
Access to and use of health care		
Personal doctor	✓	
Visits to medical doctor, specialist in past year	✓	✓

Barriers to care	✓	
Unmet needs for care or prescriptions	✓	✓
Communication with doctor	✓	✓

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

Health insurance	Adult	Child
Current health insurance coverage and source	✓	✓
Coverage of prescription drugs, dental and behavioral health services	✓	✓
Coverage over past 12 months	✓	✓
Availability of coverage through employment	✓	*
Medical debt and its effects	✓	
Housing and neighborhood	**Adult**	**Child**
Type of housing	✓	
Neighborhood safety	✓	✓
Family nearby	✓	
Availability of food shopping, cultural facilities	✓	
Characteristics of neighbors	✓	
Use of parks	✓	
Volunteer service	✓	
Social interactions	✓	
Parental involvement	**Adult**	**Child**
Parental presence after school		
Marital status of parents		
Child's activities with family		✓
Child care and school attendance	**Adult**	**Child**
Current child care arrangements		✓
Difficulty finding care		✓
Quality of child care		✓
Employment	**Adult**	**Child**
Employment status, spouse's employment status	✓	*
Work in last week	✓	*
Hours worked at all jobs, spouse hours	✓	*
Tenure	✓	*
Employer size, spouse employer size	✓	*
Income	**Adult**	**Child**
Household income (annual before taxes)	✓	*
Number of persons supported by household income	✓	*
Food, housing insecurity	✓	
Receipt of Social Security disability, SSI	✓	
Participation in TANF	✓	✓
Participation in food stamps, WIC		✓
Respondent characteristics	**Adult**	**Child**
Age, gender	✓	✓
Race and ethnicity	✓	✓
Marital status	✓	
Education	✓	
Sexual orientation	✓	
Household composition	✓	
First language, English proficiency	✓	
Languages spoken at home		✓
Country of birth	✓	✓
Military service	✓	
County of residence	✓	✓

1 Sample Design

This chapter describes the sampling methods used in the 2010 Arizona Health Survey (AHS). AHS consisted of a stratified telephone random-digit-dialing (RDD) sample drawn using a list-assisted approach, with a supplemental sample drawn from listed telephone numbers expected to reach households with children aged 0-5. The supplemental sample was used to increase the number of child interviews because it was expected that fewer than 8 percent of contacted households in the RDD sample would have eligible children.

The first section describes list-assisted RDD sampling and the procedures implemented in AHS to save costs by reducing the number of calls to unproductive 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 landline 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.

1.2 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., 2010). By 2009, about 25 percent of households nationally had only cellular telephones in the US. 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.

1.3 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 surname to improve the efficiency of the sample.

1.3.1 Population of Interest

The AHS sample was intended to represent the adult (age 18 and older) residential population of Arizona, as well as young children (age 5 and under) in Arizona. 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 living in a household without a parent or legal guardian.

1.3.2 Adult and Child Sample Targets

The 2010 AHS is a dual frame survey with two components: a base random digit dialing (RDD) landline telephone sample and a supplemental sample from lists of telephone numbers likely to have households with children. It covers six geographic areas as shown in Table 1-1. The goal of the study was to complete 8,100 adult interviews (18 years of age or older) and 2,000 child interviews (under age 6), with targets for each of the six areas (five for children).

There are two methods for increasing the number of child interviews. The first method screens households from the landline frame for the presence of children in the household. Since it is expected less than 8 percent of contacted households would have eligible children, the cost of screening for households with children is high. An alternative is to use a different frame that consists of telephone numbers of households likely to have children.

In order to meet the goal of number of completed child interviews, we supplemented the sample from a special list of telephone numbers of households likely to have a child. This sampling strategy is based on the concept of a dual frame design, where the landline sample is supplemented with a much less expensive sample drawn from a list of telephone numbers likely to include members of the target group (i.e., households with children). The list frame does not have to be complete to be useful, although the more complete the list is, the greater the potential for overall efficiency. The composition of the list affects its efficiency (that is, the proportion of sampled numbers that lead to a member of the target group), but not the ability to produce unbiased estimates. Unbiased estimates can be produced if the list membership of every sampled unit (telephone number) from the other frame (landline in our case) can be determined.

Table 1-1. 2010 AHS stratum definitions and target sample sizes

Geographic Area (Stratum)	Counties	Adult interviews	Child interviews
1	Mohave, Coconino, Navajo, Apache, Yavapai	1,000	400
2	Yuma and La Paz	700	400
3	Graham, Greenlee, Cochise, Santa Cruz	700	200†
4	Pinal and Gila	700	200†
5	Pima	2,000	400
6	Maricopa	3,000	400
Total		8,100	2,000

†Strata 3 and 4 were combined for the child sample.

1.3.3 Sampling Households

Before sample selection, all telephone exchanges in Arizona were stratified into 6 strata defined by the county or group of counties as indicated in Table 1-1. Generally, a proportional allocation of the sample to area code-exchange within each primary stratum is expected to yield the most efficient sample design. However, this type of allocation does not produce the required number of interviews in the smaller strata. To achieve the survey goals, the sample was be disproportionately allocated, oversampling the smaller strata.

Since the geographic information required to stratify the sampling frame is only available at the exchange level[1], 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 the regions). To allocate these exchanges we used the coverage report for each region with the list of exchanges that provided service in the region. 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 a region if 50 percent or more of the serviced households in the exchange were located in the county.

Once the sampling strata were created, the sample of telephone numbers was drawn. Table 1-2 shows the numbers of cases drawn from the exchanges for the adult sample. The table also shows the number of telephone numbers drawn for the lists of telephone numbers of households likely to have children for the child sample. The number of telephone numbers selected

[1] 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.

had 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 were not residential. We increased the initial sample draw to deal with losses due to the following sources:

- Oversampling of Hispanic substrata;
- Nonworking, nonresidential, and never-answered numbers;
- Nonresponse to the screening interview; and
- Nonresponse to the extended interview.

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 the previous AHS and 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, 425,366 telephone numbers were drawn for the adult sample and 20,863 telephone numbers for the child sample 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. Table 1-2 summarizes the number of telephone numbers drawn in each stratum.

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

Geographic Area (Stratum)	Counties	Landline RDD sample	Child list sample
1	Mohave, Coconino, Navajo, Apache, Yavapai	42,350	4,807
2	Yuma and La Paz	73,547	2,803
3	Graham, Greenlee, Cochise, Santa Cruz	34,531	1,980
4	Pinal and Gila	30,510	5,755
5	Pima	90,746	4,208
6	Maricopa	159,139	1,310
Total		430,823	20,863

In order to oversample Hispanic adults, we created subsampling strata using available information about each telephone number in the sample. First we matched an address to each number using reverse telephone lookup services. This process created 2 groups within each stratum based on the results of the address match. The group that includes the telephone numbers with a matched address was further divided based on the surname. We used a Census Bureau surname file to classify Hispanic surnames of residents in the sample. The list contained Hispanic surnames of areas with high geographic concentration of Hispanics based on the 2000 decennial Census. Using

this file, we created an additional subgroup that includes all telephone numbers where the surname of the resident was likely to be Hispanic.

After these processes, three final substrata were created for a second phase of sampling. The substrata were (1) telephone numbers with a matched address where the surname of the resident was Hispanic, (2) telephone numbers with a matched address where the surname of the resident was not Hispanic, and (3) telephone numbers without a matched address. In order to oversample Hispanic households, all telephone numbers in the Hispanic stratum were kept in the sample while the other two strata are subsampled at the same rate in the second phase. The sampling rate was computed so that the expected number of Hispanic adult interviews was 20 percent of the total sample. Table 1-3 shows the number of telephone numbers by subsampling strata, sampling rates and number of telephone numbers retained in the sample.

Table 1-3. Substrata sample sizes and subsampling rates

Stratum	Substratum	Drawn sample	Subsampling rate	Retained sample
1	Hispanic	1,012	1.00	1,012
	Non-Hispanic	11741	0.65	7,665
	No address	29,597	0.65	19,321
2	Hispanic	7,706	1.00	7,706
	Non-Hispanic	15,218	0.62	9,426
	No address	50,623	0.62	31,355
3	Hispanic	2,193	1.00	2,193
	Non-Hispanic	5,583	0.58	3,258
	No address	26,755	0.58	15,614
4	Hispanic	1,654	1.00	1,654
	Non-Hispanic	8923	0.63	5,662
	No address	19,933	0.63	12,649
5	Hispanic	5,457	1.00	5,457
	Non-Hispanic	24,057	0.62	15,008
	No address	61,232	0.62	38,200
6	Hispanic	6,661	1.00	6,661
	Non-Hispanic	39,133	0.61	23,767
	No-address	113,345	0.61	68,839
Total	Hispanic	24,683	1.00	24,683
	Non-Hispanic	104,655	0.62	64,786
	No address	301,485	0.62	185,978
Total		430,823		275,447

1.3.4 Increasing the Efficiency of Data Collection

Special data collection procedures are often implemented before data collection to reduce costs and to increase efficiency. We used tritone purges and directory matching to remove 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 from this survey. Table 1-4 shows the CSS result codes as well as the distribution of the sampled telephone numbers. Approximately 57 percent of the sampled numbers (result codes CP, LB, FM, NR NW, and some UB) were excluded from dialing. The remaining 126,340 numbers were loaded into the CATI system.

Table 1-4. CSSR result codes and their distribution in the AHS sample

CSSR result code	Description	Number of telephones	Percentage
CP	Agent dispositioned cell phone	83	0.03
DK	Undetermined	82,825	27.95
FM	Fax/modem	8,175	2.76
LA	Language barrier	1,896	0.64
LB	Listed business	8,164	2.76
NR	No ring back	3,014	1.02
NW	Nonworking	137,055	46.25
PM	Privacy manager	9,672	3.26
RS	Residence	29,987	10.12
UB	Unlisted business	13,621	4.60
WR	Pro-t-s wireless detection	1,818	0.61
Total		296,310	100.00

1.4 Within-Household Sampling

Once the sample of telephone numbers was selected, interviewers called the numbers, and conducted interviews with sampled persons within the household. For the adult sample, the AHS design called for the random selection of one adult from all the adults in each sampled household. In addition, in those households with children (under age 6), one child was sampled and

a parent was interviewed about the child. In the supplemental child sample, no adult was selected and one child was sampled among all eligible children in the household.

To reduce the burden on the respondents and increase the utility of the data, children in the landline RDD sample were selected using a *linked* sampling approach. In this approach, children 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).

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

The linked approached was not used in the child supplemental samples because no adult was selected. In this case, children are selected in the first phase. The probability of selection of the sampled child is computed as the inverse of the number of eligible children in the household.

1.4.1 Child First Procedure

To increase the number of child interviews, they could be conducted before 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 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 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 could be sampled and the appropriate interview conducted without waiting for the completion of the adult interview.

1.4.2 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.

1.5 Achieved Sample Sizes

The goals for AHS were stated in terms of the total number of completed adults and children 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 1-5 shows the number of completed adult interviews by two methods of classifying the geographic area in which the sampled person 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 before, 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 fifth column in the table uses the self-reported residence county of the 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. Table 1-4 shows that the overall target for the number of completed AHS adult interviews was met, and each of the stratum targets was met except for Maricopa County.

Table 1-5. Number of completed adult interviews by sampling and self-reported stratum

Geographic Area	Counties	Sampling location		Self-reported location	
		Completed interviews	Percentage	Completed interviews	Percentage
1	Mohave, Coconino, Navajo, Apache, Yavapai	1,059	12.9	1,053	12.8
2	Yuma and La Paz	759	9.2	743	9.0
3	Graham, Greenlee, Cochise, Santa Cruz	770	9.4	755	9.2
4	Pinal and Gila	777	9.5	798	9.7
5	Pima	2,168	26.4	2,143	26.1
6	Maricopa	2,682	32.7	2,723	33.2
Total		8,215	100.0	8,215	100.0

Table 1-6 shows the number of completed child interviews by sampling and self-reported stratum. As with the adult interviews, each of the targets was met except for one, Stratum 1.

Table 1-6. Number of completed child interviews by sampling and self-reported stratum

Geographic Area (Stratum)	Counties	Sampling location		Self-reported location	
		Completed interviews	Percentage	Completed interviews	Percentage
1	Mohave, Coconino, Navajo, Apache, Yavapai	362	16.9	364	17.0
2	Yuma and La Paz	416	19.4	407	19.0
3	Graham, Greenlee, Cochise, Santa Cruz, Pinal and Gila	571	26.6	537	25.0
5	Pima	400	18.6	412	19.2
6	Maricopa	399	18.6	428	19.9
Total		2,148	100.0	2,148	100.0

2 Survey Questionnaires

The 2010 AHS interview structure was modeled on that of the 2008 survey, with additional items drawn from the Behavioral Risk Factor Surveillance Survey (BRFSS), the Medical Expenditure Panel Survey (MEPS), and other sources. AHS interviews could include, for a given household, an adult or child extended questionnaire or both. 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 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 "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 6 years of age living in the household; and
- The number and use (home, business) of telephone numbers ringing into the household.

For the listed telephone number sample, the screening interview was similar, except that no adults were sampled. If a household did not include one or more children under 6, it was considered ineligible.

2.2 Overall Structure of AHS Interviews

Given the number of 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; and
- A "most knowledgeable adult" (MKA) for the child extended interview.

In practice, one adult often filled multiple roles in households with 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. 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. 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 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.

For cases other than those meeting the child-first criteria, the screening interview resumed in the middle of Section F of the Adult Extended Questionnaire, with enumeration of children under age 6 in the household and determination of which age-eligible children the sampled adult was the parent or legal guardian.

This information was used by the CATI program to select one eligible child among those for whom the sampled adult was the parent or legal guardian. Children who did not have a parent or legal guardian in the household were not eligible for selection.

Because sampling children was part of the adult interview except for child-first cases, the adult interview had to be completed first. After an adult interview was completed for non-child-first cases, if a child was selected the sampled adult was asked to identify the respondent for the child interview.

2.3 Extended Interviews

AHS included separate extended interviews for adults and children. 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 30 minutes in mean administration time, and a child questionnaire that would not exceed 25 minutes. Child interviews without an adult interview having been done first were not expected to exceed 40 minutes.

In March 2010, SLHI provided Westat with drafts of the adult 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 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.

In order to achieve cost savings in data collection, Westat used a different CATI system than had been used for the 2008 survey. Thus, the AHS 2010 questionnaires were programmed

from scratch in Voxco, following the draft instruments provided by SLHI. Specifications documenting the CATI program were prepared after the fact.

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), and data preparation staff. After the pilot test and then again during the first few weeks of the full 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 includes 9 sections:

A. **Health Status and Conditions** – General health, height and weight, asthma, diabetes, high blood pressure, heart disease, chronic inflammatory conditions, gastrointestinal disorders, bi-polar disorder, anxiety disorder, depression, limitations of activity.

B. **Healthy Behaviors** – Physical activity, dietary intake.

C. **Behavioral Health** – Mental health status, psychological distress, experiences of stress, emotional functioning, interpersonal relationships, perceived need for and use of behavioral health services, reasons for not seeking treatment.

D. **Risky Behaviors** – Smoking, drinking, use of illegal drugs, abuse of prescription drugs or steroids.

E. **Demographics** – Age, gender, race, ethnicity, marital status, country of birth, first language, English proficiency, household composition, education, sexual orientation, employment status of self and spouse, household income and receipt of Social Security disability, SSDI, or TANF.

F. **Health Insurance –** Current coverage by public or private plans, coverage of prescription drugs, dental care, and behavioral health services, coverage over past 12 months, availability of coverage through employment, and medical debt.

G. **Health Care** – Personal doctor, medical doctor and specialist visits in past year, barriers to care, communication with doctor, unmet needs for care or prescriptions, most recent dental care.

H. **Housing and Neighborhood** – Type of housing, feeling safe in neighborhood, characteristics of neighborhood and its residents, whether family nearby, availability of food shopping and cultural facilities, use of parks, volunteer service, social interactions.

I. **Closing** – County of residence, willingness to participate in follow-up study.

The child extended questionnaire comprises ten sections:

A. **Demographics and Health Conditions –** Gender, age, height, and weight, whether in school, asthma, developmental disorders, limitations of activity.

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

C. **Diet, Physical Activity, and Park Use –** Types of food eaten, physical activity, use of parks and library.

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

E. **Parental Involvement, Concerns, Mental Health –** Parental involvement with child; developmental and behavioral concerns of parent, school, and doctor; treatment for developmental issues; emotional health and seeking help for emotional issues.

F. **Demographics, Part II –** Race and ethnicity, use of Indian health programs, country of birth for child and parents, respondent's English proficiency and level of education.

G. **Public Programs –** Participation in TANF, Food Stamps, and WIC.

H. **Child Care and Early Childhood Education –** Types of child care used, difficulty finding care, quality of child care.

I. **Health Insurance Coverage –** Current coverage and source, coverage of prescription drugs, dental and behavioral health services, coverage over past 12 months.

J. **Closing –** County of residence, willingness to participate in follow-up study.

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.

2.3.3 Translation of Questionnaires

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 the 2008 AHS, 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.

2.3.4 Pilot Test

Westat conducted a small pilot test of the AHS interviews the week of May 3, 2010, essentially starting the statewide data collection with a small group of interviewers. The purpose of this test was to ensure that the CATI instrument was working as planned and that responses were being stored appropriately in the database.

3 Data Collector Recruiting and Training

3.1 Organization of the Telephone Research Center

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

Westat's computing systems and telephony capabilities enable the networked combination of geographically diverse data collector 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 data collectors, 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 experienced Westat data collectors. Eighty-one of the data collectors had previously worked on the 2009 California Health Interview Survey (CHIS). Due to the similarity of AHS to CHIS, these data collectors were able to make this transition smoothly after a two-hour training on the CATI system used for AHS, which differed from that used for CHIS. One hundred sixty-seven data collectors who had not worked on CHIS completed the full 6½-hour project specific training, which included two hours of web-based training, 1½ hours of a live WebEx session and 3 hours of a dyad role play session. After completion of project trainings, 248 data collectors worked on the study.

Two hours of project-specific training involved interviewers completing a web-based distance learning session. This training included completion of a full adult and child 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 a pronunciation guide, answers to commonly asked questions, refusal avoidance ideas, the toll-free numbers and web site for reference.

Following completion of the web based distance learning session, a live WebEx session lasting 1½ hours was required. This session brought groups of trainees together providing a forum for asking questions and sharing ideas for gaining cooperation. The trainer shared his/her screen for reviewing contact procedures covering situations for recording outcomes of calls when the interview was not completed. The WebEx session was followed by trainees being paired and conducting two scripted role plays for practicing administration prior to live production. The role plays were monitored by supervisory staff to assess readiness for data collection.

Trainings began on March 20, 2010, and continued until the required number of data collectors were working on AHS. After all data collectors started production, they received supplemental training on specific questionnaire issues that arose after training. Data collectors already trained in refusal conversion for CHIS 2009 were recruited to work on refusal conversion cases for AHS.

3.3 Training for Spanish-Language Interviewing

All Spanish bilingual data collectors 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 data collectors initially worked in English. Once the Spanish-language instrument was ready, bilingual data collectors 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 data collectors began interviewing at the TRCs in Spanish, they were monitored closely by Spanish-speaking supervisors. The first priority in CATI for Spanish bilingual data collectors were cases from the work class identified as speaking Spanish. Bilingual Spanish data collectors 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 data collectors 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 data collector.

3.4 Quality Control and Post-Training Support for Data collectors

3.4.1 Evaluating Data Collector Performance

Data collector 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 data collectors 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. Data collectors whose performance fell below acceptable levels attended additional coaching sessions with an emphasis on gaining respondent cooperation and answering respondent questions.

3.4.2 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.

3.4.3 Communication with Data Collectors

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

During work shifts, all data collectors 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. Data collectors 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 data collector to provide instantaneous guidance and support or post-interview feedback.

A website was provided which data collectors 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 data collectors 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, data collector questions and monitoring notes.

Each data collector 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.

Data collectors 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 data collectors and to ask any questions about AHS or general procedures.

Data Collection 4

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, and handling inbound calls to Westat's AHS toll-free number. Data collection began May 4, 2010, and ended August 15, 2010. For the last 3 weeks of the field period, only child interviews were attempted.

4.1.1 Preparation and Release of Sampled Numbers

As described in Chapter 1, the AHS sample draw comprised 430,823 telephone numbers in the main RDD sample and 20,863 in the child supplemental sample. After subsampling as described in Chapter 1, about 57 percent of the subsampled numbers were purged as nonworking or business. After the purge and subsampling, 126,340 numbers were available to be called in CATI.

4.1.2 Sample Management

All of the numbers released to CATI were divided into "release groups," or random subsets of the overall samples.

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 data collectors 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 data collectors. 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 data collectors;

- **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 data collectors selected to work and trained as refusal converters. There were three different refusal work classes: screener initial refusal, extended refusal (other than adolescent and adolescent permission), 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; and

- **Language (Spanish)—**Any case determined or suspected to require a Spanish bilingual data collector to re-contact; available only to the appropriate bilingual data collectors.

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. Data collectors 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.

When study telephone numbers were called, the program was configured to display an Arizona telephone number on caller ID systems (602-252-1288) with the text "AZ Health." Respondents calling the Arizona number would automatically have the call forwarded to the toll-free number to be received by an operator knowledgeable about the study so all questions could be answered or requested appointments set.

Between the start of data collection in May 2010 and the end in August 2010, 1,696 calls were made to Westat staff fielding respondent calls to the 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 available on the AHS web site. Westat data collectors 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 calling for St. Luke's Health Initiatives in Phoenix. We are doing a study about the health of the people of Arizona and about health care. I am not asking for money--this is a scientific study called the Arizona Health Survey.
>
> We will call you back in the next few days.

Overall, about 47.5 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 50.7 percent, and it was about 46 percent for the remainder of Arizona.

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. However, actual sample management was considerably affected by the disproportionate pace in which goals by strata (both geographic and child vs. adult) were filled. Many active sample records were not worked to completion because they were in a stratum that met its goals before the record received the maximum attempts. Also, towards the end of data collection, additional sample was loaded to help reach specific goals by specific deadlines and with reasonable efficiency and many of these sample records were not worked through the full protocol.

For non-final call results (other than contacts that led to an appointment date and time for the next call attempt), call-back rules ensured that a minimum amount of time elapsed before the next call attempt to avoid excessive calls to respondents in a relatively short time window. General call-back delays were as follows:

- Ring No Answer, Answering Machine – 12 hours;
- Busy signal - 15 minutes (with a 12 hour delay after the 4th consecutive busy signal);
- Fax/Modem, Dropped Calls, Questionable Rings, Dead Air – 12 hours.

Additionally, the CATI program assigned final results based on call histories as follows:

- Any 3 results indicating a non-residential or non-working number (Fax/Modem, Questionable Rings, Etc,) were finalized with the NW (non-working code);
- Cases that were never answered by either a person or machine after 7 attempts were finalized as NA or NZ (NZ cases are NA cases that include a predictive dropped call);
- Cases that were never answered by a person but were answered by an answering machine were coded NM or NX (NX cases are NM cases that include a predictive dropped call);
- Cases with a second refusal during conversion attempts were coded RB;
- Cases worked for a second refusal conversion attempt and encountered a third refusal were coded R3.

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 data collector 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 of this high limit and the way sample was worked through the course of the study to meet strata-specific goals, very few cases received this many attempts.

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 data collectors 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. There were 22,334 matching numbers in all. About 20.2percent of these cases resulted in completed screeners and 8.3 percent resulted in completed adult and/or child interviews Forty percent of the screeners completed with this group were done in Spanish. Of those that completed the adult interview, 93 percent reported being of Hispanic or Latino origin, and they represented 76 percent of the total number of AHS adult respondents reporting Hispanic/Latino ethnicity.

Other sample cases determined to require a Spanish bilingual data collector were put into the Spanish-language work class, and became available to bilingual data collectors. Virtually all Spanish-language cases reported being Hispanic and about 37% of completed adults interviews with Hispanics were conducted in Spanish.

Overall, some 655 adult interviews were conducted in Spanish, about 8.0 percent of the total. The sample stratum including Yuma and La Paz Counties had the highest proportion of

Spanish-language interviews 28.3 percent), and the stratum including Mohave and Coconino Counties had the lowest (2.2 percent). Among child interviews, 335 were completed in Spanish (15.6%). As with adult interviews, the sample stratum including Yuma and La Paz had the highest proportion of Spanish interviews, with 38.5 percent. The lowest percentage of child interviews completed in Spanish was in the sample stratum including Pinal and Gila, at 3.2 percent.

The mean interview times differed by language. The adult interview was about 35.3 minutes in English and 46 minutes in Spanish. The child interview averaged 30.7 minutes in English, 40.7 minutes in Spanish

4.2.5 Refusal Conversion

At each stage of the interview process, Westat data collectors made extensive conversion efforts for refusals that were not judged to be hostile or abusive. When a refusal was first encountered, the data collector 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 were not called again for a minimum of 8 days Initial refusals that were considered hostile or abusive received were removed them from conversion attempts. Data collectors 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.

4.3 Detailed Results by Outcome

This section will present detailed tables of outcomes at each interview level – screener, adult, and child. Data collectors 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, (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.

As described above, many cases were still in interim status when we stopped calling them because either their stratum was at its goal or because we ended the study due to the deadline.

To simplify reporting, these cases are reported below with the IN result code to indicate they were interim cases.

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. Ineligibility was rare except when goals were met for adult completes and eligibility criteria included the presence of at least one child under the age of 6 in the household. Completed screeners accounted for about 17 percent of total numbers called. Refusals accounted for about 20 percent of all sample.

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

	N	Percentage
COMPLETE SCREENER	**21,912**	**17.3%**
INELIGIBLE	**6,309**	**5.0%**
No one 18+	88	0.1%
No Child in Household	6,162	4.9%
More than 9 Un-related adults	3	0.0%
Respondent Deceased	4	0.0%
Out of State	52	0.0%
OUT OF SCOPE	**15,621**	**12.4%**
Non-Working	9,278	7.3%
Non-residential	6,343	5.0%
NON-CONTACT	**39,073**	**30.9%**
Final Never Answered	17,936	14.2%
Final Never Contacted	21,137	16.7%
REMAINING INTERIM	**17,536**	**13.9%**
REFUSALS	**25,089**	**19.9%**
Hostile Refusals	261	0.2%
Inbound Refusals	148	0.1%
Second Refusals (after conversion attempt)	20,152	16.0%
Third Refusals (after second conversion attempt)	1,095	0.9%
Not attempted for Conversion	3,433	2.7%
PROBLEMS	**800**	**0.6%**
Final Language Problem	366	0.3%
Final Speech/Hearing Problem	19	0.0%
Not available in Field Period	77	0.1%
Respondent Sick or Mentally Incapable	65	0.1%
Other Problems	268	0.2%
Enumeration Error	4	0.0%
Duplicate Record	1	0.0%

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 42 percent of sampled adults. However, many pending adult completes were abandoned as goals for adult completes by strata and overall were met.

4.3.3 Child Extended Interview

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. For some households only child interviews were conducted and no adults were selected for an adult interview. For the last three weeks of the study only child interviews were conducted for all households called. 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 54 percent of sampled children.

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

	N	Percentage
COMPLETE ADULT INTERVIEW	**8,217**	**42.0%**
INELIGIBLE	**92**	**0.5%**
No one 18+	2	0.0%
No Child in Household	60	0.3%
Respondent Deceased	12	0.1%
Out of State	18	0.1%
OUT OF SCOPE	**428**	**2.2%**
Non-Working	331	1.7%
Non-residential	97	0.5%
NON-CONTACT	**2,016**	**10.3%**
Final Never Answered (After Screener)	275	1.4%
Final Never Contacted (After Screener)	1,741	8.9%
REMAINING INTERIM	**3,723**	**19.0%**
REFUSALS	**4,141**	**21.2%**
Hostile Refusals	74	0.4%
Inbound Refusals	39	0.2%
Second Refusals (after conversion attempt)	3,144	16.1%
Third Refusals (after second conversion attempt)	8	0.0%
Not attempted for Conversion	876	4.5%
PROBLEMS	**934**	**4.8%**
Final Language Problem	70	0.4%
Final Speech/Hearing Problem	64	0.3%
Not available in Field Period	74	0.4%
Respondent Sick or Mentally Incapable	196	1.0%
Other Problems	186	1.0%
Enumeration Error	51	0.3%
Duplicate Record	1	0.0%
No Longer at this Number	292	1.5%

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

	N	Percentage
COMPLETE CHILD INTERVIEW	**2,148**	**54.1%**
INELIGIBLE	**90**	**2.3%**
No Child in Household	87	2.2%
Respondent Deceased	2	0.1%
Out of State	1	0.0%
OUT OF SCOPE	**84**	**2.1%**
Non-Working	78	2.0%
Non-residential	6	0.2%
NON-CONTACT	**545**	**13.7%**
Final Never Answered (After Screener)	76	1.9%
Final Never Contacted (After Screener)	469	11.8%
REMAINING INTERIM	**523**	**13.2%**
REFUSALS	**455**	**11.5%**
Hostile Refusals	6	0.2%
Inbound Refusals	12	0.3%
Second Refusals (after conversion attempt)	253	6.4%
Third Refusals (after second conversion attempt)	39	1.0%
Not attempted for Conversion	145	3.7%
PROBLEMS	**127**	**3.2%**
Final Language Problem	5	0.1%
Final Speech/Hearing Problem	1	0.0%
Not available in Field Period	5	0.1%
Respondent Sick or Mentally Incapable	2	0.1%
Other Problems	74	1.9%
Enumeration Error	5	0.1%
Duplicate Record	1	0.0%
No Longer at this Number	34	0.9%

4.4 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. The RR4 and RR3 definitions 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 the unweighted AAPOR RR4. Note that this response rate does not reflect the fact that telephone numbers were sampled with different selection probabilities in the AHS strata from different frames in the response rate computation. Table 4-5 shows the screener, extended and overall response rates for the adult and child interviews by stratum. The adult response rate was computed using all records used to screen for adults while the child response rates includes the adult and child samples. The overall response rate is computed as the product of the screener and extended response rates. The screener response rates range from 44 to 52 percent for the adult sample, and 25 to 47 percent for the child sample. The Stratum 2 rate for Yuma and La Paz Counties reflects the fact that a large sample was worked only near the end of the field period to reach the child interview target. Adult interview response rates ranged from 39 to 45 percent, while child interview rates ranged from 54 to 61 percent.

Table 4-4. AHS 2010 response rates by sample (adult or child), stratum, and instrument

Stratum	Response rate		
	Screener	Extended	Overall
Adult Sample			
Statewide	0.46	0.42	0.19
1. Mohave, Coconino, Navajo, Apache, Yavapai	0.48	0.43	0.21
2. Yuma, La Paz	0.51	0.41	0.21
3. Graham, Greenlee, Cochise, Santa Cruz	0.52	0.45	0.23
4. Pinal, Gila	0.50	0.43	0.22
5. Pima	0.46	0.44	0.20
6. Maricopa	0.44	0.39	0.17
Child Sample			
Statewide	0.39	0.57	0.22
1. Mohave, Coconino, Navajo, Apache, Yavapai	0.40	0.59	0.24
2. Yuma, La Paz	0.25	0.58	0.14
3. Graham, Greenlee, Cochise, Santa Cruz	0.47	0.61	0.29
4. Pinal, Gila	0.40	0.60	0.24
5. Pima	0.43	0.54	0.23
6. Maricopa	0.43	0.51	0.22

Data Processing 5

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 included;

- Racial/ethnic identification (adult and child interviews);
- country of birth (adult and child interviews);
- first language/language spoken at home (adult and child interviews);
- sexual orientation (adult interview);
- why person not covered by an employer's health plan (adult and child interviews);
- why person had no health insurance (adult and child interviews);
- child's limiting health condition (child interview).

Westat data preparation staff reviewed these responses and up-coded them to the existing categories whenever possible. For the race/ethnicity items, some "other" responses were standardized. Responses such as Mexican, Latino, and Hispanic, were updated to HISPANIC-LATINO.

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.

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.

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.

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 SPSS data sets. The Voxco CATI program has a utility to export SPSS files directly. Survey weights and other variables created outside of the CATI system, in SAS, were added to the SPSS files . . .

5.3 Race and Ethnicity Coding

The AHS survey items about ethnicity and race were changed from those used in the 2008 survey. 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 AE129 (CE084 in the child questionnaire) asked respondents for their race/ethnicity: "Please tell me which of the following **best** describes you. Do you consider yourself: Caucasian or White; Latino or Hispanic; Black or African American; Asian, Pacific Islander or Native Hawaiian; Native American or American Indian?" The "other specify" race was recorded in text. The following question asked "Is there any other group that you use to describe yourself?" A "yes" response would then lead to the first question being asked again, with the category picked the first time eliminated. In 2008, the questions were consistent with those asked in the U.S. Census: the first item was about Hispanic ethnicity, and the second question about racial identity, with a somewhat different set of response categories and an instruction to "code all that apply." Many Hispanic respondents in 2008 did not select one of the racial categories, but reported their ethnicity as an "other race."

Despite the differences in how the questions were asked, the AHS procedures for coding the 2010 responses were consistent with the ones used to code the 2000 Census data as 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.

Upcoding 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 "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.

- If the "other specify" response to the second race question would have been upcoded to the same category that was recorded in the first question, then the response was deleted.

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).

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 from multiple frames. 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.

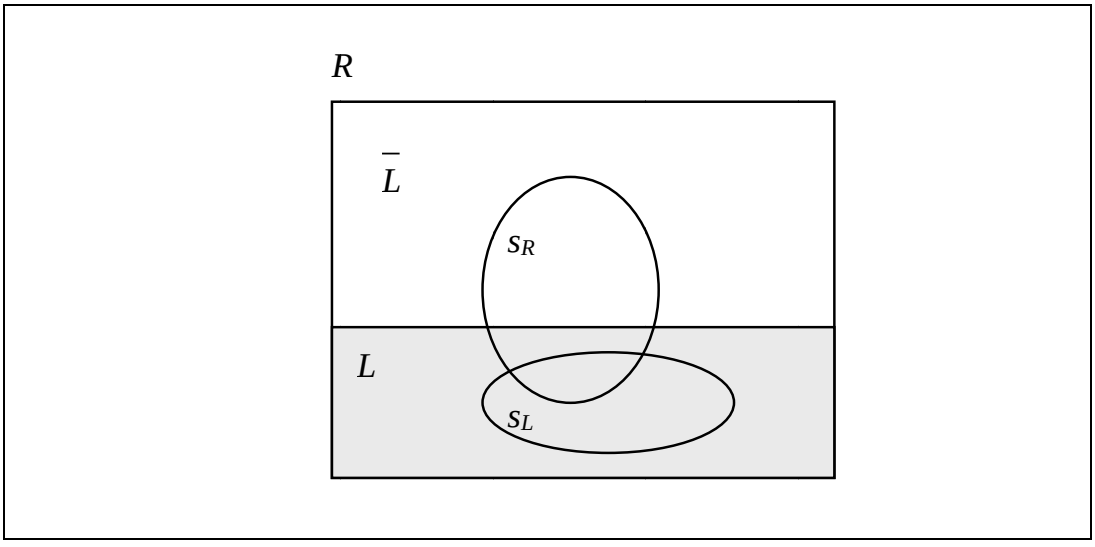
6.2 Household Weighting

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. In AHS 2010, telephone numbers were drawn from a landline frame and mutually exclusive list frames. The base weights reflect the multiple probability of selection of telephone numbers among the landline and list frames.

Figure 6-1 shows the relationship between the landline frame and a single list frame for a single sampling stratum. The figure also shows the relationship between the samples drawn from each frame. In order to create the household base weights, we consider all landline telephone households in Arizona as either being on the list (L) or as only being eligible for sampling from the landline sample ($\overline{L}$) as shown in Figure 3-1. The relationships are discussed in detail below.

Figure 6-1. Relationship between the landline frame (R), landline sample (s_R), list frame (L), and list sample (s_L) for a single stratum



* The figure is not drawn to scale. The sizes of the list frame (L) and list and landline samples (s_L and s_R) are smaller than shown in the figure.

The notation in the figure follows:

R	the landline frame containing all telephone numbers;
L	the list frame (i.e., telephone of households likely to have children);
$\bar{L}$	all telephone numbers not found on the list – we assume that all the numbers in the list are found in R, and $R = L \cup \bar{L}$;
s_R	the simple random sample drawn from the frame *R*; and
s_L	the simple random sample drawn from the frame *L*.

Define the following:

N_R	the number of telephone numbers in the frame R;
N_L	the number of telephone numbers in the frame L;
n_R	the sample size (number of telephone numbers) of s_R; and
n_L	the sample size (number of telephone numbers) of s_L.

Notice that the landline sample s_R can be separated into two parts: s_{RL}, the portion of s_R that is found in the list (L) and $s_{R\bar{L}}$, the portion of s_R that is not found in the list $(\bar{L})$. The sample sizes for each portion are n_{RL} and $n_{R\bar{L}}$, respectively, and $n_R = n_{RL} + n_{R\bar{L}}$.

Consider L and $\bar{L}$ as two separate strata within the frame R. Since s_R is a simple random sample within R, the sample $s_{R\bar{L}}$ can be viewed as a simple random sample of size $n_{R\bar{L}}$ drawn from the $N_{\bar{L}}$ elements from stratum $\bar{L}$. Similarly, the sample s_{RL} can be viewed as a simple random sample of size n_{RL} drawn from the N_L elements from stratum L. In stratum L, there is a second sample s_L (the list sample). Since both samples s_L and s_{RL} are simple random samples, we can view them as a single sample of size $n_{RL} + n_L$ drawn from the N_L elements from stratum L. Notice that s_{RL} and s_L are not necessarily mutually exclusive; i.e., s_{RL} and s_L may contain duplicate telephone numbers. These numbers were removed from s_L during the sample selection.

The landline and list base weights can be expressed as follows:

- For sampled records that could only be sampled from the landline frame (landline numbers not found in the list L):

$$HHBW_{\bar{L}i} = \frac{N_{\bar{L}}}{n_{R\bar{L}}};$$

- For sampled records from the list and sampled records from the landline frame that are found in the list L (duplicate telephone numbers were eliminated from the list):

$$HHBW_{Li} = \frac{N_L}{n_{RL} + n_L}.$$

Creating these weights required classification of every telephone number by whether or not it was on the list irrespective of how it was sampled. It is easy to show that the resulting weights are composite weights derived by averaging the landline and list samples using a composite factor proportional to the sample sizes. Thus, this base weight produces an unbiased estimate in the traditional design-based framework.

6.2.2 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:

$$HHA2W_i = HHA2F_i * HHBW_i,$$

where $HHA2F_i$ is the multiple telephone adjustment factor computed as:

$$HHA2F_i = \begin{cases} 0.5 & \text{If household } i \text{ has more than one residential telephone number} \\ 1 & \text{Otherwise} \end{cases}.$$

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 and child who completed the extended interview. 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 and child interviews. Furthermore, the selection of the child is different for cases from the landline and the list samples.

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) in the landline sample. 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

$$ADA0W_j = ADCNT_i \cdot HHA2W_i ,$$

where $ADCNT_i$ is the total number of adults in household *i*, and $HHA2W_i$ is the adjusted household weight.

6.3.2 Child Weight

The creation of the child base weight depends on how the child was sampled. In the landline sample, children linked to the sample adult were first identified and then sampled. In the list samples, children were sampled directly from among all eligible children in the household.

In the landline sample, children not selected in the child-first procedure were sampled in Section E of the adult extended interview. Consequently, weights for those children must be further adjusted to account for nonresponse at the adult interview level. On the other hand, weights of the children sampled during the child first procedure were not adjusted for adult nonresponse.

Telephone numbers in the landline sample were classified into completion groups by Section E completion status and their child-first interview status as shown in Table 6-1.

Table 6-1. Section E completion groups

Section E completion group (SECGST)	Child-first interview?	Section E completed by adult?	Description
C1st	Yes	N/A	Households where the child-first interview procedures occurred
NC1stEC	No	Yes	Households where the child-first interview procedures did not occur and section E was completed
NC1stENC	No	No	Households where the child-first interview procedures did not occur and section E was not completed

To account for adults who did not complete Section E of the adult interview (hence, no child could be sampled), the household weight $HHA2W_i$ was adjusted. We refer to this adjusted weight as the Section E adjusted household weight, $HHA2W_i$, and it is

$$HHA3W_i = HHA3F_c * HHA2W_i,$$

where

$$HHA3F_c = \begin{cases} \dfrac{\sum_{i \in SECGST = NC1stGC, NC1stGNC} \delta_i(c)}{\sum_{i \in SECGST = NC1stGC,} \delta_i(c)} & \text{If } i \in SECEST = NC1stEC \\ 0 & \text{If } i \in SECEST = NC1stENC \\ 1 & \text{If } i \in SECEST = C1st \end{cases},$$

and where the section E completion groups *C1st*, *NC1stEC*, and *NC1stENC* are defined in Table 6-1, c denotes the Section E nonresponse adjustment cell, and $\delta_i(c)=1$ if the adult belongs to cell c and is zero otherwise. Following this adjustment, the weights were positive for all households with sampled adults who completed section E and either completed, partially completed, or did not complete the adult interview[2]. Note that this adjustment can be considered as a household adjustment in addition to the household weight.

The Section E nonresponse adjustment cells were created within sampling strata considering the presence of children. The information about the presence of children is collected during the screener interview.

[2] If the adult interview was not completed in a case that did not use the child-first procedure, no attempt was made to complete a child/teen interview.

The child weight in the landline sample 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 in the landline sample 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 eligible children present in the household.

Since selecting a child in the landline sample 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[3] 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

$$CHPROB_j = \frac{1}{ADLTCNT} \cdot \frac{1}{SACH} \ ,$$

where $ADLTCNT$ is the number of adults in the household and $SACH$ is the total number of children associated to the sampled adult. 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

$$CHPROB_j = \frac{1}{ADLTCNT}\left(\frac{1}{SACH} + \frac{1}{SPCH}\right) \ ,$$

where $ADLTCNT$ and $SACH$ are defined as before and $SPCH$ is the total number of children associated with the spouse/partner of the sample adult.

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

Children in the list samples were selected directly among all eligible children in the household. Accordingly, $CHPROB_j$ is the probability of selecting a child associated computed as

$$CHPROB_j = \frac{1}{CHLDCNT} ,$$

where $CHLDCNT$ is the total number of eligible children in the household.

6.3.3 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 $0 < t_i < 1$.

Eleven adult and 21 child weights were trimmed[4]. 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.

6.3.4 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.

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,

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

[4] 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.

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 for the adults sample and five raking dimensions for the child sample are shown in Table 6-3. The dimensions were created by combining demographic variables (age, sex, race, and ethnicity) and geographic areas. The dimensions include variables such as household tenure (rent, own) and educational attainment 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 except for those cases where control totals were not available.

Table 6-2. Definitions of the dimensions used in raking in adult sample

Dimension	Level	Description	Categories	
1	Stratum	Sex (2) x Age groups (3)	11	18 to 34 years, male
			12	35 to 64 years, male
			13	65 or older, male
			21	18 to 34 years, female
			22	35 to 64 years, female
			23	65 or older, female
2	Stratum	Age groups (8) (collapsed where necessary)	1	18 to 24 years
			2	25 to 29 years
			3	30 to 39 years
			4	40 to 49 years
			5	50 to 64 years
			6	65 years or older
3	Stratum	Race-ethnicity (3)	1	Hispanic, 18 years or older
			2	Non-Hispanic White, 18 years or older
			3	Other, under 18 years or older
4	Stratum	Education attainment (4) (collapsed where necessary)	111	Less than High School, 18 years or older
			121	High School or equivalent, 18 years or older
			131	More than High School, 18 years or older

Table 6-2. Definitions of the dimensions used in raking in adult sample (continued)

Dimension	Level	Description	Categories	
5	Stratum	Household tenure (2) x Age group (collapsed where necessary)	111 111	Own, 18 years or older Rent, 18 years or older
6	State	Age group (14) (collapsed where necessary)	1 2 3 4 3 4 5 6 7 8 9 10 11	18 to 19 years 20 to 24 years 25 to 29 years 30 to 34 years 25 to 29 years 30 to 34 years 35 to 44 years 45 to 54 years 55 to 54 years 55 to 64 years 65 to 74 years 75 to 84 years 85 years old or older

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

Dimension	Level	Description	Categories	
1	Stratum	Sex (2)	1 2	Male, under 6 years old Female, under 6 years old
2	Stratum	Race-ethnicity (3) x Age groups (1)	1 2 3	Hispanic, under 6 years old Non-Hispanic White, under 6 years old Other, under 6 years old
3	Stratum	Household tenure (3) x Age groups (1)	1 2	Own, 18 years or older Rent, 18 years or older
2	Stratum	Age groups (8) (collapsed where necessary)	1 2 3	Under 2 years 2 to 3 years 4 to 5 years
4	State	Age	1 2 3 3 3 3	Under 1 year old 1 year 2 years 3 years 4 years 5 years

6.4.2 Raking Factors

Table 6-4 shows the overall adjustment factors for the adult and child 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 and child interviews by sample characteristics

Characteristic	Adult	Child
Total	15.63	13.54
Self-reported stratum		
Mohave, Coconino, Navajo, Apache, Yavapai	15.73	13.03
Yuma and La Paz	15.49	14.61
Graham, Greenlee, Cochise, Santa Cruz	10.53	13.92
Pinal and Gila	12.80	8.01
Pima	14.69	16.37
Maricopa	16.68	13.84
Sex		
Male	20.80	11.85
Female	12.55	15.91
Age		
0 – 5 years		13.54
18-24 years	34.18	
25-29 years	42.13	
30-39 years	27.53	
40-49 years	19.41	
50-64 years	11.42	
65 years and over	8.24	
Race/Ethnicity		
Hispanic	21.84	16.21
Non-Hispanic		
White alone	13.90	12.42
Other	17.00	10.72
Educational Attainment		
N/A (age < 18 years)		
Less than High School	25.35	
HS grad or GED	17.01	
Some college or more	13.54	
Household Tenure		
Owner	14.41	
Renter	19.58	

Table 6-4 shows that 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 (18 to 24 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 2006-2008 American Community Survey (ACS) estimates for Arizona (U.S. Census Bureau, 2010). Tables with population demographics were downloaded from the Census Bureau website. For more detailed group breakdown, the distribution was derived using the 2006-2008 ACS Public Use Micro data files (PUMS). The total population was applied to the distribution of the detailed breakdown from to produce the control totals. Tables in Appendix A show the sample size, sum of weights and control totals for adults and children for the variables used as control totals.

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.

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.

In AHS, random allocation and hot-deck imputation were used to fill in the missing responses. Random selection from the observed distribution was used to impute age and sex, which had 35 missing values and 2 missing values respectively 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.

Household tenure and the adult respondent's educational attainment were used to create two raking dimensions. Household tenure had 52 missing responses and educational attainment had 16 missing cases. 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, poverty, and the percentages of college graduates, home owners, African Americans and Hispanics in the telephone exchange.

A hot-deck imputation was also developed to deal with race/ethnicity within households. There were 65 cases with missing race-ethnicity. The hot-deck imputations were done separately by the completed extended interview structure of the household.

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:

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

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[5]), 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

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

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 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 Westat. 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 or children 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 R

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

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.

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 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.

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.

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.

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.

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

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.

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

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

Appendix A

Weighting Tables

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics

Characteristic	All state					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	2,884	35.1	2,319,205	49.7	2,319,206	49.7
2. Female	5,331	64.9	2,348,157	50.3	2,348,156	50.3
Total						
Tenure						
1. Own	6,298	76.7	3,290,240	70.5	3,214,609	70.5
2. Other	1,917	23.3	1,377,122	29.5	1,346,921	29.5
Total			4,667,362		4,561,530	
Race-ethnicity						
1. Hispanic	1,704	20.7	1,169,834	25.1	1,169,835	25.1
2. Non-Hispanic White only	5,794	70.5	3,003,895	64.4	3,003,894	64.4
3. Other	717	8.7	493,633	10.6	493,633	10.6
Age group						
1. 18-29	559	6.8	1,067,138	22.9	1,046,219	22.4
2. 30-44	1,200	14.6	1,297,811	27.8	1,318,731	28.3
3. 45-54	1,344	16.4	822,883	17.6	738,707	15.8
4. 55-64	1,844	22.5	654,531	14.0	738,707	15.8
5. 65 or older	3,268	39.8	824,999	17.7	824,999	17.7
Education attainment						
1. Less than HS	992	12.1	779,333	16.7	795,970	17.1
2. HS graduate	2,086	25.4	1,279,318	27.4	1,256,673	26.9
3. More than HS	5,137	62.5	2,608,711	55.9	2,614,719	56.0

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics (continued)

Characteristic	1. Mohave, Coconino					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	388	37%	264,011	49%	264,011	49%
2. Female	665	63%	276,051	51%	276,051	51%
Total						
Tenure						
1. Own	792	75%	379,639	70%	479,895*	72%
2. Other	261	25%	160,424	30%	188,003*	28%
Total						
Race-ethnicity						
1. Hispanic	90	9%	55,381	10%	53,642	10%
2. Non-Hispanic White only	841	80%	375,956	70%	375,956	70%
3. Other	122	12%	108,725	20%	110,464	20%
Age group						
1. 18-29	62	6%	114,915	21%	114,012	21%
2. 35-44	142	13%	123,739	23%	124,642	23%
3. 45-54	153	15%	95,858	18%	90,879	17%
4. 55-64	256	24%	85,899	16%	90,879	17%
5. 65 or older	440	42%	119,651	22%	119,651	22%
Education attainment						
1. Less than HS	109	10%	95,963	18%	95,963	18%
2. HS graduate	299	28%	168,971	31%	168,971	31%
3. More than HS	645	61%	275,128	51%	275,128	51%

*Since the two strata have counties that share the same PUMA code (the geographic location code) it is impossible to separate the estimates for the control totals. The values here are the combined control totals of the first two strata (strata 1. Mohave, Coconino, Navajo, Apache, Yavapai and strata 2. Yuma, La Paz).

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics (continued)

Characteristic	2. Yuma, La Paz					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	242	33%	74,903	49%	74,903	49%
2. Female	501	67%	77,093	51%	77,093	51%
Total						
Tenure						
1. Own	573	77%	117,616	77%	479,895*	72%
2. Other	170	23%	34,380	23%	188,003*	28%
Total						
Race-ethnicity						
1. Hispanic	360	48%	65,859	43%	67,598	44%
2. Non-Hispanic White only	336	45%	75,719	50%	75,719	50%
3. Other	47	6%	10,418	7%	8,679	6%
Age group						
1. 18-29	66	9%	30,993	20%	32,097	21%
2. 35-44	144	19%	37,245	25%	36,141	24%
3. 45-54	126	17%	23,475	15%	21,249	14%
4. 55-64	137	18%	19,022	13%	21,249	14%
5. 65 or older	270	36%	41,261	27%	41,261	27%
Education attainment						
1. Less than HS	183	25%	30,678	20%	40,874	27%
2. HS graduate	244	33%	57,479	38%	47,283	31%
3. More than HS	316	43%	63,839	42%	63,839	42%

*Since the two strata have counties that share the same PUMA code (the geographic location code) it is impossible to separate the estimates for the control totals. The values here are the combined control totals of the first two strata (strata 1. Mohave, Coconino, Navajo, Apache, Yavapai and strata 2. Yuma, La Paz).

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics (continued)

Characteristic	3. Graham, Greenlee					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	292	39%	74,616	49%	74,616	49%
2. Female	463	61%	77,246	51%	77,246	51%
Total						
Tenure						
1. Own	577	76%	108,614	72%	105,842	72%
2. Other	178	24%	43,249	28%	42,145	28%
Total						
Race-ethnicity						
1. Hispanic	229	30%	58,394	38%	58,394	38%
2. Non-Hispanic White only	464	61%	80,412	53%	80,412	53%
3. Other	62	8%	13,056	9%	13,056	9%
Age group						
1. 18-29	51	7%	38,265	25%	33,178	22%
2. 35-44	75	10%	30,341	20%	35,427	23%
3. 45-54	109	14%	27,410	18%	25,833	17%
4. 55-64	199	26%	24,256	16%	25,833	17%
5. 65 or older	321	43%	31,591	21%	31,591	21%
Education attainment						
1. Less than HS	134	18%	31,738	21%	31,738	21%
2. HS graduate	202	27%	48,572	32%	42,564	28%
3. More than HS	419	55%	71,552	47%	77,560	51%

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics (continued)

Characteristic	4. Pinal, Gila					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	270	34%	136,210	52%	136,210	52%
2. Female	528	66%	125,550	48%	125,550	48%
Total						
Tenure						
1. Own	653	82%	199,855	76%	184,980	76%
2. Other	145	18%	61,905	24%	57,298	24%
Total						
Race-ethnicity						
1. Hispanic	120	15%	64,071	24%	64,071	24%
2. Non-Hispanic White only	594	74%	168,873	65%	168,873	65%
3. Other	84	11%	28,816	11%	28,816	11%
Age group						
1. 18-29	44	6%	67,096	26%	61,318	23%
2. 35-44	94	12%	65,804	25%	71,583	27%
3. 45-54	123	15%	40,962	16%	38,743	15%
4. 55-64	185	23%	36,524	14%	38,743	15%
5. 65 or older	352	44%	51,374	20%	51,374	20%
Education attainment						
1. Less than HS	95	12%	42,384	16%	48,825	19%
2. HS graduate	233	29%	89,741	34%	83,300	32%
3. More than HS	470	59%	129,635	50%	129,635	50%

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics (continued)

Characteristic	5. Pima					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	758	35%	365,505	48%	365,505	48%
2. Female	1,385	65%	391,874	52%	391,874	52%
Total						
Tenure						
1. Own	1,671	78%	524,114	69%	510,212	69%
2. Other	472	22%	233,265	31%	227,078	31%
Total						
Race-ethnicity						
1. Hispanic	386	18%	211,917	28%	211,917	28%
2. Non-Hispanic White only	1,582	74%	477,081	63%	477,081	63%
3. Other	175	8%	68,381	9%	68,381	9%
Age group						
1. 18-29	114	5%	172,054	23%	168,661	22%
2. 35-44	278	13%	190,473	25%	193,867	26%
3. 45-54	356	17%	136,082	18%	123,429	16%
4. 55-64	497	23%	110,775	15%	123,429	16%
5. 65 or older	898	42%	147,995	20%	147,995	20%
Education attainment						
1. Less than HS	212	10%	104,493	14%	104,493	14%
2. HS graduate	453	21%	194,981	26%	194,981	26%
3. More than HS	1,478	69%	457,905	60%	457,905	60%

Table A-1. Sample sizes, sums of weights, and control totals for adult interviews by sample characteristics (continued)

Characteristic	6. Maricopa					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	934	34%	1,403,960	50%	1,403,961	50%
2. Female	1,789	66%	1,400,343	50%	1,400,342	50%
Total						
Tenure						
1. Own	2,032	75%	1,960,403	70%	1,933,680	70%
2. Other	691	25%	843,900	30%	832,397	30%
Total						
Race-ethnicity						
1. Hispanic	519	19%	714,212	25%	714,213	25%
2. Non-Hispanic White only	1,977	73%	1,825,853	65%	1,825,853	65%
3. Other	227	8%	264,237	9%	264,237	9%
Age group						
1. 18-29	222	8%	643,815	23%	636,954	23%
2. 35-44	467	17%	850,210	30%	857,072	31%
3. 45-54	477	18%	499,096	18%	438,576	16%
4. 55-64	570	21%	378,055	13%	438,576	16%
5. 65 or older	987	36%	433,127	15%	433,127	15%
Education attainment						
1. Less than HS	259	10%	474,077	17%	474,077	17%
2. HS graduate	655	24%	719,574	26%	719,574	26%
3. More than HS	1,809	66%	1,610,652	57%	1,610,652	57%

Table A-2. Sample sizes, sums of weights, and control totals for child interviews by sample characteristics

Characteristic	All state					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	1,137	52.9	301,329	51.1	301,329	51.1
2. Female	1,011	47.1	288,902	48.9	288,902	48.9
Total						
Tenure						
1. Own	1,466	68.2	379,906	64.4	379,906	64.4
2. Other	682	31.8	210,325	35.6	210,325	35.6
Total						
Race-ethnicity						
1. Hispanic	829	38.6	266,803	45.2	266,803	45.2
2. Non-Hispanic White only	1,111	51.7	238,984	40.5	238,984	40.5
3. Other	208	9.7	84,444	14.3	84,444	14.3
Age group						
1. 0-1	698	32.5	193,708	32.8	193,708	32.8
2. 2-3	772	35.9	202,929	34.4	202,929	34.4
3. 4-5	678	31.6	193,594	32.8	193,594	32.8

Table A-2. Sample sizes, sums of weights, and control totals for child interviews by sample characteristics (continued)

Characteristic	1. Mohave, Coconino					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	170	46.7	29,428	51.3	29,428	51.3
2. Female	194	53.3	27,941	48.7	27,941	48.7
Total						
Tenure						
1. Own	251	69.0	45,249	78.9	62,043*	79.7
2. Other	113	31.0	12,120	21.1	15,769*	20.3
Total						
Race-ethnicity						
1. Hispanic	74	20.3	12,086	21.1	12,086	21.1
2. Non-Hispanic White only	240	65.9	26,829	46.8	26,829	46.8
3. Other	50	13.7	18,454	32.2	18,454	32.2
Age group						
1. 0-1	122	33.5	18,365	32.0	24,325*	31.3
2. 2-3	131	36.0	19,173	33.4	26,166*	33.6
3. 4-5	111	30.5	19,832	34.6	27,321*	35.1

*Since the two strata have counties that share the same PUMA code (the geographic location code) it is impossible to separate the estimates for the control totals. The values here are the combined control totals of the first two strata (strata 1. Mohave, Coconino, Navajo, Apache, Yavapai and strata 2. Yuma, La Paz).

Table A-2. Sample sizes, sums of weights, and control totals for child interviews by sample characteristics (continued)

Characteristic	2. Yuma, La Paz					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	202	49.6	10,364	50.7	10,364	50.7
2. Female	205	50.4	10,079	49.3	10,079	49.3
Total						
Tenure						
1. Own	228	56.0	16,795	82.2	62,043*	79.7
2. Other	179	44.0	3,649	17.8	15,769*	20.3
Total						
Race-ethnicity						
1. Hispanic	300	73.7	15,054	73.6	15,054	73.6
2. Non-Hispanic White only	90	22.1	4,258	20.8	4,258	20.8
3. Other	17	4.2	1,131	5.5	1,131	5.5
Age group						
1. 0-1	131	32.2	5,961	29.2	24,325*	31.3
2. 2-3	133	32.7	6,993	34.2	26,166*	33.6
3. 4-5	143	35.1	7,489	36.6	27,321*	35.1

*Since the two strata have counties that share the same PUMA code (the geographic location code) it is impossible to separate the estimates for the control totals. The values here are the combined control totals of the first two strata (strata 1. Mohave, Coconino, Navajo, Apache, Yavapai and strata 2. Yuma, La Paz).

Table A-2. Sample sizes, sums of weights, and control totals for child interviews by sample characteristics (continued)

Characteristic	3 & 4. Graham, Greenlee, Pinal, Gila					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	280	52.1	24,909	51.3	24,909	51.3
2. Female	257	47.9	23,606	48.7	23,606	48.7
Total						
Tenure						
1. Own	389	72.4	31,533	65.0	31,533	65.0
2. Other	148	27.6	16,983	35.0	16,983	35.0
Total						
Race-ethnicity						
1. Hispanic	168	31.3	20,371	42.0	20,371	42.0
2. Non-Hispanic White only	314	58.5	20,826	42.9	20,826	42.9
3. Other	55	10.2	7,318	15.1	7,318	15.1
Age group						
1. 0-1	179	33.3	15,296	31.5	15,297	31.5
2. 2-3	213	39.7	16,821	34.7	16,821	34.7
3. 4-5	145	27.0	16,397	33.8	16,397	33.8

Table A-2. Sample sizes, sums of weights, and control totals for child interviews by sample characteristics (continued)

Characteristic	5. Pima					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	251	60.9	41,208	50.8	41,208	50.8
2. Female	161	39.1	39,868	49.2	39,868	49.2
Total						
Tenure						
1. Own	315	76.5	40,370	49.8	40,370	49.8
2. Other	97	23.5	40,706	50.2	40,706	50.2
Total						
Race-ethnicity						
1. Hispanic	124	30.1	40,844	50.4	40,844	50.4
2. Non-Hispanic White only	246	59.7	30,145	37.2	30,145	37.2
3. Other	42	10.2	10,087	12.4	10,087	12.4
Age group						
1. 0-1	131	31.8	26,374	32.5	26,374	32.5
2. 2-3	134	32.5	28,796	35.5	28,796	35.5
3. 4-5	147	35.7	25,907	32.0	25,907	32.0

Table A-2. Sample sizes, sums of weights, and control totals for child interviews by sample characteristics (continued)

Characteristic	6. Maricopa					
	Sample size		Sum of weights		Control total	
	N	%	Sum	%	Total	%
Gender						
1. Male	234	54.7	195,420	51.0	195,420	51.0
2. Female	194	45.3	187,407	49.0	187,407	49.0
Total						
Tenure						
1. Own	283	66.1	245,960	64.2	245,960	64.2
2. Other	145	33.9	136,867	35.8	136,867	35.8
Total						
Race-ethnicity						
1. Hispanic	163	38.1	178,448	46.6	178,448	46.6
2. Non-Hispanic White only	221	51.6	156,926	41.0	156,926	41.0
3. Other	44	10.3	47,453	12.4	47,453	12.4
Age group						
1. 0-1	135	31.5	127,712	33.4	127,712	33.4
2. 2-3	161	37.6	131,146	34.3	131,146	34.3
3. 4-5	132	30.8	123,969	32.4	123,969	32.4