



June 9, 2014

CHIS 2011-2012 Methodology Report Series

Report 1

Sample Design

CALIFORNIA HEALTH INTERVIEW SURVEY

CHIS 2011-2012 METHODOLOGY SERIES

REPORT 1

SAMPLE DESIGN

VERSION DATE: JUNE 9, 2014

This report was prepared for the California Health Interview Survey by Ismael Flores Cervantes and J. Michael Brick of Westat.



www.chis.ucla.edu

This report provides analysts with information about the sampling methods used for CHIS 2011-2012, including both the household and person (within household) sampling. This report also provides a discussion on achieved sample size and how it compares to the planned sample size.

Suggested citation:

California Health Interview Survey. *CHIS 2011-2012 Methodology Series: Report 1 - Sample Design*. Los Angeles, CA: UCLA Center for Health Policy Research, 2014.

Copyright © 2014 by the Regents of the University of California.

The California Health Interview Survey is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Public Health, and the Department of Health Care Services. Funding for CHIS 2011-2012 came from multiple sources: the Blue Shield of California Foundation, the California Department of Health Care Services, the California Department of Mental Health, the California Department of Public Health, the California Endowment, First 5 California, the California Office of the Patient Advocate, Kaiser Permanente, the NIH National Cancer Institute, NIH Office of Research on Women's Health, NIH Office of Behavioral and Social Sciences Research, San Diego County Health and Human Services Agency, United American Indian Involvement, Inc./CA Indian Health Service, Archstone Foundation, Susan B. Komen for the Cure, Centers for Disease Control and Prevention, NCI American Cancer Society, NIH RAND, and the California Wellness Foundation.

PREFACE

Sample Design is the first in a series of methodological reports describing the 2011-2012 California Health Interview Survey (CHIS 2011-2012). The other reports are listed below.

CHIS is a collaborative project of the University of California, Los Angeles (UCLA) Center for Health Policy Research, the California Department of Public Health, and the Department of Health Care Services. Westat was responsible for data collection and the preparation of five methodological reports from the 2011-2012 survey. The survey examines public health and health care access issues in California. The telephone survey is the largest state health survey ever undertaken in the United States.

Methodological Report Series for CHIS 2011-2012

The methodological reports for CHIS 2011-2012 are as follows:

- Report 1: Sample Design;
- Report 2: Data Collection Methods;
- Report 3: Data Processing Procedures;
- Report 4: Response Rates; and
- Report 5: Weighting and Variance Estimation.

The reports are interrelated and contain many references to each other. For ease of presentation, the references are simply labeled by the report numbers given above. After the Preface, each report includes an “Overview” chapter ([Chapter 1](#)) that is nearly identical across reports, followed by detailed technical documentation on the specific topic of the report.

Report 1: Sample Design (this report) describes the procedures used to design and select the sample from CHIS 2011-2012. An appropriate sample design is a feature of a successful survey, and CHIS 2011-2012 presented many issues that had to be addressed at the design stage. This report explains why the design features of CHIS were selected and presents the alternatives that were considered and provides analysts information about the sampling methods used for both the household and person (within household) sampling. In general terms, once a household was sampled, an adult within that household was sampled. If there were children and/or adolescents in the household, one child and/or one adolescent was eligible for sampling. This report also provides a discussion on achieved sample size and how it compares to the planned sample size.

The purposes of this report are:

- To serve as a reference for researchers using CHIS 2011-2012 data;
- To document data collection procedures so that future iterations of CHIS, or other similar surveys, can replicate those procedures if desired;
- To describe lessons learned from the data collection experience and make recommendations for improving future surveys; and
- To evaluate the level of effort required for the various kinds of data collection undertaken.

For further methodological details not covered in this report, refer to the other methodological reports in the series at <http://healthpolicy.ucla.edu/chis/design/Pages/methodology.aspx>. General information on CHIS data can be found on the California Health Interview Survey Web site at <http://www.chis.ucla.edu> or by contacting CHIS at CHIS@ucla.edu.

TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
PREFACE.....	i
1. CHIS 2011-2012 SAMPLE DESIGN AND METHODOLOGY SUMMARY	1-1
1.1 Overview	1-1
1.2 Switch to a Continuous Survey.....	1-2
1.3 Sample Design Objectives.....	1-2
1.4 Data Collection	1-5
1.5 Response Rates	1-9
1.6 Weighting the Sample	1-11
1.7 Imputation Methods.....	1-12
1.8 Methodology Report Series	1-14
2. TELEPHONE SAMPLING METHODS	2-1
2.1 List-Assisted Random Digit Dial Sampling of Landlines	2-2
2.2 Households without Landline Telephones.....	2-3
2.3 Increasing the Efficiency of Data Collection.....	2-4
2.4 Supplemental Sampling.....	2-5
3. SAMPLING HOUSEHOLDS	3-1
3.1 Population of Interest.....	3-1
3.2 Sample Design.....	3-1
3.2.1 Landline Sample	3-5
3.2.2 Cell Phone Sample.....	3-9
3.2.3 Supplemental Geographic Sample.....	3-14
3.2.4 Supplemental Surname List Samples	3-15
3.2.5 American Indian and Alaska Native List Sample.....	3-16
3.3 Sample Selection	3-17
4. WITHIN-HOUSEHOLD SAMPLING	4-1
4.1 Sampling Alternatives	4-1
4.2 Child-first Procedure	4-2
4.3 Adult Sampling.....	4-4
4.4 Child Sampling.....	4-6
4.5 Adolescent Sampling.....	4-7
5. ACHIEVED SAMPLE SIZES	5-1

TABLE OF CONTENTS (CONTINUED)

<u>Chapter</u>	<u>Page</u>
REFERENCES	R-1
APPENDIX A.....	A-1

List of Tables

<u>Table</u>	<u>Page</u>
1-1. California County and County Group Strata Used in the CHIS 2011-2012 Sample Design.....	1-4
1-2. Number of completed CHIS 2011-2012 interviews by type of sample and instrument.....	1-6
1-3. CHIS 2011-2012 survey topic areas by instrument.....	1-7
2-1. CSS result codes and their distribution in the CHIS 2011-2012 sample.....	2-4
3-1. Initial targets for completed adult interviews by county (excluding the AIAN supplemental sample).....	3-2
3-2. Final targets for completed adult interviews by county (excluding the AIAN supplemental sample).....	3-4
3-3. Final targets for completed adult interviews from the landline sample by county (excluding the AIAN supplemental sample).....	3-6
3-4. Definition of sampling substratum, number of exchanges, and total number of households for Los Angeles County, San Diego County, Orange County, and Santa Clara County.....	3-9
3-5. Final targets for completed adult interviews from the cell phone sample by county.....	3-10
3-6. Definition of cell phone sampling strata for complete area codes.....	3-12
3-7. Definition of cell phone sampling strata based on area code and county combinations.....	3-12
3-8. Number of cell telephone numbers drawn by sampling stratum.....	3-13
3-9. Targeted number of completed adult interviews for the Korean and Vietnamese samples.....	3-15
3-10. Surname frame sizes.....	3-16
3-11. Frame size and targeted number of completed adult interviews for the American Indian/Alaska Native supplemental list sample.....	3-17
3-12. Number of telephone numbers drawn by sample type.....	3-18
3-13. Release groups of telephone numbers by sample type.....	3-18

TABLE OF CONTENTS (CONTINUED)

List of Tables (continued)

<u>Table</u>		<u>Page</u>
4-1.	Effect of the child-first procedure on completed child and adolescent interviews in the landline sample	4-4
4-2.	Distribution of households with children by type of child sampling.....	4-7
5-1.	Number of completed interviews by type of sample	5-1
5-2.	Number of completed adult interviews by self-reported stratum	5-2
5-3.	Number of completed child interviews and by self-reported stratum.....	5-3
5-4.	Number of completed adolescent interviews by self-reported stratum	5-5
5-5.	Number of completed adult interviews by ethnicity and sample type.....	5-6
A-1.	Stratum definitions for CHIS 2001, 2003, 2005, 2007, 2009, and 2011-2012.....	A-1
A-2.	Number of telephone numbers and addresses drawn by sample frame and sampling stratum.....	A-3
A-3.	Number of adult completed interviews by sample type and self-reported stratum ...	A-5
A-4.	Number of child completed interviews by self-reported stratum	A-7
A-5.	Number of adolescent completed interviews by self-reported stratum	A-9

1. CHIS 2011-2012 SAMPLE DESIGN AND METHODOLOGY SUMMARY

1.1 Overview

The California Health Interview Survey (CHIS) is a population-based telephone survey of California conducted every other year since 2001 and continually beginning in 2011. CHIS is the largest state health survey conducted and one of the largest health surveys in the nation. CHIS is conducted by the UCLA Center for Health Policy Research (UCLA-CHPR) in collaboration with the California Department of Public Health, the Department of Health Care Services, First 5 California, The California Endowment, the National Cancer Institute, and Kaiser Permanente. CHIS collects extensive information for all age groups on health status, health conditions, health-related behaviors, health insurance coverage, access to health care services, and other health and health related issues.

The sample is designed to meet and optimize two objectives:

- 1) Provide estimates for large- and medium-sized counties in the state, and for groups of the smallest counties (based on population size), and
- 2) Provide statewide estimates for California's overall population, its major racial and ethnic groups, as well as several Asian and Latino ethnic subgroups.

The CHIS sample is representative of California's non-institutionalized population living in households. CHIS data and results are used extensively by federal and State agencies, local public health agencies and organizations, advocacy and community organizations, other local agencies, hospitals, community clinics, health plans, foundations, and researchers. These data are used for analyses and publications to assess public health and health care needs, to develop and advocate policies to meet those needs, and to plan and budget health care coverage and services. Many researchers throughout California and the nation use CHIS data files to further their understanding of a wide range of health-related issues (visit the CHIS Research Clearinghouse at <http://healthpolicy.ucla.edu/chis/research/Pages/default.aspx> for many examples of these studies).

This series of reports describes the methods used in collecting data for CHIS 2011-2012, the sixth CHIS data collection cycle, which was conducted between June 2011 and January 2013. The previous CHIS cycles (2001, 2003, 2005, 2007, and 2009) are described in similar series, available at <http://healthpolicy.ucla.edu/chis/design/Pages/methodology.aspx>.

1.2 Switch to a Continuous Survey

From the first CHIS cycle in 2001 through 2009, CHIS data collection was biennial, with data collected during a 7-9 month period every other year. Beginning in 2011, CHIS data are collected continually over each 2-year cycle. This change was driven by several factors including the ability to track and release information about health in California on a more frequent and timely basis and to eliminate potential seasonality in the biennial data.

The CHIS 2011-2012 data included in these files were collected between June 2011 and January 2013. Approximately half of the interviews were conducted during the 2011 calendar year and half during the 2012 calendar year. As in previous CHIS cycles, weights are included with the data files and are based on the State of California's Department of Finance population estimates and projections, adjusted to remove the population living in group quarters (such as nursing homes, prisons, etc. and not eligible to participate in CHIS). When the weights are applied to the data, the results represent California's residential population during that one year period for the age group corresponding to the data file in use (adult, adolescent, or child).

See what else is new in the 2011-2012 CHIS sampling and data collection here: <http://healthpolicy.ucla.edu/chis/design/Documents/whats-new-chis-2011-2012.pdf>

In order to provide CHIS data users with more complete and up-to-date information to facilitate analyses of CHIS data, additional information on how to use the CHIS sampling weights, including sample code, is available at: <http://healthpolicy.ucla.edu/chis/analyze/Pages/sample-code.aspx>

Additional documentation on constructing the CHIS sampling weights is available in CHIS 2011-2012 Methods Report #5—Weighting and Variance Estimation, available at: <http://healthpolicy.ucla.edu/chis/design/Pages/methodology.aspx>. Other helpful information for understanding the CHIS sample design and data collection processing can be found in the four other methodology reports for each CHIS cycle year, described in the Preface above.

1.3 Sample Design Objectives

The CHIS 2011-2012 sample was designed to meet two sampling objectives discussed above: (1) provide estimates for adults in most counties and groups of counties with small populations; and

(2) provide estimates for California’s overall population, major racial and ethnic groups, and for several smaller ethnic subgroups.

To achieve these objectives, CHIS employed a dual-frame, multi-stage sample design. The random-digit-dial (RDD) sample included telephone numbers assigned to both landline and cellular service. The random-digit-dial (RDD) sample was approximately 80% landline and 20% cellular phone numbers. For the landline RDD sample, the 58 counties in the state were grouped into 44 geographic sampling strata, and 14 sub-strata were created within two of the largest metropolitan areas in the state (Los Angeles and San Diego). The Los Angeles County stratum included 8 sub-strata for Service Planning Areas, and the San Diego County stratum included 6 sub-strata for Health Service Regions. Most of the strata (39 of 44) are made up of a single county with no sub-strata (counties 3-41 in [Table 1-1](#)), with three multi-county strata comprised of the 17 remaining counties (see [Table 1-1](#)). A sufficient number of adult interviews were allocated to each stratum and sub-stratum to support the first sample design objective—to provide health estimates for adults at the local level. The same geographic stratification of the state has been used since CHIS 2005. In the first two CHIS cycles (2001 and 2003) there were 47 total sampling strata, including 33 individual counties and one county with sub-strata (Los Angeles).

Within each geographic stratum, residential telephone numbers were selected, and within each household, one adult respondent (age 18 and over) was randomly selected. In those households with adolescents (ages 12-17) and/or children (under age 12), one adolescent and one child were randomly selected; the adolescent was interviewed directly, and the adult most knowledgeable about the child’s health completed the child interview.

The RDD CHIS sample is of sufficient size to accomplish the second objective (produce estimates for the state’s major racial/ethnic groups, as well as many ethnic subgroups). To increase the precision of estimates for Koreans and Vietnamese, areas with relatively high concentrations of these groups were sampled at higher rates. These geographically targeted oversamples were supplemented by telephone numbers associated with group-specific surnames drawn from listed telephone directories to further increase the sample size for Koreans and Vietnamese.

Table 1-1. California county and county group strata used in the CHIS 2011-2012 sample design

1. Los Angeles	7. Alameda	27. Shasta
1.1 Antelope Valley	8. Sacramento	28. Yolo
1.2 San Fernando Valley	9. Contra Costa	29. El Dorado
1.3 San Gabriel Valley	10. Fresno	30. Imperial
1.4 Metro	11. San Francisco	31. Napa
1.5 West	12. Ventura	32. Kings
1.6 South	13. San Mateo	33. Madera
1.7 East	14. Kern	34. Monterey
1.8 South Bay	15. San Joaquin	35. Humboldt
2. San Diego	16. Sonoma	36. Nevada
2.1 N. Coastal	17. Stanislaus	37. Mendocino
2.2 N. Central	18. Santa Barbara	38. Sutter
2.3 Central	19. Solano	39. Yuba
2.4 South	20. Tulare	40. Lake
2.5 East	21. Santa Cruz	41. San Benito
2.6 N. Inland	22. Marin	42. Colusa, Glen, Tehama
3. Orange	23. San Luis Obispo	43. Plumas, Sierra, Siskiyou, Lassen, Modoc, Trinity, Del Norte
4. Santa Clara	24. Placer	44. Mariposa, Mono, Tuolumne, Alpine, Amador, Calaveras, Inyo
5. San Bernardino	25. Merced	
6. Riverside	26. Butte	

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

To help compensate for the increasing number of households without landline telephone service, a separate RDD sample was drawn of telephone numbers assigned to cellular service. In CHIS 2011-2012, the goal was to complete approximately 8,000 interviews (20% of all RDD interviews statewide) with adults from the cell phone sample. Telephone numbers assigned to cellular service cannot be geographically stratified at the county level with sufficient precision, so the cell RDD sample was geographically stratified into 28 strata using 7 CHIS regions and telephone area codes. If a sampled cell number was shared by two or more adult members of a household, one household member was selected for the adult interview. Otherwise, the adult owner of the sampled number was selected. Cell numbers used exclusively by children under 18 were considered ineligible. About 550 teen interviews and 1,500 child interviews were completed from the cell phone sample in CHIS 2011-2012.

The CHIS 2011-2012 and 2009 cell phone sampling method differed from that used in CHIS 2007 in two significant ways. First, in CHIS 2011-2012, all cell phone sample numbers used for non-business purposes by adults living in California were eligible for the extended interview, while in 2007 only cell numbers belonging to adults in cell-only households were eligible. Thus, adults in households with landlines who had their own cell phones or shared one with another adult household member could

have been selected through either the cell or landline sample. The second change to the cell phone sample was the inclusion of child and adolescent extended interviews.

Unlike both CHIS 2007 and CHIS 2009, where the cell phone sample quotas were treated separately from the landline sample, the CHIS 2011-2012 cell sample respondents were included in the overall and county specific target sample sizes. Twenty-eight cell phone sampling strata were created using CHIS 2007 and 2009 cell phone respondents' data and their pre-assigned FIPS county code, supplied by the sampling vendor. The statewide target of 8,000 adult cell phone interviews was also supplemented with an oversample to yield approximately 1,150 adult cell phone interviews. The oversample focused on six counties; Los Angeles, Orange, Santa Clara, Alameda, San Francisco, and San Mateo.

Finally, the CHIS 2011-2012 sample included an American Indian/Alaska Native (AIAN) oversample. This oversample was sponsored by Urban American Indian Involvement, Inc., and California Indian Health Services. The purpose of this oversample was to increase the number of AIAN participants and improve the statistical stability and precision of estimates for this group. The oversample was conducted using a list provided by Indian Health Services.

1.4 Data Collection

To capture the rich diversity of the California population, interviews were conducted in five languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, and Korean. These languages were chosen based on analysis of 2000 Census data to identify the languages that would cover the largest number of Californians in the CHIS sample that either did not speak English or did not speak English well enough to otherwise participate.

Westat, a private firm that specializes in statistical research and large-scale sample surveys, conducted CHIS 2011-2012 data collection under contract with the UCLA Center for Health Policy Research. For all samples, Westat staff interviewed one randomly selected adult in each sampled household, and sampled one adolescent and one child if they were present in the household and the sampled adult was the parent or legal guardian. Thus, up to three interviews could have been completed in each household. In landline sample households with children where the sampled adult was not the screener respondent, children and adolescents could be sampled as part of the screening interview, and the extended child (and adolescent) interviews could be completed before the adult interview. This "child-first" procedure was new for CHIS 2005 and has been continued in subsequent CHIS cycles; this

procedure substantially increases the yield of child interviews. While numerous subsequent attempts were made to complete the adult interview for child-first cases, there are completed child and/or adolescent interviews in households for which an adult interview was not completed. [Table 1-2](#) shows the number of completed adult, child, and adolescent interviews in CHIS 2011-2012 by the type of sample (landline RDD, surname list, cell RDD, and American Indian/Alaska Native list).

Table 1-2. Number of completed CHIS 2011-2012 interviews by type of sample and instrument

Type of sample	Adult	Child	Adolescent
Total all samples	42,935 ¹	7,334	2,799
Landline RDD	32,692	5,600	2,164
Surname list	825	161	57
Cell RDD	9,151	1,523	557
American Indian/Alaska Native list	267	50	21

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

Interviews in all languages were administered using Westat’s computer-assisted telephone interviewing (CATI) system. The average adult interview took about 35 minutes to complete. The average child and adolescent interviews took about 15 minutes and 23 minutes, respectively. For “child-first” interviews, additional household information asked as part of the child interview averaged about 9 minutes. Interviews in non-English languages generally took longer to complete. More than 14 percent of the adult interviews were completed in a language other than English, as were about 27 percent of all child (parent proxy) interviews and 7 percent of all adolescent interviews.

[Table 1-3](#) shows the major topic areas for each of the three survey instruments (adult, child, and adolescent).

¹ Numbers in this table represent the data publically released and available through our Data Access Center. Total sample sizes may differ for specific calculations within the five methodology reports, or for specific analyses based on CHIS data.

Table 1-3. CHIS 2011-2012 survey topic areas by instrument

Health status	Adult	Teen	Child
General health status	✓	✓	✓
Days missed from school due to health problems	✓	✓	✓
Health-related quality of life (HRQOL)		✓	
Health conditions	Adult	Teen	Child
Asthma	✓	✓	✓
Diabetes, gestational diabetes, pre- /borderline diabetes	✓		
Heart disease, high blood pressure, stroke	✓		
Arthritis, physical disability	✓		
Epilepsy		✓	
Physical, behavioral, and/or mental conditions			✓
Mental health	Adult	Teen	Child
Mental health status	✓	✓	
Perceived need, access and utilization of mental health services	✓	✓	
Functional impairment, stigma	✓		
Suicide ideation and attempts	✓		
Health behaviors	Adult	Teen	Child
Dietary intake, fast food	✓	✓	✓
Physical activity and exercise, commute from school to home		✓	✓
Walking for transportation and leisure	✓		
Doctor discussed nutrition/physical activity		✓	✓
Flu Shot	✓		✓
Alcohol and cigarette use	✓	✓	
Illegal drug use		✓	
Sexual behavior	✓	✓	
HIV/STI testing		✓	
Elderly falls	✓		
Women's health	Adult	Teen	Child
Mammography screening	✓		
Pregnancy	✓	✓	
Dental health	Adult	Teen	Child
Last dental visit, main reason haven't visited dentist		✓	✓
Neighborhood and housing	Adult	Teen	Child
Safety, social cohesion	✓	✓	✓
Homeownership, length of time at current residence	✓		
Park use		✓	✓
Civic engagement	✓	✓	
Access to and use of health care	Adult	Teen	Child
Usual source of care, visits to medical doctor	✓	✓	✓
Emergency room visits	✓	✓	✓
Delays in getting care (prescriptions and medical care)	✓	✓	✓
Medical home, timely appointments, hospitalizations	✓	✓	✓
Communication problems with doctor	✓		✓
Internet use for health information	✓		✓

Table 1-3. CHIS 2011-2012 survey topic areas by instrument (continued)

Food environment	Adult	Teen	Child
Access to fresh and affordable foods	✓		
Where teen/child eats breakfast/lunch, fast food at school		✓	✓
Availability of food in household over past 12 months	✓		
Health insurance	Adult	Teen	Child
Current insurance coverage, spouse's coverage, who pays for coverage	✓	✓	✓
Health plan enrollment, characteristics and plan assessment	✓	✓	✓
Whether employer offers coverage, respondent/spouse eligibility	✓		
Coverage over past 12 months, reasons for lack of insurance	✓	✓	✓
Difficulty finding private health insurance	✓		
High deductible health plans	✓	✓	✓
Partial scope Medi-Cal	✓		
Public program eligibility	Adult	Teen	Child
Household poverty level	✓		
Program participation (CalWORKs, Food Stamps, SSI, SSDI, WIC, TANF)	✓	✓	✓
Assets, alimony/child support, social security/pension	✓		
Medi-Cal and Healthy Families eligibility	✓	✓	✓
Reason for Medi-Cal non-participation among potential beneficiaries	✓	✓	✓
Bullying and interpersonal violence	Adult	Teen	Child
Bullying, personal safety, interpersonal violence		✓	
Parental involvement/adult supervision	Adult	Teen	Child
Adult presence after school, role models, resiliency		✓	
Parental involvement			✓
Child care and school attendance	Adult	Teen	Child
Current child care arrangements			✓
Paid child care	✓		
First 5 California: Kit for New Parents			✓
Preschool/school attendance, name of school		✓	✓
Preschool quality			✓
School instability		✓	
Employment	Adult	Teen	Child
Employment status, spouse's employment status	✓		
Hours worked at all jobs	✓		
Income	Adult	Teen	Child
Respondent's and spouse's earnings last month before taxes	✓		
Household income, number of persons supported by household income	✓		

Table 1-3. CHIS 2011-2012 survey topic areas by instrument (continued)

Respondent characteristics	Adult	Teen	Child
Race and ethnicity, age, gender, height, weight	✓	✓	✓
Veteran status	✓		
Marital status, registered domestic partner status (same-sex couples)	✓		
Sexual orientation	✓		
Language spoken with peers, language of TV, radio, newspaper used	✓		
Education, English language proficiency	✓		
Citizenship, immigration status, country of birth, length of time in U.S., languages spoken at home	✓	✓	✓

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

1.5 Response Rates

The overall response rate for CHIS 2011-2012 is a composite of the screener completion rate (i.e., success in introducing the survey to a household and randomly selecting an adult to be interviewed) and the extended interview completion rate (i.e., success in getting one or more selected persons to complete the extended interview). To maximize the response rate, especially at the screener stage, an advance letter in five languages was mailed to all landline sampled telephone numbers for which an address could be obtained from reverse directory services. An advance letter was mailed for 48.3 percent of the landline RDD sample telephone numbers not identified by the sample vendor as business or nonworking numbers, 81.1 percent of surname list sample numbers, and 94.3 percent of the AIAN list with landline numbers after removing nonworking and business numbers. Addresses were not available for the cell sample. As in all CHIS cycles since CHIS 2005, a \$2 bill was included with the CHIS 2011-2012 advance letter to encourage cooperation.

The CHIS 2011-2012 screener response rate for the landline sample was 31.6 percent, and was higher for households that were sent the advance letter. For the cell phone sample, the screener response rate was 33.0 percent in all households. The extended interview response rate for the landline sample varied across the adult (47.4 percent), child (73.2 percent) and adolescent (42.7 percent) interviews. The adolescent rate includes getting permission from a parent or guardian. The adult interview response rate for the cell sample was 53.8 percent, the child rate was 73.4 percent, and the adolescent rate 42.6 percent. Multiplying the screener and extended rates gives an overall response rate for each type of interview. The percentage of households completing one or more of the extended interviews (adult, child, and/or adolescent) is a useful summary of the overall performance of the landline sample. For CHIS 2011-2012, the landline/list sample household response rate was 17.0 percent (the product of the screener response

rate and the extended interview response rate at the household level of 53.9 percent). The cell sample household response rate was 18.3 percent, incorporating a household-level extended interview response rate of 55.5 percent. All of the household and person level response rates vary by sampling stratum. For more information about the CHIS 2011-2012 response rates please see *CHIS 2011-2012 Methodology Series: Report 4 – Response Rates*.

Historically, the CHIS response rates are comparable to response rates of other scientific telephone surveys in California, such as the California Behavioral Risk Factor Surveillance System (BRFSS) Survey. However, comparing the CHIS and BRFSS response rates requires recomputing the CHIS response rates so they match the BRFSS response rate calculation methods. The 2011 California BRFSS landline response rate is 37.4 percent, the cell phone response rate is 20.4 percent, and the combined landline and cell phone rate is 35.4 percent.² In contrast, the CHIS 2011-2012 landline response rate is 39.5, cell phone response rate is 32.1 percent, and the combined landline and cell phone response rate is 35.1 percent, all these computed using the BRFSS methodology. California as a whole and the state's urban areas in particular are among the most difficult parts of the nation in which to conduct telephone interviews. The 2011 BRFSS, for example, shows the refusal rate for California (31.4%) is the highest in the nation and twice the national median (16.0%). Survey response rates tend to be lower in California than nationally, and over the past decade response rates have been declining both nationally and in California.

Further information about CHIS data quality and nonresponse bias is available at <http://healthpolicy.ucla.edu/chis/design/Pages/data-quality.aspx>.

After all follow-up attempts to complete the full questionnaire were exhausted, adults who completed at least approximately 80 percent of the questionnaire (i.e., through Section K which covers employment, income, poverty status, and food security), were counted as “complete.” At least some responses in the employment and income series, or public program eligibility and food insecurity series were missing from those cases that did not complete the entire interview. They were imputed to enhance the analytic utility of the data.

Proxy interviews were conducted for frail and ill persons over the age of 65 who were unable to complete the extended adult interview in order to avoid biases for health estimates of elderly persons that might otherwise result. Eligible selected persons were re-contacted and offered a proxy option. For 283

² As reported in the Behavioral Risk Factor Surveillance System 2011 Summary Data Quality Report (Version #5--Revised: 2/04/2013, available online at http://www.cdc.gov/brfss/pdf/2011_Summary_Data_Quality_Report.pdf.)

elderly adults, a proxy interview was completed by either a spouse/partner or adult child. A reduced questionnaire, with questions identified as appropriate for a proxy respondent, was administered.

1.6 Weighting the Sample

To produce population estimates from CHIS 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. The weighting procedures used for CHIS 2011-2012 accomplish the following objectives:

- Compensate for differential probabilities of selection for households and persons;
- Reduce biases occurring because non-respondents may have different characteristics than respondents;
- Adjust, to the extent possible, for under-coverage in the sampling frames 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 an iterative proportional fitting method or raking, as it is commonly called, so that the CHIS estimates are consistent with the marginal population control totals. This iterative procedure forces the CHIS weights to sum to known population control totals from an independent data source (see below). The procedure requires iteration to make sure all the control totals, or raking dimensions, are simultaneously satisfied within a pre-specified tolerance.

Population control totals of the number of persons by age, race, and sex at the stratum level for CHIS 2011-2012 were created primarily from the California Department of Finance’s (DOF) 2012 Population Estimates and 2012 Population Projections. The raking procedure used 12 raking dimensions, which are combinations of demographic variables (age, sex, race, and ethnicity), geographic variables (county, Service Planning Area in Los Angeles County, and Health Region in San Diego County), household composition (presence of children and adolescents in the household), and socio-economic

variables (home ownership and education). The socio-economic variables are included to reduce biases associated with excluding households without landline telephones from the sample frame. One limitation of using Department of Finance (DOF) data is that it includes about 2.4 percent of the population of California who live in “group quarters” (i.e., persons living with nine or more unrelated persons and includes, for example nursing homes, prisons, dormitories, etc.). These persons were excluded from the CHIS target population and, as a result, the number of persons living in group quarters was estimated and removed from the Department of Finance control totals prior to raking.

DOF control totals used to create the CHIS 2011-2012 weights are based on 2010 Census counts, while those in previous CHIS cycles were based on Census 2000 counts (with adjustments made by the Department of Finance). Please pay close attention when comparing estimates using CHIS 2011-2012 data with estimates using data from previous CHIS cycles. The most accurate California population figures are available when the US population count is conducted (every 10 years). Population-based surveys like CHIS must use estimates and projections based on the decennial population count data between Censuses. For example, population control totals for CHIS 2009 were based on DOF estimates and projections, which were based on Census 2000 counts with adjustments for demographic changes within the state between 2000 and 2009. These estimates become less accurate and more dependent on the models underlying the adjustments over time. Using the most recent Census population count information to create control totals for weighting produces the most statistically accurate population estimates for the current cycle, but it may produce unexpected increases or decreases in some survey estimates when comparing survey cycles that use 2000 Census-based information and 2010 Census-based information. See *CHIS 2011-2012 Methodology Series: Report 5 – Weighting and Variance Estimation* for more information on the weighting process.

1.7 Imputation Methods

Missing values in the CHIS data files were replaced through imputation for nearly every variable. This was a massive task designed to enhance the analytic utility of the files. Westat imputed missing values for those variables used in the weighting process and UCLA-CHPR staff imputed values for nearly all other variables.

Two different imputation procedures were used by Westat to fill in missing responses for items essential for weighting the data. The first imputation technique was a completely random selection from the observed distribution of respondents. This method was used only for a few variables when the percentage of the items missing was very small. The second technique was hot deck imputation without

replacement. The hot deck approach is one of the most commonly used method for assigning values for missing responses. 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. The characteristics defining “similar” vary for different variables. To carry out hot deck imputation, the respondents who answer a survey item form a pool of donors, while the item non-respondents are a group of recipients. A recipient is matched to the subset pool of donors based on household and individual characteristics. A value for the recipient is then randomly imputed from one of the donors in the pool. Once a donor is used, it is removed from the pool of donors for that variable. Hot deck imputation was used to impute the same items in CHIS 2003, CHIS 2005, CHIS 2007, CHIS 2009, and CHIS 2011-2012 (i.e., race, ethnicity, home ownership, and education).

UCLA-CHPR imputed missing values for nearly every variable in the data files other than those imputed by Westat and some sensitive variables in which nonresponse had its own meaning. Overall, item nonresponse rates in CHIS 2011-2012 were low, with most variables missing valid responses for less than 2% of the sample. However, there were a few exceptions where item nonresponse rate was greater than 20%, such as household income.

The imputation process conducted by UCLA-CHPR started with data editing, sometimes referred to as logical or relational imputation: for any missing value, a valid replacement value was sought based on known values of other variables of the same respondent or other sample(s) from the same household. For the remaining missing values, model-based hot-deck imputation with donor replacement was used. This method replaces a missing value for one respondent using a valid response from another respondent with similar characteristics as defined by a generalized linear model with a set of control variables (predictors). The link function of the model corresponds to the nature of the variable being imputed (e.g., linear regression for continuous variables, logistic regression for binary variables, etc.). Donors and recipients are grouped based on their predicted values from the model.

Control variables (predictors) used in the model to form donor pools for hot-decking always included the following: gender, age group, race/ethnicity, poverty level (based on household income), educational attainment, and region. Other control variables were also used depending on the nature of the imputed variable. Among the control variables, gender, age, race/ethnicity and regions were imputed by Westat. UCLA-CHPR then imputed household income and educational attainment in order to impute other variables. Household income, for example, was imputed using the hot-deck method within ranges from a set of auxiliary variables such as income range and/or poverty level.

The imputation order of the other variables followed the questionnaire. After all imputation procedures were complete, every step in the data quality control process is performed once again to ensure consistency between the imputed and non-imputed values on a case-by-case basis.

1.8 Methodology Report Series

A series of five methodology reports is available with more detail about the methods used in CHIS 2011-12:

- Report 1 – Sample Design;
- Report 2 – Data Collection Methods;
- Report 3 – Data Processing Procedures;
- Report 4 – Response Rates; and
- Report 5 – Weighting and Variance Estimation.

For further information on CHIS data and the methods used in the survey, visit the California Health Interview Survey Web site at <http://www.chis.ucla.edu> or contact CHIS at CHIS@ucla.edu.

2. TELEPHONE SAMPLING METHODS

This chapter describes the sampling methods used in the CHIS 2011-2012 telephone survey. CHIS 2011-2012 employed a dual-frame with two main components and several supplemental samples. The main components are a landline random digit dialing (RDD) sample with approximately 80 percent of the dialed telephone numbers and a cell phone RDD sample with the remaining 20 percent of dialed numbers. The supplemental samples include one geographic sample in San Diego County, Korean and Vietnamese surname list samples, and an American Indian and Alaska Native list sample. The landline sample, geographic supplemental sample, and cell phone sample were drawn using RDD approaches while the list samples were drawn from separate lists of telephone numbers. Beginning in 2011, CHIS data are collected continuously across the two-year data collection cycle. CHIS 2011-2012 data collection began on June 15, 2011 and concluded on January 14, 2013.

The first section describes the list-assisted RDD sampling methodology for the landline sample component. It also discusses some sources of undercoverage associated with landline telephone samples, such as persons who cannot be interviewed because of language limitations.

The second section describes the cell phone sampling methodology used to address the problems associated with the increasing noncoverage of landline samples due to greater reliance on cellular telephone use and a drop in landline telephone services.

The third section describes the methods used to increase the efficiency of the landline sample through the use of tritone and business purges of unproductive numbers to reduce the number of calls to sampled but ineligible telephone numbers.

The last section reviews the supplemental samples in CHIS 2011-2012. As in previous cycles of CHIS, geographic areas with high concentrations of Korean and Vietnamese populations of interest were oversampled in the landline sample. The sample yield for these groups was also increased by sampling lists of telephone numbers where the owner is likely to be Korean or Vietnamese based on surname. CHIS 2011-2012 included an American Indian and Alaska Native supplemental list sample drawn from a list of telephone numbers of users served by Indian Health Service (IHS) health clinics in California.

2.1 List-Assisted Random Digit Dial Sampling of Landlines

List-assisted RDD sampling has been the primary method for landline telephone samples for all cycles of CHIS. This method was designed to produce an unclustered sample that has good operational features (Tucker, Lepkowski, & Piekarski, 2002). In 100 series list-assisted sampling, the set of all landline telephone numbers in operating telephone prefixes is composed of 100-banks, each containing 100 telephone numbers with the same first eight digits. All 100-banks with at least one residential number listed in a published telephone directory comprise the sampling frame. A simple random or a systematic sample of telephone numbers is selected from the landline frame. Initially, this method had a small amount of noncoverage because telephone numbers in 100-banks with no listed telephone numbers (i.e., zero banks) were not sampled. Brick, Waksberg, Kulp, & Starer (1995) showed that the bias from this approach was negligible for most estimates.

Changes in the structure of the U.S. telecommunications industry and an increasing number of residential exchanges have had a large impact on the 100 series list-assisted methodology. Fahimi, Kulp, & Brick (2008) found that the exclusion of 100-banks without any listed telephone number could result in coverage losses of up to 20 percent of the households with a landline. Although there is no current information on the characteristics of the households, it is likely that these households have different characteristics from those in the traditionally sampled banks. Methods for addressing this problem are being studied for implementation in future cycles of CHIS. Although the CHIS 2011-2012 landline sample does not have a specific method to address this potential undercoverage bias directly, the weighting methods using control totals representing the entire population in California should mitigate its effects.

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 typically treated as nonresponse, but could 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 CHIS 2011-2012 and previous cycles, significant efforts have been made to limit this potential bias by interviewing in multiple languages (Lee, Nguyen, Jawad, & Kurata, 2008). In CHIS 2011-2012, interviews were conducted in five languages: English, Spanish, Korean, Vietnamese, and Chinese (Cantonese and Mandarin dialects). This effort eliminates a potentially large source of the bias that might result if interviews had only been conducted in English.

2.2 Households without Landline Telephones

In landline telephone surveys, households with no access to landline telephones (households with only cellular telephones and households with no telephone service of any type) are not sampled. For estimates correlated with socioeconomic measures such as health insurance coverage, food security, and poverty, this undercoverage introduces biases. The bias depends on the number of households with no landline telephones and the difference in characteristics of persons in households with and without a landline telephone.

Households with only cellular service account for the largest proportion of those without a landline. The numbers of households and persons in the United States who have cell phones have greatly increased in the last few years. The most recent estimate of cell-phone-only households is 39.4 percent nationally for the first 6 months of 2013 (Blumberg & Luke, 2013). They also reported that a sizeable proportion of households may be difficult to reach even though they have a landline because they rely on cell phones for most of their calls. This source of bias is likely to continue to grow along with the prevalence of cell phones.

The characteristics of persons in cell-phone-only households are different from those in households with landlines. For example, the cell-phone-only adults are much less likely to be insured than the adults in households with landlines. Demographic differences such as age and gender are also associated with cell-phone-only households, although some of these characteristics are changing as more people use cell phones. Additionally, adults living in cell-only households are more likely than those in households with landlines to be renters or living with unrelated adults. Since this population is excluded from landline-based telephone surveys, there is increasing concern about the quality of estimates from this type of survey.

CHIS 2011-2012 included a cell phone sample component that addresses the biases from excluding cell-phone-only households in landline telephone surveys. Similar to CHIS 2009, the CHIS 2011-2012 cell sample collected information from households with landlines who were contacted through the cell phone sample. The cell phone sample was also used to collect information on children and adolescents as in CHIS 2009. Additional details on the selection of this sample are presented in [Section 3.2.2](#).

2.3 Increasing the Efficiency of Data Collection

When landline telephone numbers are sampled, special procedures are often implemented before data collection to reduce costs and to increase the efficiency of sampling and data collection. These techniques have been used in all previous cycles of CHIS, although some of the details of the procedures may have evolved over time.

The CHIS 2011-2012 landline sample was processed using tritone tests (the distinctive three-bell sound heard when dialing a nonworking number) and business purge methods to reduce the number of unproductive numbers (i.e., business and nonworking numbers). The procedure, called Comprehensive Screening Service (CSS), is offered by Marketing Systems Group (MSG), the vendor that also provided the sampling frames for CHIS. CSS is an attended screening process that first removes all listed business telephone numbers. The remaining numbers are then dialed to screen out nonworking and additional business numbers. The procedure also identifies cell phone numbers that were ported from landline exchanges. These ported numbers have been included as part of the cell sample since CHIS 2009.

[Table 2-1](#) shows the CSS result codes as well as the distribution of the sampled telephone numbers in CHIS 2011-2012. Approximately 55 percent of the sampled numbers (CSS result codes LB, FM, NR NW, and some UB) were excluded from dialing. This was 6 percentage points higher than the 49 percent purged in CHIS 2009.

Table 2-1. CSS result codes and their distribution in the CHIS 2011-2012 sample

CSS result code	Description	Number of telephones	Percentage
CP	Agent identified cell phone	19	0.00
DK	Undetermined	257,580	32.56
FM	Fax/modem	22,828	2.89
LA	Language barrier	4,116	0.52
LB	Listed business	28,079	3.55
NR	No ring-back	9,241	1.17
NW	Nonworking	338,699	42.81
PM	Privacy manager	4,325	0.55
RS	Residence	86,904	10.98
UB	Unlisted business	37,220	4.70
WR	Wireless number	2,203	0.28
Total		791,214	100.00

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

2.4 Supplemental Sampling

The first type of supplemental sample implemented in CHIS 2011-2012 was geographic sampling used to increase the sample size in specified geographic areas. CHIS 2011-2012 included one supplemental sample in San Diego County for a target of 4,800 completed adult interviews. The selection of this sample is described in [Section 3.2.3](#).

The second type of supplemental sampling was used to improve the sample size and precision of the estimates for specific race and ethnic groups. As mentioned in [Chapter 1](#), one of the goals of CHIS 2011-2012 and previous cycles was to produce reliable estimates for Koreans and Vietnamese in California. These two ethnic groups are important for analytical reasons, but constitute a small proportion of the total California population. The expected sample yield from the landline sample was too small to support inferences for these groups at the desired level of precision. Since CHIS 2003, two sampling strategies have been used to meet a target sample yield of 500 for Korean and 500 Vietnamese adult interviews per cycle (Edwards, Brick, Flores Cervantes, DiSogra, & Yen, 2002): disproportionate stratified sampling and multiple frame sampling. These strategies are mainly used to oversample rare or small populations (Flores Cervantes & Kalton, 2007).

The first strategy for oversampling Korean and Vietnamese populations was geographic targeting using the same substrata used since 2003. These strata were created classifying exchanges based on the concentration of Korean and Vietnamese residing in the exchange³ within selected counties. Under disproportionate stratified sampling, telephone numbers in exchanges located in areas with a relatively high proportion of members (high-density strata) were sampled at a higher rate than the numbers in the other areas (low-density strata). Since the stratification was based on information from the 2000 Census, we examined the observed sample from the previous cycles and reclassified the telephone exchanges using the sample distribution of these populations in previous cycles of CHIS. Reclassifying exchanges reflected changes in the Korean and Vietnamese populations in these areas.

The second strategy to increase the number of Korean and Vietnamese interviews included supplemental samples from other frames (i.e., surname lists of these groups). This sampling strategy is based on the concept of multiple frame design. In this approach, 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(s). The list frame does not have to be complete to be useful, although the more complete the list is, the greater the potential for increasing the precision of the estimates. The composition of the list

³ Refer to the *CHIS 2003 Methodology Series: Report 1 Sample Design* for additional details on the creation of the substrata.

affects its efficiency (that is, the proportion of sampled numbers that leads 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. The cost associated with the use of the surname lists is much lower than the cost for locating and interviewing members of the groups from the landline sample.

The identification of eligible (i.e., Korean or Vietnamese) adults in the list samples was done through a question in the screener interview. This strategy was relatively simple to implement and has good statistical properties, except for any measurement error that may be introduced by asking a question about the ethnicity of the adults at the beginning of the interview. Screening was not necessary for the cases sampled from the high/low density strata because these cases were part of the base landline sample where all households are eligible for further interviewing. Although the use of surname lists was an effective way to increase the number of completed interviews for these groups, the variance of estimates for these groups is not greatly reduced by this approach.

CHIS 2011-2012 also included a supplemental sample of American Indian/Alaska Native residents of California to increase the representation of this group. The oversample was produced using a list of users who had been served by the Indian Health Service (HIS) health clinics in California. This supplemental sample was treated in the same way as the surname samples, including a self-identification question in the screener. As with the surname samples, this approach increased the number of American Indian or Alaska Native cases, but the variance of estimates for these groups is not greatly reduced because the list includes only a small proportion of the target population.

3. SAMPLING HOUSEHOLDS

This chapter describes the sample design and selection of households for CHIS 2011-2012. We begin by defining the target population and the persons included in and excluded from the survey. Target numbers of completed adult interviews by county and for the supplemental samples are then described. The remainder of the chapter describes the types of supplemental samples and the selection of telephone numbers in order to achieve the stated goals.

3.1 Population of Interest

As in previous CHIS cycles, the 2011-2012 sample was intended to represent the adult (age 18 and older) residential population of California, as well as adolescents (age 12-17) and children (age 11 and younger). 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 [Chapter 2](#), some individuals who were part of the residential population did not have a chance of selection. These include those living in households without any telephone service, and children and adolescents living in a household without a parent or legal guardian.

3.2 Sample Design

The principal goals of the CHIS 2011-2012 sample design were (1) to produce reliable statewide estimates for the total population in California and for its larger race/ethnic groups, as well as for several smaller ethnic groups (i.e., Koreans and Vietnamese), and (2) to produce reliable estimates at the county level for as many counties as possible. In CHIS 2011-2012, a landline sample, a cell phone sample, and surname list samples were drawn in order to meet these goals. These samples are described in the following sections.

The goals of the survey required compromises allocating the sample into strata and frame type. To achieve the most reliable statewide estimates, the optimal design is to allocate the sample to counties proportionately to their population. On the other hand, the optimal allocation for producing individual county-level estimates is to assign each county an equal sample size. Different allocations of the sample by stratum and telephone sample (i.e., landline or cell phone) consistent with the available budget were evaluated at the beginning of the study. The UCLA CHIS staff consulted with various constituencies to assess the relative importance of particular types of estimates. Westat statistical staff helped evaluate each alternative and examined the consequences of the sample allocations.

The initial goal for CHIS 2011-2012 was to complete 48,000 adult interviews as shown in [Table 3-1](#). This goal included 500 adult interviews each for Koreans and Vietnamese, including those from the RDD samples and those sampled from surname lists. Unlike previous cycles of CHIS, the landline and cell phone sample targets were defined separately by the 44 geographic sampling strata as indicated in the table. The initial sample was allocated so 25 percent of the adult interviews would be completed from the cell phone frame.

Table 3-1. Initial targets for completed adult interviews by county (excluding the AIAN supplemental sample)

	Stratum	Landline sample	Cell phone sample	Total	Population size
	State total	36,000	12,000	48,000	
1	Los Angeles	7,204	2,401	9,605	Over 9,000,000
2	San Diego	2,449	816	3,265	
3	Orange	2,208	736	2,944	
4	Santa Clara	1,341	447	1,788	
5	San Bernardino	1,416	472	1,888	1,200,000 or greater
6	Riverside	1,623	541	2,164	
7	Alameda	1,213	404	1,617	
8	Sacramento	1,179	393	1,572	
9	Contra Costa	852	284	1,136	800,000 to 1,200,000
10	Fresno	662	221	883	
11	San Francisco	760	253	1,013	
12	Ventura	596	199	795	
13	San Mateo	579	193	772	500,000 to 800,000
14	Kern	593	198	791	
15	San Joaquin	500	167	667	

Table 3-1. Initial targets for completed adult interviews by county (excluding the AIAN supplemental sample) (continued)

	Stratum	Landline sample	Cell phone sample	Total	Population size
16	Sonoma	450	150	600	
17	Stanislaus	450	150	600	
18	Santa Barbara	450	150	600	
19	Solano	450	150	600	
20	Tulare	450	150	600	
21	Santa Cruz	450	150	600	
22	Marin	450	150	600	
23	San Luis Obispo	450	150	600	
24	Placer	450	150	600	
25	Merced	450	150	600	Medium counties 100,000 to 500,000
26	Butte	450	150	600	
27	Shasta	450	150	600	
28	Yolo	450	150	600	
29	El Dorado	450	150	600	
30	Imperial	450	150	600	
31	Napa	450	150	600	
32	Kings	450	150	600	
33	Madera	450	150	600	
34	Monterey	450	150	600	
35	Humboldt	450	150	600	
36	Nevada	450	150	600	
37	Mendocino	450	150	600	Small counties less than 100,000 population per county
38	Sutter	450	150	600	
39	Yuba	450	150	600	
40	Lake	450	150	600	
41	San Benito	450	150	600	
42	Colusa, Glenn, Tehama	375	125	500	
43	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	375	125	500	Small counties combined
44	Amador, Alpine, Calaveras, Inyo, Mariposa, Mono, Tuolumne	375	125	500	

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

The initial overall goal was reduced by 8,000 interviews by the end of 2011. The sample was reallocated so that 20 percent of adult interviews would be completed from the cell phone frame. During 2012, the overall goal was increased with supplemental samples that included additional landline cases in Los Angeles County (see [Section 3.2.1](#)), additional cell phone cases in seven counties with a large concentration of Asians (see [Section 3.2.2](#)), a landline/cell phone sample supplemental sample in the San

Diego County (see [Section 3.2.3](#)), and an American Indian and Alaska Native (AIAN) list sample (see [Section 3.2.5](#)). The cell phone goals were also revised to take into account the distributions of adults with a cell phone at the county level. The state-level allocation was maintained at 80 percent landline to 20 percent cell phone, but this allocation varied across counties. The CHIS 2011-2012 final goal was 42,656 statewide adult interviews, as shown in [Table 3-2](#). This final goal was 756 interviews more than the CHIS 2009 goal (41,900 adult interviews).

Although the number of child and adolescent interviews was not predetermined, we expected to get approximately 3,000 completed adolescent interviews (depending on compliance since parental consent and adolescent agreement are required) and approximately 8,000 child interviews based on CHIS 2009.

Table 3-2. Final targets for completed adult interviews by county (excluding the AIAN supplemental sample)

	Stratum	Landline sample	Cell phone sample	Total
	State total	32,578	9,457	42,035
1	Los Angeles	6,506	2,114	8,620
2	San Diego	3,840	960	4,800
3	Orange	1,749	687	2,436
4	Santa Clara	1,083	480	1,563
5	San Bernardino	1,088	346	1,434
6	Riverside	1,353	292	1,645
7	Alameda	998	359	1,357
8	Sacramento	981	214	1,195
9	Contra Costa	704	160	864
10	Fresno	412	259	671
11	San Francisco	579	194	773
12	Ventura	471	133	604
13	San Mateo	479	288	767
14	Kern	479	121	600
15	San Joaquin	372	134	506
16	Sonoma	354	146	500
17	Stanislaus	395	105	500
18	Santa Barbara	414	86	500
19	Solano	395	105	500
20	Tulare	378	122	500
21	Santa Cruz	407	93	500
22	Marin	442	58	500

Table 3-2. Final targets for completed adult interviews by county (excluding the AIAN supplemental sample) (continued)

	Stratum	Landline sample	Cell phone sample	Total
23	San Luis Obispo	405	95	500
24	Placer	383	117	500
25	Merced	433	67	500
26	Butte	376	124	500
27	Shasta	400	100	500
28	Yolo	352	148	500
29	El Dorado	383	117	500
30	Imperial	427	73	500
31	Napa	461	39	500
32	Kings	456	44	500
33	Madera	456	44	500
34	Monterey	290	210	500
35	Humboldt	287	213	500
36	Nevada	410	90	500
37	Mendocino	420	80	500
38	Sutter	430	70	500
39	Yuba	437	63	500
40	Lake	452	48	500
41	San Benito	461	39	500
42	Colusa, Glenn, Tehama	344	56	400
43	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	291	109	400
44	Amador, Alpine, Calaveras, Inyo Mariposa, Mono, Tuolumne	345	55	400

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

3.2.1 Landline Sample

The revised CHIS 2011-2012 statewide landline goal was 32,578 adult interviews as shown in [Table 3-3](#). This goal includes the landline portion of the geographic supplemental sample for San Diego (1,855 interviews) and the Korean and Vietnamese surname samples, but excludes the AIAN supplemental list sample. During data collection, the initial landline goal in Los Angeles County was increased by 676 cases to be completed in the Antelope Valley and Metro Service Planning Areas (SPAs). After the additional release, the final target was 600 completed adult interviews in Antelope Valley and 1,300 adult interviews in the Metro SPA.

The stratification of the landline frame for the California's 58 counties used in CHIS 2011-2012 was the same as that used since 2005. The design consisted of 44 strata, with 41 single-county strata and 3

strata with multiple counties. The multiple-county strata were created by grouping the remaining counties into three geographic areas. The stratum assignment was based on the population residing in the county. [Table A-1](#) in the Appendix shows the assignment of counties to geographic strata across the CHIS cycles.

Because of the need to produce reliable estimates at the county level, the sample allocation was not proportional to the population in the counties. With a proportional allocation, the estimates from the smaller counties would be based on small sample sizes and would not be adequate for the envisioned analyses. To achieve the goal of producing local or county estimates, the target sample sizes from medium and smaller counties was fixed at 500 or 400 interviews. The remaining sample was allocated proportionately by population size. More details about the landline sample are given after discussing the designs for the other samples.

Table 3-3. Final targets for completed adult interviews from the landline sample by county (excluding the AIAN supplemental sample)

	Stratum	Revised goal	Geographic supplemental sample	Additional release	Final landline goal
	State total	30,047	1,855	676	32,578
1	Los Angeles	5,830	0	676	6,506
2	San Diego	1,985	1,855	0	3,840
3	Orange	1,749	0	0	1,749
4	Santa Clara	1,083	0	0	1,083
5	San Bernardino	1,088	0	0	1,088
6	Riverside	1,353	0	0	1,353
7	Alameda	998	0	0	998
8	Sacramento	981	0	0	981
9	Contra Costa	704	0	0	704
10	Fresno	412	0	0	412
11	San Francisco	579	0	0	579
12	Ventura	471	0	0	471
13	San Mateo	479	0	0	479
14	Kern	479	0	0	479
15	San Joaquin	372	0	0	372
16	Sonoma	354	0	0	354
17	Stanislaus	395	0	0	395
18	Santa Barbara	414	0	0	414
19	Solano	395	0	0	395
20	Tulare	378	0	0	378
21	Santa Cruz	407	0	0	407
22	Marin	442	0	0	442
23	San Luis Obispo	405	0	0	405
24	Placer	383	0	0	383
25	Merced	433	0	0	433

Table 3-3. Final targets for completed adult interviews from the landline sample by county (excluding the AIAN supplemental sample) (continued)

	Stratum	Revised goal	Geographic supplemental sample	Additional release	Final landline goal
26	Butte	376	0	0	376
27	Shasta	400	0	0	400
28	Yolo	352	0	0	352
29	El Dorado	383	0	0	383
30	Imperial	427	0	0	427
31	Napa	461	0	0	461
32	Kings	456	0	0	456
33	Madera	456	0	0	456
34	Monterey	290	0	0	290
35	Humboldt	287	0	0	287
36	Nevada	410	0	0	410
37	Mendocino	420	0	0	420
38	Sutter	430	0	0	430
39	Yuba	437	0	0	437
40	Lake	452	0	0	452
41	San Benito	461	0	0	461
42	Colusa, Glenn, Tehama	344	0	0	344
43	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	291	0	0	291
44	Amador, Alpine, Calaveras, Inyo, Mariposa, Mono, Tuolumne	345	0	0	345

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

The landline sampling frame was created by stratifying 100-banks with one or more listed telephone numbers into nonoverlapping strata, each corresponding to a county or a group of counties as shown in [Table 3-3](#). The procedure for assigning the numbers to strata was the same as that used in previous CHIS cycles. The geographic information required for stratification was available only at the exchange level,⁴ so 100-banks could not be assigned directly to a single stratum. All banks within an exchange were stratified indirectly by mapping the exchanges to a county represented by the stratum. However, some telephone exchanges actually service households in more than one county.

To solve the stratification problem, the procedure used coverage reports for each county produced by MSG, the sampling vendor. The coverage reports listed all the exchanges in the county. For each exchange, the report showed the total number of listed households in the exchange and the proportion of listed households that were within the county. After combining information from the coverage reports for all 58 counties, we created a frame of exchanges with variables for the number of listed households in

⁴ 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.

each county that the exchange covers. Each exchange was then assigned to the county with the most listed households. The telephone exchanges in Los Angeles County were stratified in 8 substrata each representing a SPA using ZIP Code information. Telephone exchanges that crossed SPAs were assigned to the SPA with the most listed households. At the beginning of the study there were no targets for individual SPAs, so the sample for Los Angeles was allocated proportionally by these substrata, except for the sample for Antelope Valley. The initial sample for Antelope Valley included an additional sample to yield 250 adult interviews more than what would be expected from proportional allocation. After additional funding, the sample in the Antelope Valley and Metro SPAs was increased to meet the new goals.

As mentioned in [Chapter 2](#), disproportionate stratified sampling was used to oversample Koreans and Vietnamese without increasing the sample size allocated to any stratum (the stratum sample size was fixed). An analysis done in CHIS 2003 to help with the allocation found that six percent or more Korean or Vietnamese in the exchanges was optimal for the creation of the substrata. In addition, the analysis showed that oversampling the substrata with high concentration at twice the rate of the low concentration strata did not inordinately inflate the design effect nor decrease the effective sample sizes for other race-ethnic groups of interest. See *CHIS 2003 Methodology Series: Report 1 - Sample Design* for additional details of the analysis for the creation of high- and low-density substrata.

Since the creation of the high/low density designation used information from Census 2000, the assignment of telephone exchanges has been revised in past cycles of CHIS using tabulations of the number of Korean or Vietnamese interviews by telephone exchange. Using this information, some exchanges have been reallocated to the high/low density strata depending on the number of interviews completed from adults of Korean or Vietnamese descent. The high/low density subsampling strata were created in San Diego County, Orange County, and Santa Clara County. Fourteen substrata were created in Los Angeles County by classifying the SPAs into high/low density substrata.

[Table 3-4](#) shows the definition of the substrata for Los Angeles County, San Diego County, Orange County, and Santa Clara County. The table also shows the number of telephone exchanges and the estimated number of households in the substrata.

Table 3-4. Definition of sampling substratum, number of exchanges, and total number of households for Los Angeles County, San Diego County, Orange County, and Santa Clara County

Stratum	Substratum	SPA/Service region	Density	Number of telephone exchanges	Number of households
1. Los Angeles	1.12	San Fernando SPA	High	36	35,516
	1.13	San Gabriel SPA	High	81	73,393
	1.14	Metro SPA	High	120	50,226
	1.17	South SPA	High	29	20,931
	1.18	South Bay SPA	High	53	39,194
	1.21	Antelope Valley SPA	Low	52	52,000
	1.22	San Fernando SPA	Low	440	332,669
	1.23	San Gabriel SPA	Low	258	188,759
	1.24	Metro SPA	Low	179	109,035
	1.25	West SPA	Low	266	128,384
	1.26	South SPA	Low	170	134,468
	1.27	East SPA	Low	192	157,737
	1.28	South Bay SPA	Low	265	203,469
2. San Diego	2.12	North Central	High	62	33,641
	2.13	Central SR	High	33	37,326
	2.21	North Coastal SR	Low	87	80,223
	2.22	North Central SR	Low	104	55,820
	2.23	Central SR	Low	82	40,951
	2.24	South SR	Low	98	72,208
	2.25	East SR	Low	69	77,049
	2.26	North Inland SR	Low	121	90,513
3. Orange	3.1	N/A	High	281	167,502
	3.2	N/A	Low	418	317,223
4. Santa Clara	4.1	N/A	High	164	82,001
	4.2	N/A	Low	329	201,297
Total				3,989	2,781,535

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

3.2.2 Cell Phone Sample

The CHIS 2011-2012 cell phone sample had a final state-wide target of 9,457 adult interviews. Unlike both CHIS 2007 and CHIS 2009, where the cell phone sample targets were set at the region level, the CHIS 2011-2012 cell sample targets were set to the same strata (county and groups of counties) defined for the landline sample shown in [Table 3-5](#).

Table 3-5. Final targets for completed adult interviews from the cell phone sample by county

Stratum		Initial target	Geographic supplemental sample	Additional release	Final target
	State total	7,634	464	1,359	9,457
1	Los Angeles	1,469	0	645	2,114
2	San Diego	496	464	0	960
3	Orange	489	0	198	687
4	Santa Clara	276	0	204	480
5	San Bernardino	346	0	0	346
6	Riverside	292	0	0	292
7	Alameda	231	0	128	359
8	Sacramento	214	0	0	214
9	Contra Costa	160	0	0	160
10	Fresno	259	0	0	259
11	San Francisco	190	0	4	194
12	Ventura	133	0	0	133
13	San Mateo	108	0	180	288
14	Kern	121	0	0	121
15	San Joaquin	134	0	0	134
16	Sonoma	146	0	0	146
17	Stanislaus	105	0	0	105
18	Santa Barbara	86	0	0	86
19	Solano	105	0	0	105
20	Tulare	122	0	0	122
21	Santa Cruz	93	0	0	93
22	Marin	58	0	0	58
23	San Luis Obispo	95	0	0	95
24	Placer	117	0	0	117
25	Merced	67	0	0	67
26	Butte	124	0	0	124
27	Shasta	100	0	0	100
28	Yolo	148	0	0	148
29	El Dorado	117	0	0	117
30	Imperial	73	0	0	73
31	Napa	39	0	0	39
32	Kings	44	0	0	44
33	Madera	44	0	0	44
34	Monterey	210	0	0	210
35	Humboldt	213	0	0	213
36	Nevada	90	0	0	90
37	Mendocino	80	0	0	80
38	Sutter	70	0	0	70

Table 3-5. Final targets for completed adult interviews from the cell phone sample by county (continued)

Stratum		Initial target	Geographic supplemental sample	Additional release	Final target
39	Yuba	63	0	0	63
40	Lake	48	0	0	48
41	San Benito	39	0	0	39
42	Colusa, Glenn, Tehama	56	0	0	56
43	Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	109	0	0	109
44	Amador, Alpine, Calaveras, Inyo, Mariposa, Mono, Tuolumne	55	0	0	55

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

The final statewide target includes 464 cell phone adult interviews from the San Diego geographic sample and 1,359 interviews from an additional cell phone sample release made during the middle of data collection. The later supplemental sample targeted six counties with a large concentration of Asians: Los Angeles County, Orange County, Santa Clara County, Alameda County, San Francisco County, and San Mateo County.

The cell phone sample design was different from the landline design and presented its own challenges. The main cell phone sample was drawn by the sampling vendor using the latest Telcordia database. Unlike the landline sample where the numbers were drawn from banks with 100 numbers, the cell phone numbers were drawn from groups of 1,000 numbers (i.e., 1,000-series blocks) in California dedicated to wireless service.⁵ Telephone numbers that were ported from a landline to a cell phone could not be selected from these exchanges because these numbers were in exchanges assigned to landlines. To address this problem, telephone numbers identified as ported cell phones in the base landline sample were included as part of the cell phone sample. The ported numbers were identified by disposition code in the CSS (see codes WR and CP in [Table 2-1](#)). There were close to 3,000 ported cell phone numbers identified in the landline sample. This is similar to the number of ported numbers identified in 2009. The remainder of this section discusses the sampling of the main cell sample.

A difference between the landline and cell phone samples is the lack of detailed demographic and socio-economic information (e.g., number of households, percentage of homeowners, African Americans, etc.) on the geographic area where the cell phone is sampled. Although cell phone numbers are sampled from exchanges assigned to wireless service, the geographic area covered by the exchange does not

⁵ There are some additional, technical restrictions in the sampling, such as making sure the number can be dialed into and that toll-free numbers are excluded.

necessarily indicate where the owner of the number resides. This is because the cell phone exchange generally corresponds to where the cell phone was purchased or activated. Thus, the cell sample could not easily be stratified to match the stratification of the landline sample.

Since there is no precise information on the geographic area covered by the cell phone exchange, the cell phone sampling strata were created in an indirect way. Utilizing data from the CHIS 2007 and 2009 cell phone respondents and their pre-assigned FIPS county code, supplied by the sampling vendor, we were able to define combinations of area codes and/or counties that closely predicted the self-reported county of the respondent. Using this information we created 28 cell sample strata for CHIS 2011, an increase from the 7 region strata used in CHIS 2007 and 2009, but fewer than the 44 strata that were used for the landline sample.

[Table 3-6](#) and [Table 3-7](#) show the area code and/or FIPS county code combinations used to define the CHIS 2011-2012 cell phone sample strata. [Table 3-8](#) shows those counties that made up each of the 28 cell phone sample strata along with the total number of records sampled for each stratum.

Table 3-6. Definition of cell phone sampling strata for complete area codes

Area code(s)	Stratum
213, 310, 323, 424, 562, 626, 747, 818	1
619, 858	2
657, 714, 949	3
408	4
909	5
951	6
510	7
916	8
925	9
559	10
650	13
209	15

Source: UCLA Center for Health Policy Research, 2011 California Health Interview Survey.

Table 3-7. Definition of cell phone sampling strata based on area code and county combinations

Area code	FIPS code(s)	Counties included	Stratum
415	06001, 06081	Alameda, San Mateo	22
415	Any other	All except Alameda & San Mateo	11
530	06007	Butte	26
530	06089	Shasta	27

Table 3-7. Definition of cell phone sampling strata based on area code and county combinations (continued)

Area code	FIPS code(s)	Counties included	Stratum
530	06015, 06035, 06049, 06093	Del Norte, Lassen, Modoc & Siskiyou	43
530	Any other	All except Butte, Shasta, Del Norte, Lassen, Modoc & Siskiyou	28
661	06029	Kern	14
661	Any other	All except Kern	1
707	06023	Humboldt	35
707	06015, 06035, 06049, 06093	Del Norte, Lassen, Modoc & Siskiyou	43
707	Any other	All except Humboldt, Del Norte, Lassen, Modoc & Siskiyou	16
760	06073	San Diego	30
760	06027	Sierra Counties	44
760	All other	All except San Diego & Sierra Counties	6
805	06083	Santa Barbara	18
805	06079	San Luis Obispo	23
805	Any other	All except Santa Barbara & San Luis Obispo	12
831	06053	Monterey	34
831	Any other	All except Monterey	21

Source: UCLA Center for Health Policy Research, 2011 California Health Interview Survey.

Table 3-8. Number of cell telephone numbers drawn by sampling stratum

Sampling stratum	Counties covered	Total sampled
State total	All	134,648
1	Los Angeles	31,174
2	San Diego	6,618
3	Orange	9,541
4	Santa Clara	5,783
5	San Bernardino	5,782
6	Riverside	5,843
7	Alameda	4,614
8	Sacramento, Placer	2,728
9	Contra Costa	3,179
10	Fresno, Tulare, Kings, Madera	6,457
11	San Francisco	2,264
12	Ventura	2,815
13	San Mateo	4,232
14	Kern	1,336
15	San Joaquin, Stanislaus, Merced	5,203
16	Sonoma, Solano, Napa	4,037
18	Santa Barbara	1,352

Table 3-8. Number of cell telephone numbers drawn by sampling stratum (continued)

Sampling stratum	Counties covered	Total sampled
21	Santa Cruz	1,118
22	San Francisco, Marin	1,372
23	San Luis Obispo	1,377
26	Butte, Tehama, Glenn, Colusa	1,563
27	Shasta	1,199
28	Yolo, El Dorado, Nevada, Sutter, Yuba	9,389
30	San Diego, Imperial	7,152
34	Monterey, San Benito	3,494
35	Humboldt, Mendocino, Lake	3,392
43	Del Norte, Siskiyou, Trinity, Modoc, Lassen, Plumas, Sierra	1,350
44	Amador, Alpine, Calaveras, Tuolumne, Mariposa, Mono, Inyo	284

Source: UCLA Center for Health Policy Research, 2011 California Health Interview Survey.

[Table A-2](#) in the Appendix shows the numbers drawn by sampling stratum and type of sample. When determining the sample size to draw, we used the observed response rates within the sampling strata from the cell sample in CHIS 2007 and 2009. Further, the misclassification rate between the sampled stratum and the self-reported county observed in CHIS 2007 and 2009 was taken into account when allocating the sample. The sample was allocated to cell phone strata by adjusting the number of cases to draw so that the expected number of self-reported cases was approximately 20 percent of each stratum's overall target.

There were also differences in the way the cell sample was processed after it was selected. Unproductive numbers in the cell phone sample (i.e., nonworking and business telephone numbers) could not be purged using directory matching because no cell phone directories exist. In addition, there are prohibitions on predictive dialing of cell phone numbers, so the other components of the CSS purging for nonproductive or nonworking numbers could not be done. Thus, the full selected cell phone sample was sent to be dialed by interviewers.

3.2.3 Supplemental Geographic Sample

In CHIS 2011-2012 one supplemental geographic sample was added at the request of San Diego County after funding was arranged. The final target was 800 completed adult interviews in each of the seven Health Districts in San Diego County. Officials in this county were interested in larger samples for a more detailed analysis. The San Diego geographic sample had both landline and cellphone targets as

shown in [Table 3-3](#) and [Table 3-4](#). Since the landline and cell phone components of the geographic sample covered an entire county, it was selected as part of the core landline and cell phone sample. Unlike some supplemental geographic samples in previous cycles, screening was not used to determine if the case was in the county of interest. As a result, there was no difference between the main and supplemental geographic sample instruments.

3.2.4 Supplemental Surname List Samples

Supplemental surname samples were used to increase the number of completed interviews of adults of Korean and Vietnamese descent. The statewide goal was 500 completed adult interviews from each ethnic group from the combined landline, cell phone, and surname samples. As in previous cycles of CHIS, the Korean and Vietnamese supplemental samples were drawn from lists, maintained by the sampling vendor, of telephone numbers associated with likely Korean and Vietnamese surnames. We screened the telephone numbers in these samples to determine eligible adults (i.e., adults of Korean or Vietnamese descent) in the household. If there were no eligible adults, the interview was terminated and the case was coded as ineligible.

[Table 3-9](#) shows the sampling goals for completed adult interviews with Koreans and Vietnamese in CHIS 2011-2012. The targets of the surname list sample were adjusted during data collection, as the actual landline and surname samples yields became known. The landline sample did not produce the expected number of Korean interviews; therefore, we drew additional numbers from the list frame during data collection.

Table 3-9. Targeted number of completed adult interviews for the Korean and Vietnamese samples

Subgroup	Targeted number of adult interviews		
	Landline & Cell phone sample	Supplemental list sample	Total
Korean	358	142	500
Vietnamese	233	267	500
Total	591	409	1,000

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

The surname list frames were created by the sampling vendor from telephone directories in California. The vendor provided three nonoverlapping surname frames; the first two frames included telephone numbers whose associated surnames were very likely to be Korean *only* or Vietnamese *only*,

and not any other ethnic group. The last frame included those surnames likely to be either Korean or some other ethnic group but not Vietnamese.

Separate samples were drawn from each of the three frames. A household was eligible for the extended interview if it included an adult who was either Korean or Vietnamese, regardless of which frame the number was drawn from. [Table 3-10](#) shows the size of the surname list frames used in 2011-2012 and the number of telephone numbers drawn from each frame.

Table 3-10. Surname frame sizes

Surname frame	Number of records
Korean only	33,924
Vietnamese only	76,202
Korean and some other race but not Vietnamese	93,625

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

3.2.5 American Indian and Alaska Native List Sample

CHIS 2011-2012 included a supplemental list sample to increase the number of American Indian/Alaska Native (AIAN) adult interviews and improve the statistical precision of estimates for this group. Unlike the Korean and Vietnamese surname samples, the sampling vendor does not maintain a list of telephone numbers of persons likely to be AIAN that can be used as sampling frame. The AIAN frame was created using a list provided by the Indian Health Service (IHS). It identified users of 19 IHS health clinics in California. The list was processed to remove users residing outside California, users without a telephone number, users with invalid or duplicate telephone numbers, and users with numbers with telephone exchanges not found in any of the landline and cell phone sample frames. The final frame had approximately 37 percent of telephones without an area code. The area code was completed using the most prevalent area code of the users without a missing area code in the clinic.

[Table 3-11](#) shows the frame size, the target and the number of telephone numbers drawn for the AIAN supplemental list sample. After sample selection, the sample was processed to identify landline and cell phone cases. Approximately 12 percent of the drawn cases were called using the cell phone sample protocols.

Table 3-11. Frame size and targeted number of completed adult interviews for the American Indian/Alaska Native supplemental list sample

	Frame size	Target
American Indian/Alaska Native	41,508	300

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

3.3 Sample Selection

The number of telephone numbers selected in any telephone survey has to be greater than the targeted number of completed interviews to account for a variety of factors. For example, a substantial percentage of the sampled telephone numbers are not residential. For CHIS 2011-2012 the sample of telephone numbers was inflated to deal with losses due to the following sources:

- Nonworking, nonresidential, and never answered numbers;
- Nonresponse to the screening interview;
- Nonresponse to the extended interview; and
- Ineligible households in the supplemental list samples.

The first three sources noted above are typical of all telephone surveys. To deal with these losses we used information from earlier CHIS cycles to estimate the percentage of telephone numbers that would not be residential and the percentage that would not respond to the screener and extended interviews, and increased the sample size accordingly. Estimates of the eligibility rates for the surname samples were taken from the corresponding CHIS 2009 samples.

Taking all of these factors into consideration, 928,739 telephone numbers were sampled for CHIS 2011-2012. Not all the telephone numbers were selected at the same time, as the sample design was modified several times during the field period to reflect the observed yield and changes in the targeted number of completed interviews. After each selection, duplicate telephone numbers (those numbers that had been previously sampled) were removed from the samples. [Table 3-12](#) summarizes the size of each sample. [Table A-2](#) in the Appendix shows the sample size by sampling stratum for the different samples. The data collection procedures are discussed in *CHIS 2011-2012 Methodology Series: Report 2 - Data Collection Methods*.

Table 3-12. Number of telephone numbers drawn by sample type

Sample type	Number of telephone numbers drawn
Landline sample	767,799
Main landline sample	698,218
Geographic sample	47,771
Supplemental landline sample	21,810
Cell phone sample	134,648
Main Cell phone sample	110,487
Geographic sample	5,105
Supplemental cell phone sample	19,056
List samples	26,292
Korean only surname	5,142
Korean and other surname	5,527
Vietnamese only surname	5,608
American Indian and Alaska Native user	10,015
Total	928,739

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

Release groups of telephone numbers were drawn and called throughout 2011 and 2012. Each release consisted of several replicates or random groups of telephone numbers drawn separately by sampling stratum and sample type (i.e., landline sample, cell phone) as indicated in [Table 3-13](#). The size of the major releases was computed as the total telephone numbers needed to complete predetermined six-month interim targets. The smaller releases included random groups of numbers in those strata with expected shortage in the number of completed interviews. All telephone numbers in a random group were called before a new random group was released. There were only two releases in the Korean/Vietnamese surname supplemental samples, one in 2011 and one in 2012. We expected to complete half of total surname goal in 2011 and the remaining half in 2012. There was only one release for the AIAN supplemental list sample because it was funded closer to the end of data collection.

Table 3-13. Release groups of telephone numbers by sample type

Sample type	Release	Date	Sample size
Landline	1	06/15/2011	241,881
	2	08/15/2011	80,474
	3	09/22/2011	62,243
	4	11/15/2011	198,887
	5	02/21/2012	37,207
	6	06/14/2012	100,255
	7	09/04/2012	22,678
	8	10/16/2012	24,174
Total			767,799

Table 3-13. Release groups of telephone numbers by sample type (continued)

Sample type	Release	Date	Sample size	
Cell sample	1	06/22/2011	60,000	
	2	10/27/2011	11,848	
	3	11/16/2011	53,073	
	4	05/29/2012	3,800	
	5	09/06/2012	341	
	6	11/07/2012	5,586	
	Total		134,648	
List samples				
	Korean/Vietnamese	1	08/29/2011	8,801
	Korean/Vietnamese	2	03/22/2012	7,476
	American Indian Alaska Native	3	12/11/2012	10,015
	Total		26,292	
Total			928,739	

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

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. This chapter describes the procedures for selecting the sample of persons within households for CHIS 2011-2012. Samples of adults, children, and adolescents within the households were selected using different sampling procedures, but one adult and up to one child and one adolescent were sampled within a given household. The within-household sampling procedures were developed to maximize the analytic utility of the data collected from the respondents.

The next section describes the within-household sampling alternatives we evaluated and the reasons for choosing the specific method of sampling. The second section describes the operational “child-first” procedure used to increase the number of child interviews. The last sections describe the methodology used for sampling adults, children, and adolescents in CHIS 2011-2012.

4.1 Sampling Alternatives

The general idea for the sample design over the CHIS cycles has been to sample one adult randomly from all the adults in a household associated with a sampled telephone number. In addition, in those households with adolescents (ages 12-17) and/or children (under age 12), one adolescent and one child were to be sampled and interviewed (a parent of the child was interviewed about the child). One approach to accomplishing these goals is simply to list all the persons in the age group (adult, child, and adolescent) in the household and select one person randomly from each group. We call this the *completely random* sampling method.

The completely random sampling method is not a problem in most households because most households have only one family. However, in households with two or more families, the completely random method could result in selecting persons who were not members of the same family. This situation is undesirable because the adult interview collects data about the family of the sampled adult. The data from the adult interview are of great value for the analysis of the data from the child and adolescent interviews. If the sampled child and/or sampled adolescent were not members of the same family as the sampled adult, then the data collected about them would be of very limited utility.

To solve this analytic problem, a second sampling alternative was adopted and has been used since CHIS 2001. We call this method the *linked* sampling approach. In this approach, children and/or

adolescents for whom a sampled adult was a blood or adoptive parent or a legal guardian were considered as linked to or “associated” with that adult.

In the linked sampling method, persons are sampled in two phases. In the first phase, an adult is randomly sampled from all eligible adults in the household. In the second phase, a child is sampled from all eligible children associated with the sampled adult. Since the sampling of children is a two-phase procedure, the probability of selection of the child is the product of the probability of selecting the adult (phase one) and the probability of selecting the child from all children associated with that adult (phase two). Adolescents are sampled in the same way, that is, one adolescent is selected from all adolescents associated with the adult sampled in the first phase.

To use the linked sampling method, data are needed to link children and adolescents in a household to the sampled adult and his/her spouse/partner (children or adolescents linked to both the sampled adult and spouse/partner could be selected if either adult was sampled). These data were collected during the screener interview or the adult interview in CHIS 2011-2012. We expected that in a very few households it would not be possible to link or associate a child or adolescent to an adult because of unusual household structures. A child or adolescent not associated with an adult does not have a chance of being selected. Beginning in 2003, the UCLA Institutional Review Board (IRB) directed that only children and adolescents of the sampled adult could be interviewed. Therefore, unassociated children and adolescents in a household could not be randomly linked to an adult in the household in this and most previous cycles of the survey. The bias due to excluding unassociated children and adolescents was expected to be very small; however, it is not possible to evaluate it.

4.2 Child-first Procedure

In the first two cycles of CHIS, children and adolescents were enumerated and sampled during the adult extended interview. The child and/or adolescent interviews were then conducted following the adult interview. Beginning in 2005, the child and adolescent interviews could be conducted prior to the adult interview under certain conditions. These changes in the order a child and/or adolescent was selected and interviewed are called the “child-first” procedure. This procedure was an operational method (not a sampling method) used to increase the sample yield for child interviews.

In 2001 and 2003, children and adolescents were enumerated and sampled at about the midpoint of the adult interview (section G). If the adult did not complete the extended interview, the child and adolescent could not be interviewed. The child-first procedure was used only when the screener

respondent was the spouse or partner of the sampled adult and there were children in the household associated with the sampled adult. If these conditions were met, a child and/or adolescent could be sampled and the appropriate interview conducted without waiting for the completion of the adult interview. When the child-first criteria were not met, the sampling for children or adolescents was not done until the adult was interviewed. For the first time in 2009, children and adolescents were selected in the cell sample; however, the child-first procedure has not been used for the cell sample, since almost all adults selected in the cell sample are also the screener respondent.

[Table 4-1](#) shows the distribution of completed screener interviews for households with children where children were related to the sampled adult and the number of these households where the child-first procedure was used in the landline sample in CHIS 2011-2012. This table was produced using the variables created during the CATI interview; numbers may not match the final counts of households with children and adolescents shown in other places. In CHIS 2011-2012, there were 5,350 landline sample households reporting children related to the sampled adult in the screener. The child-first procedure was used in 74.6 percent of these households. A child interview was completed in 65.9 percent (2,629 interviews) of such households where the child-first procedure was used. In comparison, a child interview was completed in only 17.9 percent (235 cases) of these households where the procedure was not used. The cases not completing the child interview may have broken off before the end of the screener, not completed the adult interview, or not completed the child interview after the adult was completed.

Although the child-first procedure was intended to increase the number of child interviews, it also had an effect on the number of adolescent interviews. The child-first procedure was used in 72.8 percent (1,454 cases) of the households reporting children related to the sampled adult as well as adolescents in the screener. In 31.8 percent of these households, the adolescent interview was completed. In contrast, in only 6.6 percent of these households where the child-first procedure was not used the adolescent interview was completed. See *CHIS 2011-2012 Methodology Series: Report 2 - Data Collection Methods* for more detail on the child-first procedures and further evaluation of the yields.

Table 4-1. Effect of the child-first procedure on completed child and adolescent interviews in the landline sample*

Households reporting children related to the sampled adult in the screener	Count	Percentage
Total	5,350	100.0
Participated in the child-first procedure	3,990	74.6
<i>Completed the child extended interview</i>	2,629	65.9
<i>Did not complete the child extended interview</i>	1,361	34.1
Did not participate in the child-first procedure	1,360	25.4
<i>Completed the extended interview</i>	235	17.3
<i>Did not complete the extended interview</i>	1,125	82.7
<hr/>		
Households reporting children related to the sampled adult and adolescents in the screener	Count	Percentage
Total	1,997	100.0
Participated in the child-first procedure	1,454	72.8
<i>Completed the adolescent extended interview</i>	463	31.8
<i>Did not complete the adolescent extended interview</i>	991	68.2
Did not participate in the child-first procedure	543	27.2
<i>Completed the adolescent extended interview</i>	36	6.6
<i>Did not complete the adolescent extended interview</i>	507	93.4

* Household counts include landline and surname samples

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

4.3 Adult Sampling

For CHIS, an adult is defined as any person 18 years or older residing in the household. The procedure to select adults in CHIS 2011-2012 for the landline sample and the area sample was the same used since 2003, called the Rizzo method (see Rizzo, Brick, & Park, 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 is 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 $N = 1, 2, 3, \dots$;
- 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; and
 - 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; or
 - 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 is unknown, then the screener respondent is asked to list the adults in the household (including the screener respondent) and the CATI system randomly chooses one of the adults from this roster. No other sampling steps are necessary.

Procedures for the sampling of adults within the cell-only household were developed and implemented in the CHIS 2005 cell-phone pilot and were based on principles similar to those used in landline RDD surveys (Brick, Edwards, & Lee, 2007). The same procedure was used in CHIS 2011-2012. In this approach adults were sampled during the screening interview.

In households with only one adult, no sampling was required. In households sampled from the cell frame with more than one adult, sampling adults depended on whether other adult household members shared the cell phone. If adults shared the cell phone, the same within-household sampling method used in the landline sample was implemented. That is the screener respondent (SR) is randomly selected for the adult interview with a probability equal to the inverse of the number of adults in the household. In case the SR is not selected, then one adult other than the SR is selected for the adult interview using the next birthday method. If the cell phone was not shared, then the SR is sampled.

This sampling scheme assumes that, in cell households with more than one adult, each adult has a cell phone (or shares a different cell phone) if the sampled cell phone is not shared. This weakness in the sampling scheme was recognized during the CHIS cell phone pilot in 2005. However, the alternative

approach required asking the full battery of items to ascertain the cell phone status of each adult in the household, which was viewed as a heavy burden that could detract from gaining cooperation. In other words, while the sampling scheme did not address all possible forms of within-household undercoverage, it was believed this was a good compromise between reducing the potential for increased nonresponse and coverage errors.

4.4 Child Sampling

In an earlier cycle of CHIS, the child sampling procedure was modified to increase the number of interviews for younger children (0 to 5 years old) while reducing the number of interviews for older children (6 to 11 years old). The same sampling procedure was used in all CHIS 2011-2012 samples. Previously all children were sampled at the same rate. If there were only younger or older children in the sampled households, a child was selected with equal probability of selection. In contrast, in households with both younger and older children, children were sampled with differential probabilities of selection. Younger children in such households were assigned a greater probability of selection with respect to the older children. The probability assigned to child i in the household h , p_{hi} , was assigned as

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

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

[Table 4-2](#) shows the number of households with a completed screener interview in which the enumeration and selection of children were completed (either at the end of the extended interview for child-first cases or in section G of the adult extended interview) in CHIS 2011-2012. Children were selected with unequal probability of selection in 26.6 percent of the households with children.

Table 4-2. Distribution of households with children by type of child sampling

Type of child sampling	Type of household	Number of households*	Percentage
Equal probability	Household with children only 0 to 5 years old	8,981	42.9
	Household with children only 6 to 11 years old	6,400	30.5
Unequal probability	Household with children 0 to 5 and 6 to 11 years old	5,569	26.6
Total		20,950	100.0

* Household count include all samples

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

4.5 Adolescent Sampling

The sampling method used in CHIS 2011-2012 to select an adolescent did not change from previous cycles. That is, an adolescent was sampled with equal probability from among all eligible adolescents associated with the sampled adult in a household. In the landline and list samples, adolescents were enumerated and sampled at the end of the screener interview if the child-first procedure was used or in section G of the adult extended interview otherwise. Since adolescents could be sampled and interviewed before the adult interview, there were some households with a completed adolescent interview where adult and/or child interviews were not completed.

Adolescents were also sampled for the first time in the cell phone sample in CHIS 2009. Similar to the child selection, adolescents were selected in section G of the adult extended interview and no child-first procedures were used in the cell phone sample. *CHIS 2011-2012 Methodology Series: Report 5 - Weighting and Estimation* describes how the probabilities of selection are computed for the sampled adults, children, and adolescents in the landline, surname, and cell phone samples.

5. ACHIEVED SAMPLE SIZES

This chapter summarizes the number of completed interviews in CHIS 2011-2012 for the landline cell phone, and surname samples and the relationship between the targeted and the achieved numbers. As mentioned in the previous chapters, the targeted goals for CHIS 2011-2012 were stated in terms of the total number of completed adult interviews. The actual number of completed interviews is a function of the number of telephone numbers sampled, the within-household person sampling, and different reasons for nonresponse. These reasons were discussed in more detail in [Chapter 3](#). Detailed information about the response rates is presented in *CHIS 2011-2012 Methodology Series: Report 4 – Response Rates*.

[Table 5-1](#) shows the number of completed interviews by sample type compared to the adjusted targets. The table shows that, in general, target goals for adult interviews were met in CHIS 2011-2012 at the state level.

Table 5-1. Number of completed interviews by type of sample

Sample type/interview type	Number of completed interviews	Goal	Percent of target completed
Landline sample	40,455	—	—
Adults	32,691	32,578	100.3
Child	5,600	—	—
Adolescent	2,164	—	—
Cell phone sample	11,231	—	—
Adults	9,151	9,457	96.8
Child	1,523	—	—
Adolescent	557	—	—
Surname samples	1,044	—	—
Adults	826	409	202.0
Child	161	—	—
Adolescent	57	—	—
AIAN list sample	338	—	—
Adults	267	300	89.0
Child	50	—	—
Adolescent	21	—	—
Total	53,068	—	—
Adults	42,935	42,335	101.4
Child	7,334	—	—
Adolescent	2,799	—	—

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

[Table 5-2](#) shows the number of completed interviews as percentages of the targeted number of adult interviews at the stratum level. A percentage of 100 or greater indicates the targeted number of adult

interviews was reached in the stratum. The targets were met or surpassed in 21 of the 44 strata based on the self-reported location of the respondent. Ninety five percent of the target or more was met in 38 strata based on the self-reported location.

Table 5-2. Number of completed adult interviews* by self-reported stratum**

Stratum	Landline sample		Cell phone sample		All	
	Completed interviews	% of Target	Completed interviews	% of Target	Completed interviews	% of Target
State	32,691	100.3	9,151	96.8	41,842	99.5
Los Angeles	6,792	104.4	1,962	92.8	8,754	101.6
San Diego	4,070	106.0	880	91.7	4,950	103.1
Orange	1,685	96.3	618	90.0	2,303	94.5
Santa Clara	1,039	95.9	436	90.8	1,475	94.4
San Bernardino	1,042	95.8	353	102.0	1,395	97.3
Riverside	1,327	98.1	342	117.1	1,669	101.5
Alameda	960	96.2	324	90.3	1,284	94.6
Sacramento	938	95.6	220	102.8	1,158	96.9
Contra Costa	685	97.3	188	117.5	873	101.0
Fresno	407	98.8	244	94.2	651	97.0
San Francisco	577	99.7	200	103.1	777	100.5
Ventura	466	98.9	146	109.8	612	101.3
San Mateo	479	100.0	205	71.2	684	89.2
Kern	460	96.0	114	94.2	574	95.7
San Joaquin	340	91.4	135	100.7	475	93.9
Sonoma	356	100.6	145	99.3	501	100.2
Stanislaus	391	99.0	112	106.7	503	100.6
Santa Barbara	411	99.3	97	112.8	508	101.6
Solano	391	99.0	104	99.0	495	99.0
Tulare	378	100.0	117	95.9	495	99.0
Santa Cruz	390	95.8	105	112.9	495	99.0
Marin	435	98.4	57	98.3	492	98.4
San Luis Obispo	408	100.7	102	107.4	510	102.0
Placer	377	98.4	118	100.9	495	99.0
Merced	429	99.1	76	113.4	505	101.0
Butte	364	96.8	123	99.2	487	97.4
Shasta	383	95.8	97	97.0	480	96.0
Yolo	352	100.0	158	106.8	510	102.0
El Dorado	376	98.2	115	98.3	491	98.2
Imperial	458	107.3	76	104.1	534	106.8
Napa	463	100.4	38	97.4	501	100.2
Kings	444	97.4	47	106.8	491	98.2
Madera	449	98.5	50	113.6	499	99.8
Monterey	302	104.1	207	98.6	509	101.8
Humboldt	294	102.4	212	99.5	506	101.2

Table 5-2. Number of completed adult interviews* by self-reported stratum** (continued)

Stratum	Landline sample		Cell phone sample		All	
	Completed interviews	% of Target	Completed interviews	% of Target	Completed interviews	% of Target
Nevada	424	103.4	92	102.2	516	103.2
Mendocino	414	98.6	74	92.5	488	97.6
Sutter	433	100.7	89	127.1	522	104.4
Yuba	424	97.0	53	84.1	477	95.4
Lake	450	99.6	41	85.4	491	98.2
San Benito	465	100.9	48	123.1	513	102.6
Colusa, Glenn, Tehama	342	99.4	66	117.9	408	102.0
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	275	94.5	103	94.5	378	94.5
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	346	100.3	62	112.7	408	102.0

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

** Targets exclude AIAN sample because they do not have defined targets at the stratum level.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

[Table 5-3](#) and [Table 5-4](#) show the number of completed child and adolescent interviews for the landline, cell phone, geographic, and surname samples. Because there were not predetermined targets by stratum for children and adolescents, columns for the percentages of the targeted number of interviews are not included in the table.

Table 5-3. Number of completed child interviews and by self-reported stratum*

Stratum	Landline sample	Cell phone sample	Total
State	5,600	1,523	7,123
Los Angeles	1,187	318	1,505
San Diego	744	147	891
Orange	296	96	392
Santa Clara	213	84	297
San Bernardino	195	71	266
Riverside	229	57	286
Alameda	153	43	196
Sacramento	161	28	189
Contra Costa	107	30	137
Fresno	94	54	148
San Francisco	69	29	98
Ventura	82	20	102
San Mateo	66	25	91
Kern	99	25	124
San Joaquin	64	23	87

Table 5-3. Number of completed child interviews and by self-reported stratum* (continued)

Stratum	Landline sample	Cell phone sample	Total
Sonoma	55	26	81
Stanislaus	74	27	101
Santa Barbara	65	14	79
Solano	53	18	71
Tulare	79	26	105
Santa Cruz	68	9	77
Marin	59	11	70
San Luis Obispo	53	13	66
Placer	46	19	65
Merced	94	16	110
Butte	49	21	70
Shasta	49	19	68
Yolo	64	23	87
El Dorado	58	14	72
Imperial	97	13	110
Napa	59	5	64
Kings	106	9	115
Madera	84	19	103
Monterey	57	37	94
Humboldt	45	29	74
Nevada	62	11	73
Mendocino	49	7	56
Sutter	70	25	95
Yuba	92	7	99
Lake	34	7	41
San Benito	96	9	105
Colusa, Glenn, Tehama	62	18	80
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	26	12	38
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	36	9	45

* Targets exclude AIAN sample because they do not have defined targets at the stratum level.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

Table 5-4. Number of completed adolescent interviews by self-reported stratum*

Stratum	Landline sample	Cell phone sample	Total
State	2,164	557	2,721
Los Angeles	462	118	580
San Diego	269	51	320
Orange	106	29	135
Santa Clara	70	25	95
San Bernardino	89	31	120
Riverside	87	22	109
Alameda	59	11	70
Sacramento	49	13	62
Contra Costa	36	10	46
Fresno	40	20	60
San Francisco	18	9	27
Ventura	32	8	40
San Mateo	23	10	33
Kern	39	7	46
San Joaquin	26	7	33
Sonoma	30	8	38
Stanislaus	31	8	39
Santa Barbara	31	5	36
Solano	23	2	25
Tulare	30	15	45
Santa Cruz	27	2	29
Marin	24	5	29
San Luis Obispo	23	4	27
Placer	19	8	27
Merced	38	6	44
Butte	22	5	27
Shasta	18	10	28
Yolo	22	6	28
El Dorado	22	15	37
Imperial	47	6	53
Napa	25	1	26
Kings	47	5	52
Madera	32	4	36
Monterey	22	14	36
Humboldt	22	11	33
Nevada	25	7	32
Mendocino	19	3	22
Sutter	27	7	34
Yuba	23	1	24
Lake	22	5	27
San Benito	41	2	43
Colusa, Glenn, Tehama	24	9	33
Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	9	8	17
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	14	4	18

* Targets exclude AIAN sample because they do not have defined targets at the stratum level.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

[Table 5-5](#) shows the number of completed adult interviews by ethnicity and sample type. The supplemental sample targets were revised during the data collection period as experience was gained on the actual landline sample yield. The target was exceeded for the number of completed Vietnamese adult interviews but was not met for the Korean interviews.

Table 5-5. Number of completed adult interviews by ethnicity and sample type

Sample	Number of completed interviews		
	Korean only	Vietnamese only	Other*
Landline sample	279	337	32,075
Cell phone sample	94	81	8,976
Korean only list	263	2	2
Korean and other list	79	6	1
Vietnamese only list	0	462	11
AIAN sample	0	0	267
Total	715	888	41,332
Target	500	500	N/A
Percentage of target	143	178	N/A

* Korean or Vietnamese or other group.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

[Table A-3](#), [Table A-4](#), and [Table A-5](#) in the Appendix show additional details of the number of completed interviews by self-reported stratum for the adult, child, and adolescent samples by stratum and sample type.

REFERENCES

- Blumberg, S., & Luke., J. (2013, December). *Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, January–June 2013*. Retrieved Mach 25, 2014, from <http://www.cdc.gov/nchs/nhis/releases.htm#wireless>
- Brick, J., Edwards, W., & Lee, S. (2007). Sampling telephone numbers and adults, interview length, and weighting in the California Health Interview Survey cell phone pilot study. *Public Opinion Quarterly*, 71, 793-813.
- Brick, J., Waksberg, J., Kulp, D., & Starer, A. (1995). Bias in list assisted telephone surveys. *Public Opinion Quarterly*, 59(2), pp. 218-235.
- Delnevo, C., D.A, G., & Hagman, B. (2008). Declining estimated prevalence of alcohol drinking and smoking among young adults nationally: artifacts of sample undercoverage? *167*(1).
- Edwards, W., Brick, J., Flores Cervantes, I., DiSogra, C., & Yen, W. (2002). Sampling race and ethnic groups in RDD surveys. *Proceedings of the Survey Methods Section of the American Statistical Association*.
- Fahimi, M., Kulp, D., & Brick, J. M. (2008, September 28). Bias in List-Assisted 100-Series RDD Sampling. *Survey Practice*.
- Flores Cervantes, I., & Kalton, G. (2007). Methods for Sampling Rare Populations in Telephone Surveys. In J. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japac, P. J. Lavrakas, et al. (Eds.), *Advances in Telephone Survey Methodology*. New York: John Wiley & Sons.
- Lee, S., Nguyen, H., Jawad, M., & Kurata, J. (2008). Linguistic Minorities and Nonresponse Error. *Public Opinion Quarterly*, 72., 72, 470-486.
- Rizzo, L., Brick, J., & Park, I. (2004). A Minimally Intrusive Method for Sampling Persons in Random Digit Dialing Surveys. , , 267-274. *Public Opinion Quarterly*, 68.
- Tucker, C., Lepkowski, J., & Piekarski, L. (2002). The current efficiency of list assisted telephone sampling designs. *Public Opinion Quarterly*(66), pp. 321-338.

APPENDIX A

Table A-1. Stratum definitions for CHIS 2001, 2003, 2005, 2007, 2009, and 2011-2012

County	2005 , 2007, 2009, and 2011-2012 Stratum	2001 and 2003 Stratum
Los Angeles	1	1
San Diego	2	2
Orange	3	3
Santa Clara	4	4
San Bernardino	5	5
Riverside	6	6
Alameda	7	7
Sacramento	8	8
Contra Costa	9	9
Fresno	10	10
San Francisco	11	11
Ventura	12	12
San Mateo	13	13
Kern	14	14
San Joaquin	15	15
Sonoma	16	16
Stanislaus	17	17
Santa Barbara	18	18
Solano	19	19
Tulare	20	20
Santa Cruz	21	21
Marin	22	22
San Luis Obispo	23	23
Placer	24	24
Merced	25	25
Butte	26	26
Shasta	27	27
Yolo	28	28
El Dorado	29	29
Imperial	30	30
Napa	31	31
Kings	32	32
Madera	33	33
Monterey	34	34
San Benito	41	
Lake	40	37
Mendocino	37	
Sutter	38	39
Yuba	39	

Table A-1. Stratum definitions for CHIS 2001, 2003, 2005, 2007, 2009, and 2011-2012 (continued)

County	2005 , 2007, 2009, and 2011-2012 Stratum	2001 and 2003 Stratum
Colusa Glen Tehama	42	38
Humboldt Del Norte	35	35
Lassen Modoc Siskiyou Trinity	43	36
Plumas Sierra Nevada	36	40
Alpine Amador Calaveras Inyo Mariposa Mono Tuolumne	44	41

Table A-2. Number of telephone numbers and addresses drawn by sample frame and sampling stratum

Stratum	Landline	Cell phone	Surname and list*			AIAN	Total
			Korean only	Korean and other	Vietnamese only		
State	764,922	145,287	5,156	5,527	5,594	2,253	928,739
1 Los Angeles	208,824	32,890	1,936	1,823	1,034	224	246,731
2 San Diego	98,637	8,039	229	161	383	420	107,869
3 Orange	52,057	9,890	706	463	1,400	39	64,555
4 Santa Clara	28,628	5,895	358	600	1,040	0	36,521
5 San Bernardino	24,699	5,915	163	115	126	7	31,025
6 Riverside	28,054	6,013	152	90	139	19	34,467
7 Alameda	23,862	4,723	307	594	363	4	29,853
8 Sacramento	21,881	3,529	146	163	261	196	26,176
9 Contra Costa	14,931	3,237	126	167	91	1	18,553
10 Fresno	8,976	7,235	76	55	54	172	16,568
11 San Francisco	20,979	2,359	247	719	213	1	24,518
12 Ventura	10,330	2,925	56	37	42	24	13,414
13 San Mateo	14,399	4,280	141	250	52	0	19,122
14 Kern	9,037	1,640	32	13	21	88	10,831
15 San Joaquin	7,058	5,302	49	39	84	15	12,547
16 Sonoma	5,747	4,211	43	17	31	8	10,057
17 Stanislaus	7,214	37	28	14	19	9	7,321
18 Santa Barbara	8,388	1,501	24	16	16	51	9,996
19 Solano	7,926	62	23	15	24	11	8,061
20 Tulare	6,447	38	12	8	7	8	6,520
21 Santa Cruz	7,483	1,140	19	12	8	2	8,664
22 Marin	9,493	1,415	30	23	26	0	10,987
23 San Luis Obispo	6,243	1,403	17	6	12	2	7,683
24 Placer	6,997	267	26	20	17	47	7,374
25 Merced	7,519	25	13	10	8	5	7,580
26 Butte	4,637	2,128	16	6	14	154	6,955
27 Shasta	4,898	1,426	15	2	7	70	6,418
28 Yolo	6,041	9,638	18	24	13	54	15,788
29 El Dorado	5,737	56	13	8	10	5	5,829
30 Imperial	7,170	7,251	5	3	2	0	14,431
31 Napa	8,355	36	9	3	4	3	8,410
32 Kings	7,239	105	8	2	5	21	7,380
33 Madera	6,320	127	5	3	3	34	6,492
34 Monterey	6,651	3,511	35	19	21	1	10,238
35 Humboldt	4,338	3,759	9	3	2	32	8,143
36 Nevada	5,743	111	12	4	5	22	5,897
37 Mendocino	4,985	409	5	4	4	159	5,566
38 Sutter	6,706	120	4	2	2	35	6,869

Table A-2. Number of telephone numbers and addresses drawn by sample frame and sampling stratum (continued)

Stratum	Landline	Cell phone	Surname and list*			AIAN	Total
			Korean only	Korean and other	Vietnamese only		
39 Yuba	8,678	167	8	4	4	50	8,911
40 Lake	7,433	210	7	1	2	76	7,729
41 San Benito	9,507	28	2	0	1	0	9,538
42 Colusa, Glenn, Tehama	4,414	248	5	2	4	88	4,761
43 Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	4,475	1,659	6	1	7	90	6,238
44 Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	5,786	327	15	6	13	6	6,153

* Not drawn by sampling stratum.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

Table A-3. Number of adult completed interviews by sample type and self-reported stratum

Stratum	Landline	Cell phone	Surname and list*			AIAN	Total
			Korean only	Korean and other	Vietnamese only		
State	32,691	9,151	267	86	473	267	42,935
1 Los Angeles	6,792	1,962	139	36	67	13	9,009
2 San Diego	4,070	880	10	2	40	21	5,023
3 Orange	1,685	618	46	15	132	4	2,500
4 Santa Clara	1,039	436	16	6	127	0	1,624
5 San Bernardino	1,042	353	6	5	11	2	1,419
6 Riverside	1,327	342	4	2	6	4	1,685
7 Alameda	960	324	9	4	27	0	1,324
8 Sacramento	938	220	6	1	20	26	1,211
9 Contra Costa	685	188	3	2	7	0	885
10 Fresno	407	244	1	0	1	29	682
11 San Francisco	577	200	9	3	11	0	800
12 Ventura	466	146	4	3	3	0	622
13 San Mateo	479	205	1	2	2	0	689
14 Kern	460	114	2	0	1	7	584
15 San Joaquin	340	135	0	1	11	4	491
16 Sonoma	356	145	1	0	2	1	505
17 Stanislaus	391	112	0	0	0	0	503
18 Santa Barbara	411	97	0	2	1	1	512
19 Solano	391	104	0	0	1	3	499
20 Tulare	378	117	0	0	0	0	495
21 Santa Cruz	390	105	0	1	0	0	496
22 Marin	435	57	1	0	0	0	493
23 San Luis Obispo	408	102	0	0	0	0	510
24 Placer	377	118	3	0	0	13	511
25 Merced	429	76	0	0	0	1	506
26 Butte	364	123	0	0	0	16	503
27 Shasta	383	97	0	0	0	7	487
28 Yolo	352	158	1	1	0	2	514
29 El Dorado	376	115	2	0	0	2	495
30 Imperial	458	76	0	0	0	0	534
31 Napa	463	38	0	0	1	1	503
32 Kings	444	47	0	0	0	1	492
33 Madera	449	50	0	0	0	11	510
34 Monterey	302	207	2	0	1	0	512
35 Humboldt	294	212	0	0	0	26	532
36 Nevada	424	92	0	0	1	8	525
37 Mendocino	414	74	0	0	0	14	502
38 Sutter	433	89	0	0	0	7	529
39 Yuba	424	53	0	0	0	7	484

Table A-3. Number of adult completed interviews by sample type and self-reported stratum(continued)

Stratum	Landline	Cell phone	Surname and list*				Total
			Korean only	Korean and other	Vietnamese only	AIAN	
40 Lake	450	41	0	0	0	5	496
41 San Benito	465	48	0	0	0	0	513
42 Colusa, Glenn, Tehama	342	66	1	0	0	7	416
43 Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	275	103	0	0	0	22	400
44 Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	346	62	0	0	0	2	410

* Not drawn by sampling stratum.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

Table A-4. Number of child completed interviews by self-reported stratum

Stratum	Landline	Cell phone	Surname and list*			AIAN	Total
			Korean only	Korean and other	Vietnamese only		
State	5,600	1,523	36	17	108	50	7,334
1 Los Angeles	1,187	318	15	6	14	1	1,541
2 San Diego	744	147	1	0	11	1	904
3 Orange	296	96	7	5	23	0	427
4 Santa Clara	213	84	4	1	28	0	330
5 San Bernardino	195	71	0	0	3	0	269
6 Riverside	229	57	0	0	0	1	287
7 Alameda	153	43	3	1	9	0	209
8 Sacramento	161	28	0	0	3	8	200
9 Contra Costa	107	30	0	0	4	0	141
10 Fresno	94	54	0	0	2	7	157
11 San Francisco	69	29	2	0	2	0	102
12 Ventura	82	20	2	2	2	0	108
13 San Mateo	66	25	0	1	1	0	93
14 Kern	99	25	0	0	0	1	125
15 San Joaquin	64	23	0	0	3	0	90
16 Sonoma	55	26	0	0	2	0	83
17 Stanislaus	74	27	0	0	0	0	101
18 Santa Barbara	65	14	1	0	0	0	80
19 Solano	53	18	0	0	0	1	72
20 Tulare	79	26	0	0	0	0	105
21 Santa Cruz	68	9	0	0	0	0	77
22 Marin	59	11	0	0	0	0	70
23 San Luis Obispo	53	13	0	0	0	0	66
24 Placer	46	19	0	0	0	3	68
25 Merced	94	16	0	0	0	0	110
26 Butte	49	21	0	0	0	3	73
27 Shasta	49	19	0	0	0	0	68
28 Yolo	64	23	0	1	0	0	88
29 El Dorado	58	14	1	0	0	0	73
30 Imperial	97	13	0	0	0	0	110
31 Napa	59	5	0	0	0	0	64
32 Kings	106	9	0	0	0	1	116
33 Madera	84	19	0	0	0	3	106
34 Monterey	57	37	0	0	1	0	95
35 Humboldt	45	29	0	0	0	7	81
36 Nevada	62	11	0	0	0	1	74
37 Mendocino	49	7	0	0	0	3	59
38 Sutter	70	25	0	0	0	1	96
39 Yuba	92	7	0	0	0	1	100

Table A-4. Number of child completed interviews by self-reported stratum (continued)

Stratum	Landline	Cell phone	Surname and list*				AIAN	Total
			Korean only	Korean and other	Vietnamese only			
40 Lake	34	7	0	0	0	1	42	
41 San Benito	96	9	0	0	0	0	105	
42 Colusa, Glenn, Tehama	62	18	0	0	0	3	83	
43 Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	26	12	0	0	0	3	41	
44 Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	36	9	0	0	0	0	45	

* Not drawn by sampling stratum.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.

Table A-5. Number of adolescent completed interviews by self-reported stratum

Stratum	Landline	Cell phone	Surname and list*			AIAN	Total
			Korean only	Korean and other	Vietnamese only		
State	2,164	557	16	5	36	21	2,799
1 Los Angeles	462	118	5	4	2	1	592
2 San Diego	269	51	2	0	1	0	323
3 Orange	106	29	5	1	11	1	153
4 Santa Clara	70	25	1	0	9	0	105
5 San Bernardino	89	31	0	0	2	0	122
6 Riverside	87	22	0	0	1	0	110
7 Alameda	59	11	2	0	3	0	75
8 Sacramento	49	13	0	0	2	4	68
9 Contra Costa	36	10	0	0	0	0	46
10 Fresno	40	20	0	0	0	2	62
11 San Francisco	18	9	0	0	0	0	27
12 Ventura	32	8	0	0	1	0	41
13 San Mateo	23	10	0	0	0	0	33
14 Kern	39	7	0	0	0	2	48
15 San Joaquin	26	7	0	0	2	0	35
16 Sonoma	30	8	0	0	2	0	40
17 Stanislaus	31	8	0	0	0	0	39
18 Santa Barbara	31	5	0	0	0	0	36
19 Solano	23	2	0	0	0	0	25
20 Tulare	30	15	0	0	0	0	45
21 Santa Cruz	27	2	0	0	0	0	29
22 Marin	24	5	0	0	0	0	29
23 San Luis Obispo	23	4	0	0	0	0	27
24 Placer	19	8	0	0	0	1	28
25 Merced	38	6	0	0	0	0	44
26 Butte	22	5	0	0	0	0	27
27 Shasta	18	10	0	0	0	0	28
28 Yolo	22	6	1	0	0	0	29
29 El Dorado	22	15	0	0	0	0	37
30 Imperial	47	6	0	0	0	0	53
31 Napa	25	1	0	0	0	0	26
32 Kings	47	5	0	0	0	0	52
33 Madera	32	4	0	0	0	3	39
34 Monterey	22	14	0	0	0	0	36
35 Humboldt	22	11	0	0	0	2	35
36 Nevada	25	7	0	0	0	1	33
37 Mendocino	19	3	0	0	0	0	22
38 Sutter	27	7	0	0	0	2	36
39 Yuba	23	1	0	0	0	0	24

Table A-5. Number of adolescent completed interviews by self-reported stratum (continued)

Stratum	Landline	Cell phone	Surname and list*				Total
			Korean only	Korean and other	Vietnamese only	AIAN	
40 Lake	22	5	0	0	0	0	27
41 San Benito	41	2	0	0	0	0	43
42 Colusa, Glenn, Tehama	24	9	0	0	0	2	35
43 Del Norte, Lassen, Modoc, Plumas, Sierra, Siskiyou, Trinity	9	8	0	0	0	0	17
44 Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	14	4	0	0	0	0	18

* Not drawn by sampling stratum.

Source: UCLA Center for Health Policy Research, 2011-2012 California Health Interview Survey.