



CHIS 2015-2016 SAMPLE DESIGN

Overview

The California Health Interview Survey (CHIS) began in 2001 as a biennial population-based, omnibus health survey of California. It remains the largest telephone survey in California and the largest continuous state health survey in the country. CHIS 2015-2016 is the eighth data collection cycle following CHIS 2001, 2003, 2005, 2007, 2009, 2011-2012, and 2013-2014. The 2015-2016 data files for individual years are being released separately as they become available.

CHIS became a continuous survey in 2011, with ongoing data collection throughout a two-year cycle. The change to a continuous survey enhances the role of CHIS as a critical and timely source of public health surveillance information for California. While this change significantly streamlines CHIS operations and data production, most aspects of the CHIS project, including the overall sample design, remain the same.

CHIS randomly selects one adult to interview in each randomly sampled and participating household throughout California. In addition to adults, interviews are conducted about the health of children (ages 0 to 11, by proxy with an adult sufficiently knowledgeable about the health of the selected child) and adolescents (age 12 to 17, directly interviewed following parental permission). The selected adult respondent must be the parent or legal guardian of children and/or adolescents living in the household in order for children and/or adolescents to be eligible for participation.

Major content areas for the 2015-2016 survey include health status and conditions, diet, physical activity, other health-related behaviors, health insurance coverage, mental health, and access and utilization of health care and mental health services. To capture the rich diversity of the California population, interviews are conducted in six languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, Korean, and Tagalog.

In addition to the landline RDD sample, CHIS 2015-2016 includes a statewide cell phone sample, as has each previous cycle since 2007. The most notable change from prior cycles is the increase in the cell phone sample goals to 50% of obtained interviews. More information about these samples is provided below in *The CHIS 2015-2016 Sample in Detail*.

Overall, the CHIS sample is designed to provide population-based estimates for most California counties and all major ethnic groups, including several ethnic subgroups. The sample is designed to meet and optimize two goals:

- To provide local-level estimates for counties with populations of 60,000 or more for local planning and comparisons among counties;
- To provide statewide estimates for California's overall population, its major race/ethnic groups, as well as for several Asian and Latino ethnic groups.

CHIS data and statistical estimates are publicly available to local, state and federal agencies, community-based organizations, health care providers and organizations, advocacy groups, policy makers, and individual citizens. CHIS results and data are available through multiple dissemination channels:

- CHIS website with a free, easy-to-use interactive query system, *AskCHIS*, for tailored data estimates (<http://ask.chis.ucla.edu/>)
- Public-use data files for researchers and analysts (<http://healthpolicy.ucla.edu/chis/data/Pages/public-use-data.aspx>)
- Published reports
- Scholarly publications
- The UCLA Center for Health Policy Research Data Access Center (for more-detailed data files available only through a secure environment, see: <http://healthpolicy.ucla.edu/chis/data/Pages/confidential.aspx>)

CHIS provides technical assistance for researchers and all users of web-based and public-use data (Email: dacchpr@ucla.edu, Phone: 310-794-8319).

The CHIS 2015-2016 Sample in Detail

CHIS 2015-2016 includes 2 separate samples, both administered through a computer assisted telephone interview (CATI) system:

1. a statewide landline random digit dial (RDD) sample supplemented by surname-listed sample, and
2. a statewide cell phone sample.

Landline RDD Sample

The first CHIS sample is a landline RDD sample. This sample is supplemented with a small surname list sample providing increased representation of Vietnamese, Korean, and Japanese Californians. The landline RDD sample is geographically stratified by county, groups of small counties, and some sub-county areas (see Table 2 at the end of this document). Only residential households with landline telephones are included in the landline RDD sample. The sample weights that accompany the data files statistically adjust final CHIS estimates to compensate for households without landline telephones.

Cell Phone Sample

The number of households that do not have landline telephones has grown rapidly in recent years; nearly all of this change is due to the growth of households with cell phones as their only telephone. There is considerable evidence that households that rely exclusively on cell phones are systematically different than households with landline telephones.¹ In order to avoid the bias inherent in excluding this growing population segment from CHIS, a separate RDD cell phone sample is conducted statewide. The sample is drawn using RDD methodology from banks of telephone numbers that are assigned for cellular use. Consistent with the methodology employed since CHIS 2009 the sample includes all cell phone users

¹ <https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201605.pdf>

(including those who have a landline telephone and those who do not). This method contrasts with the CHIS 2007 cell phone sample in which sampled cell phone numbers were called and screened to exclude with a landline telephone. More detail on the CHIS 2007 and 2009 cell phone sample methodologies are available online at <http://healthpolicy.ucla.edu/chis/design/Pages/methodology.aspx>.

The current sample design includes 50% landline phone numbers and 50% cell phone numbers. As the proportion of households relying mostly or exclusively on cell phones as their primary means of telephone communication continues to grow, CHIS will continue to monitor industry best-practices and adjust the proportion of the total sample that comes from cell phones.

In addition to interviews with adults, CHIS has conducted child and adolescent interviews as part of the cell phone sample since CHIS 2009 in addition to the landline sample. This contrasts with CHIS 2007, in which children and teens were enumerated as part of the CHIS screener interview, but child and teen interviews were not conducted as part of the cell phone sample.

Sample Size by Age Category

As many as three persons will be sampled from each household, one from each of three age categories, adults (age 18 and older), adolescents (age 12-17), and children (age 0-11). Adults and adolescents answer survey questions for themselves, and child data are collected by proxy interview with an adult who is sufficiently knowledgeable about that child.

Sample Size by County

The CHIS sample is stratified to provide local-level samples for most counties. California's 58 counties are arranged into sampling strata based on size and analytic goals to produce 56 sampling strata. Table 1 below describes the overall stratification approach for the 58 counties, and Table 2 at the end of this document provides target CHIS sample sizes for the 44 county and county group strata for both the landline and cell samples.

Table 1. CHIS Sampling Strata Type and Number

Stratum Type	Number of Counties and Strata
Counties with sub-county strata	2 counties, 14 strata total between them
Counties as single stratum	39 counties, 39 strata
Counties combined into single stratum	17 counties, 3 strata total
Total	58 counties, 56 strata

Los Angeles and San Diego counties have 8 and 6 sub-county strata respectively. Of the remaining 56 counties, 39 comprise single-county strata and the remaining 17 counties are arranged into three multiple-county strata. The CHIS sample is allocated to achieve the project's goals of providing estimates for as many counties as possible and to provide estimates for major race/ethnic groups and subgroups. The CHIS 2015-2016 sample, as in previous CHIS cycles, includes a minimum sample size of at least 400 households for the smallest counties and county groups; the largest county will have a sample of nearly 8,000 households (see Table 2).

The stratum sample goals in Table 2 are calculated by allocating the state target using population estimates and projections from the California Department of Finance (DoF). In this allocation, a minimum sample size is maintained to ensure sufficient sample for analysis of smaller counties. The figures in Table 2 reflect estimated sample targets based on information available at the time of this report and may be updated in consultation with the CHIS data collection subcontractor as DoF information is updated. Agencies and organizations may partner with CHIS to increase the sample size or add county-specific content in their county through supplemental funding. Such oversamples for CHIS 2015-2016 are discussed in the methodology reports available online at <http://healthpolicy.ucla.edu/chis/design/Pages/methodology.aspx>.

Samples of Race/Ethnic Groups

One of CHIS's primary sampling goals is to produce data that reflect California's ethnic and racial diversity. The CHIS sample is designed to include sufficient numbers of Whites, Latinos, Asians, Native Hawaiians and other Pacific Islanders, African Americans, and American Indians/Alaska Natives in such a way that they reflect their shares of the state's population. The CHIS sample includes a substantial number of responses from people in Latino and Asian ethnic groups because of their large population sizes in California. CHIS oversamples Vietnamese and Koreans so that each group's total sample of adults reaches 500 in order to have a sufficient number of two key Asian ethnic groups. CHIS has been oversampling Vietnamese and Koreans since its first survey in 2001. There is no need to oversample Chinese, Filipinos, and South Asians, because the CHIS sample produces a sufficient number of these groups for statistical analysis. Oversamples of American Indians/Alaska Natives have been conducted in CHIS 2001 and CHIS 2011-2012.

Table 2. CHIS 2015-2016 Initial RDD Sample Size Targets for Counties and County Groups

Stratum	Sample Type		
	Target Adult Sample	Landline (+/- 5%)	Cell (+/- 10%)
Los Angeles	7,721	3,860	3,860
San Diego	3,140	1,570	1,570
Orange	2,228	1,114	1,114
Santa Clara	1,521	760	760
San Bernardino	1,323	661	661
Riverside	2,061	1,030	1,030
Alameda	1,209	605	605
Sacramento	1,389	694	694
Contra Costa	1,054	527	527
Fresno	898	449	449
San Francisco	749	375	375
Ventura	571	286	286
San Mateo	703	351	351
Kern	735	368	368
San Joaquin	500	250	250
Sonoma	500	250	250
Stanislaus	500	250	250
Santa Barbara	500	250	250
Solano	500	250	250
Tulare	500	250	250
Santa Cruz	500	250	250
Marin	500	250	250
San Luis Obispo	500	250	250
Placer	500	250	250
Merced	500	250	250
Butte	500	250	250
Shasta	500	250	250
Yolo	500	250	250
El Dorado	500	250	250
Imperial	500	250	250
Napa	500	250	250
Kings	500	250	250
Madera	500	250	250
Monterey	500	250	250
Humboldt	500	250	250
Nevada	500	250	250
Mendocino	500	250	250
Sutter	500	250	250
Yuba	500	250	250
Lake	500	250	250
San Benito	500	250	250
Tehama-Glenn-Colusa	400	200	200
Del Norte-Siskiyou-Lassen-Trinity-Modoc-Plumas-Sierra	400	200	200
Tuolumne-Calaveras-Amador-Inyo-Mariposa-Mono-Alpine	400	200	200
TOTAL	40,000	20,000	20,000