CHIS 2013-2014 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 state health survey in the country. CHIS 2013-2014 is the seventh data collection cycle following CHIS 2001, 2003, 2005, 2007, 2009, and 2011-2012.

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 the adult most 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 2013-2014 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 five languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, and Korean. We plan to add Tagalog in CHIS 2013-2014 to further enhance the survey's representativeness.

In addition to the main landline RDD sample, CHIS 2013-2014 will include a statewide cell phone sample, as has each previous cycle since 2007. More information about these samples is provided below in *The CHIS 2013-2014 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:

1. To provide local-level estimates for counties with populations of 60,000 or more for local planning and comparisons among counties;

2. 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, and

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, *Ask*CHIS, for tailored data estimates (http://ask.chis.ucla.edu/main/default.asp)

- 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 2013-2014 Sample in Detail

CHIS 2013-2014 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 primary CHIS sample is a landline RDD sample. This sample is supplemented with a small surname list sample providing increased representation of Vietnamese and Korean Californians. The landline RDD sample is geographically stratified by county, groups of small counties, and some sub-county areas (see *Sample Size by County* in 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. While the landline RDD sample is stratified by county, country group, and sub-county areas, the cell phone RDD sample does not have sub-county stratification (i.e., it is only stratified by county or group of counties). The sample is drawn using RDD methodology from banks of telephone numbers that are assigned for cellular use. Consistent with

---

[1] Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, January–June 2012. National Center for Health Statistics. December 2012. Available from: http://www.cdc.gov/nchs/nhis.htm.

the methodology employed in CHIS 2009 and CHIS 2011-2012, CHIS 2013-2014 will include all cell phone users (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 is 80% landline sample and 20% cell phone sample (with a goal of 15% to 25% of final interviews from the cell phone sample in each stratum). 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 2013-2014 will conduct child and adolescent interviews as part of the cell phone sample as well as 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.

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 the adult who is most 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 2013-2014 sample, as in previous CHIS cycles, will include 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 10,000 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), which are available online.[2] 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.

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 included sufficient numbers of Whites, Latinos, Asians, Pacific Islanders, African Americans, and American Indians/Alaska Natives in such a way that 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, Japanese, and South Asians, because the CHIS sample produces a sufficient number of these groups for statistical analysis.

---

[2] The CHIS 2013-2014 sample design is based on DoF Table 1: E-5 County/State Population and Housing Estimates, 1/1/2012 and is available online at http://www.dof.ca.gov/research/demographic/reports/estimates/e-5/2011-20/view.php

Table 2. CHIS 2013-2014 RDD Sample Size Targets for Counties

Stratum	Total Adult Sample	Landline	Cell (+/-5%)
Los Angeles	8,113	6,491	1,622
San Diego	2,727	2,182	545
Orange	2,487	1,990	497
Riverside	1,729	1,383	346
San Bernardino	1,535	1,228	307
Santa Clara	1,520	1,216	304
Alameda	1,364	1,091	273
Sacramento	1,289	1,031	258
Contra Costa	943	754	189
San Francisco	865	692	173
Fresno	730	584	146
Ventura	669	535	134
San Mateo	646	517	129
Kern	643	514	129
San Joaquin	540	432	108
Butte	500	400	100
El Dorado	500	400	100
Humboldt	500	400	100
Imperial	500	400	100
Kings	500	400	100
Lake	500	400	100
Madera	500	400	100
Marin	500	400	100
Mendocino	500	400	100
Merced	500	400	100
Monterey	500	400	100
Napa	500	400	100
Nevada	500	400	100
Placer	500	400	100
San Benito	500	400	100
San Luis Obispo	500	400	100
Santa Barbara	500	400	100
Santa Cruz	500	400	100
Shasta	500	400	100
Solano	500	400	100
Sonoma	500	400	100
Stanislaus	500	400	100
Sutter	500	400	100
Tulare	500	400	100
Yolo	500	400	100
Yuba	500	400	100
Del Norte-Siskiyou-Lassen-Trinity-Modoc-Plumas-Sierra	400	320	80
Tehama-Glenn-Colusa	400	320	80
Tuolumne-Calaveras-Amador-Inyo-Mariposa-Mono-Alpine	400	320	80
TOTAL	40,000	32,000	8,000