

## SAMPLE DESIGN DESCRIPTION FOR CHIS 2001 (ORIGINAL PLAN)

### Overview

The California Health Interview Survey (CHIS) is the largest telephone survey in California and the largest health survey of its kind in the country. CHIS interviews one randomly selected adult in each of the 55,000 households sampled in California. In addition, in those households with adolescents (ages 12-17) and/or children (under age 12), one adolescent and one parent proxy of the child is interviewed. This results in approximately 6,000 adolescent interviews and information on approximately 13,000 young children throughout the state. Major content areas for the survey include health-related behaviors, health status and conditions, health insurance coverage, and access to health care 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 Khmer (Cambodian).

The 2001 California Health Interview Survey (CHIS 2001) sample is designed to provide population-based estimates for most California counties, all major ethnic groups, and several ethnic subgroups. The sample is also designed to meet and optimize two goals:

1. To provide local-level estimates for counties with populations of 100,000 or more — generating important information for local planning and enabling comparisons among counties;
2. To provide statewide estimates for California's overall population, its major ethnic groups, as well as for some ethnic subgroups — providing important information that has never before been available.

The sample design is created to allow both easy access to and use of CHIS data by organizations and individuals. The collected information and results is available to state and local health agencies, community-based organizations, health care providers and organizations, advocacy groups, and policy makers. CHIS results and data are available in several ways:

- CHIS web site with an interactive query system (*AskCHIS*) for tailored data estimates
- Public-use data files for researchers and analysts
- Published reports
- Scholarly publications
- The UCLA Center for Health Policy Research Data Access Center with special data files and technical assistance for researchers and technical assistance for all users of web-based and public-use data.

## **The CHIS 2001 Sample in Detail**

The CHIS sample of adults, adolescents and children will consist of 55,000 households, most (approximately 52,900) selected through a random-digit dial (RDD) telephone survey. Only residential households with telephones will be in the CHIS sample. The results, however, will be statistically adjusted to compensate for households without telephones.

CHIS 2001 also incorporates some very specific over-samples. Three city health departments have their cities over-sampled to reach a target RDD sample size of 800 per city. Also, because Asian and Pacific Islander (AAPI) subgroups are different from each other in culture and language, an additional 2,100 households are surveyed to over-sample five specific Asian subgroups. Finally, in addition to the 55,000 households, American Indian/Alaska Natives are over-sampled to raise their total sample size to 800 in order to examine important differences between urban and rural areas. This is all explained in more detail below and shown in the accompanying tables.

### ***Sample Size by Age Category***

From each household, as many as three persons, one from each of three age categories are selected to be interviewed. One adult (ages 18 and older) from each household is part of the adult sample. In households with adolescents (ages 12-17) residing at home, one adolescent is interviewed. In homes with children under age twelve, the adult who is most knowledgeable about a selected child is interviewed. Thus, CHIS 2001 has three distinct samples: adult, adolescent, and child. The estimated number of interviews for each sample is 55,000 adults, 5,000-6,000 adolescents (parental consent and adolescent ascent are required), and 14,000 children by adult proxy.

### ***Sample Size by County***

All counties in California are sampled. The 58 California counties are arranged into 41 strata. Thirty-three of the 35 counties with a population of 100,000 or more form their own sample stratum. Two (Monterey and Humboldt) are combined with smaller adjoining counties (San Benito and Del Norte, respectively). The 23 remaining counties with populations smaller than 100,000 are grouped into six sample strata in a way that is meaningful for planning purposes (see Table 1). The minimum sample size for any stratum is 800. (Sample sizes for any county or stratum can be increased with additional funding from the requesting counties.)

The sample for Los Angeles County (see Table 2) is proportionally distributed across the county's eight Service Planning Areas (SPAs).

CHIS can provide estimates for all counties and groups of counties with a population of 100,000 or more. However, in public-use data files only statewide data (no strata identifiers) are available for analysis. This restriction in the public-use data file is required in order to prevent identification of individual respondents, which would be a breach of confidentiality.

Reported estimates for medium-size and larger counties can include breakdowns by age groups and ethnic groups, depending on sample size. The CHIS design has the flexibility to expand sample sizes in any stratum to improve the precision for a stratum or county.

## ***Sample Sizes for Three City Health Departments***

The cities of Long Beach, Pasadena, and Berkeley are over-sampled in CHIS 2001 using geographically targeted RDD techniques to reach a sample size of 800 per city (see Table 3).

### ***Sample Size by Ethnicity***

The following is an approximation of the number of interviews in each major ethnic group in the random digit dial (RDD) sample:

- 34,000 non-Latino whites
- 10,500 Latinos
- 4,400 Asian American/Pacific Islanders
- 3,400 African Americans
- 440 American Indian/Alaska Native

These numbers are augmented by over-samples (see Table 4).

### ***Over-sampling Asian American Subgroups***

Approximately 2,100 additional Asian Americans are obtained in CHIS 2001 through an over-sample (based on listed surname) to achieve an ideal target sample size for each group. The over-sample number per Asian subgroup is:

- 300 Japanese to reach a target of 800
- 300 Vietnamese to reach a target of 800
- 500 Asian Indians to reach a target of 800
- 500 Koreans to reach a target of 800
- 100 Cambodians to reach a target of 200

A sufficient number of Chinese and Filipino households are interviewed through the RDD sample, thus avoiding the need for over-sampling these two groups (see Table 5).

### ***Over-sampling American Indians***

CHIS obtains interviews from 444 American Indian/Alaska Native (AIAN) households through the RDD sample (the actual number may be more). These are distributed approximately 60% in rural areas and 40% in urban areas using Indian Health Service (IHS) designated locations. The IHS is augmenting CHIS 2001 funding to obtain an over-sample in both the urban and rural areas to achieve a total sample size of 400 per area (see Table 6). Listed sample, mostly obtained from IHS is the source of these additional interviews.

**Table 1. CHIS 2001 Sample for Counties**

	<b>Geographic Strata</b>	<b>Sample Size</b>	<b>Population Size</b>
1	Los Angeles	11,292	Over 9,000,000
2	San Diego	2,660	1,200,000 or greater
3	Orange	2,540	
4	Santa Clara	1,554	
5	San Bernardino	1,527	
6	Riverside	1,359	
7	Alameda	1,232	
8	Sacramento	1,200	800,000-1,200,000
9	Contra Costa	1,200	
10	Fresno	1,000	500,000-800,000
11	San Francisco	1,000	
12	Ventura	1,000	
13	San Mateo	1,000	
14	Kern	1,000	
15	San Joaquin	1,000	
16	Sonoma	800	100,000-500,000
17	Stanislaus	800	
18	Santa Barbara	800	
19	Solano	800	
20	Tulare	800	
21	Santa Cruz	800	
22	Marin	800	
23	San Luis Obispo	800	
24	Placer	800	
25	Merced	800	
26	Butte	800	
27	Shasta	800	
28	Yolo	800	
29	El Dorado	800	
30	Imperial	800	
31	Napa	800	
32	Kings	800	
33	Madera	800	
34	Monterey (pop. >100,000) San Benito (pop. <100,000)	800	Medium and small counties combined
35	Humboldt (pop. >100,000) Del Norte (pop. <100,000)	800	
36	Siskiyou Trinity Lassen Modoc	800	Less than 100,000 population per county
37	Mendocino Lake	800	
38	Tehama Colusa Glenn	800	
39	Sutter Yuba	800	
40	Nevada Sierra Plumas	800	
41	Tuolumne Mariposa Calaveras Mono Amador Alpine Inyo	800	
<b>Total of 41 Strata</b>		<b>51,364</b>	

**Table 2. Sample Sizes in Los Angeles County Service Planning Areas (SPAs)**

Service Planning Area	Sample Size
San Fernando	2,337
San Gabriel	2,147
South Bay	1,840
East	1,433
Metro	1,244
South	1,054
West	833
Antelope Valley	404
<b>Los Angeles Total</b>	<b>11,292</b>

**Table 3. Samples for Cities with Local Health Departments**

City	RDD Sample	Over-sample	Total
Long Beach	582	218	800
Pasadena	178	622	800
Berkeley	98	702	800
<b>Total Sample</b>	<b>858</b>	<b>1,542</b>	<b>2,400</b>

**Table 4. Estimated Sample Sizes for Major Ethnic Groups**

	White	Latino	AAPI	African American	AIAN
County Samples	33,232	10,392	4,122	3,173	440
City Samples	921	191	238	188	4
Over-samples	0	0	2,094	0	356
<b>Total Samples</b>	<b>34,153</b>	<b>10,583</b>	<b>6,454</b>	<b>3,361</b>	<b>800</b>

**Table 5. Estimated Samples for Asian American Subgroups**

Group	RDD Sample	Over-sample	Total
Filipino	1,430	0	1,430
Chinese	959	0	959
Japanese	506	294	800
Vietnamese	475	325	800
Asian Indian	314	486	800
Korean	310	490	800
Cambodian	102	98	200
Other AAPI	265	0	265
<b>Total AAPI</b>	<b>4,360</b>	<b>1,694</b>	<b>6,054</b>

**Table 6. American Indian Sample**

IHS Designated Areas	RDD Sample	Over-sample	Total
Urban	187	213	400
Rural	257	143	400
<b>Total Sample</b>	<b>444</b>	<b>356</b>	<b>800</b>