California Health Interview Survey

Making California's Voices Heard on Health

CHIS 2001 Methodology Series

Report 4

Response Rates

CALIFORNIA HEALTH INTERVIEW SURVEY

CHIS 2001 METHODOLOGY SERIES

REPORT 4

RESPONSE RATES

August 2002 *(rev3/10/03)*

This report was prepared for the California Health Interview Survey by J. Michael Brick, Ismael Flores-Cervantes, W. Sherman Edwards, and R. Lee Harding of Westat, Inc

www.chis.ucla.edu

This report provides analysts with information about the response rates in CHIS 2001. The response rates are estimates of the percentage of sampled persons that participated in the survey, where the sample may be across the entire state, restricted to a county, or some other subgroup. To estimate response rates, the probability of sampling persons is taken into account. Thus, the response rates are weighted percentages of the number responding rather than simple unweighted percentages. Procedures used to increase the response rates are also discussed and, where possible, evaluated.

Suggested citation:

California Health Interview Survey. *CHIS 2001 Methodology Series: Report 4 – Response Rates*. Los Angeles, CA: UCLA Center for Health Policy Research, 2002.

The California Health Interview Survey is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Funding for CHIS 2001 came from multiple sources: the California Department of Health Services, The California Endowment, the California Children and Families Commission, the National Cancer Institute, the Centers for Disease Control and Prevention, and the Indian Health Service.

PREFACE

Response Rates in CHIS 2001 is the fourth in a series of methodological reports describing the 2001 California Health Interview Survey (CHIS 2001). 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 Health Services, and the Public Health Institute. Westat was responsible for the data collection and the preparation of five methodological reports from the 2001 survey. The survey examines public health and health care access issues in California. CHIS telephone survey is the largest state health survey ever undertaken in the United States. The plan is to monitor the health of Californians and examine changes over time by conducting periodic surveys in the future.

Methodological Reports

The first five methodological reports for the 2001 CHIS are as follows:

- Report 1: Sample Design for CHIS 2001
- Report 2: Data Collection Methods in CHIS 2001
- Report 3: Data Processing Procedures in CHIS 2001
- Report 4: Response Rates in CHIS 2001
- Report 5: Weighting and Variance Estimation for CHIS 2001

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.

This report describes the response rates from CHIS 2001. Response rates are the ratio of the number of units interviewed to the number of eligible sampled units, but the computation of response rates for CHIS 2001 is involved because of the complexity of the survey. This report presents the rates and explains the rationale for the procedures used in computing the response rates from CHIS 2001.

The primary purpose of presenting these response rates is to provide information for analysts of the data. As a result, the response rates are reported separately for the main analysis subgroups—adults (ages 18 and older), children (age less than 12), and adolescents (ages 12 to 17). The response rates are estimates of the percentage of sampled persons that participated in the survey, where the sample may be across the entire state, or it may be restricted to a county or another subgroup. To estimate response rates, the probability of sampling persons is taken into account. Thus, the response rates are weighted percentages of the number responding rather than simple unweighted percentages.

A secondary goal of this report is to examine procedures used in the survey to increase the response rates. The specific operational methods are described more completely in Report 2 (Data Collection Methods), but the methods are summarized here briefly to provide some context for the examination.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Figures

1. CHIS 2001 DESIGN AND METHODOLOGY SUMMARY

1.1 Overview

The 2001 California Health Interview Survey (CHIS 2001) is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. The focus of the survey is on a variety of public health topics, including access to health care and health insurance coverage. CHIS 2001 is the largest state health survey ever undertaken in the United States. It is a random digit dialing (RDD) telephone survey of California households designed to produce reliable estimates for the whole state, for large- and medium-sized population counties in the state, and for groups of the smallest population counties. Three California cities that have their own health departments were also sampled as part of CHIS 2001.

The survey design supports study of California's major race and ethnic groups, and a number of smaller ethnic groups within the state. Adults, parents of children below age 12, and adolescents (ages 12-17) residing in California households are the eligible respondents to the survey. CHIS 2001 collected data between November 2000 and October 2001. The plans are to conduct independent cross-sectional surveys of the California population on a biannual basis to monitor important health-related indicators and potentially track changes over time. CHIS 2001 is the first of these planned surveys.

CHIS 2001 collected information on if, where, and how people get health care in California. The goal is to provide health planners, policymakers, state, county, and city health agencies, and community organizations with information on the health and health care needs facing California's diverse population. For example, the number and characteristics of adults, children, and adolescents without access to care and lacking health insurance can be estimated from the data collected in CHIS 2001. Other key estimates on the prevalence of cancer screening, diabetes, asthma, and other health conditions can also be produced. The survey includes major content areas, such as health status and conditions, health-related behaviors, access to health care services, and health insurance coverage.

1.2 Sample Design Objectives

The CHIS 2001 sample is designed to meet two objectives: (1) provide local-level estimates for counties and groupings of counties with populations of 100,000 or more; and (2) provide statewide estimates for California's overall population and its larger race/ethnic groups, as well as for several smaller ethnic groups. To address these objectives, the sample was allocated by county and aggregates of smaller counties, with supplemental samples of selected populations and cities. Table 1-1 shows the sampling strata (i.e., counties and groups of counties that were identified in the sample design as domains for which separate estimates would be produced). A sufficient amount of sample was allocated to each of these domains to support the first sample design objective.

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

1. Los Angeles	15. San Joaquin	29. El Dorado
2. San Diego	16. Sonoma	30. Imperial
3. Orange	17. Stanislaus	31. Napa
4. Santa Clara	18. Santa Barbara	32. Kings
5. San Bernardino	19. Solano	33. Madera
6. Riverside	20. Tulare	34. Monterey, San Benito
7. Alameda	21. Santa Cruz	35. Del Norte, Humboldt
8. Sacramento	22. Marin	36. Lassen, Modoc, Siskiyou, Trinity
9. Contra Costa	23. San Luis Obispo	37. Lake, Mendocino
10. Fresno	24. Placer	38. Colusa, Glen, Tehama
11. San Francisco	25. Merced	39. Sutter, Yuba
12. Ventura	26. Butte	40. Plumas, Nevada, Sierra
13. San Mateo	27. Shasta	41. Alpine, Amador, Calaveras, Inyo,
14. Kern	28. Yolo	Mariposa, Mono, Tuolumne

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

Samples were also drawn from each of the three California cities that have their own local health department. In addition, supplemental samples were developed for three counties that contracted for additional sample to enhance their overall estimates. These city and supplemental county samples were in the following locations:

- The cities of Berkeley, Long Beach, and Pasadena; and
- The counties of San Francisco, Santa Barbara, and Solano.

The three city samples and the Solano county supplemental sample were implemented with and incorporated in the original statewide RDD sample. The separate San Francisco and Santa Barbara supplemental samples were subsequently added to the statewide RDD sample prior to constructing the sample weights and are part of the final CHIS 2001 RDD sample file.

To accomplish the second objective, larger sample sizes were allocated to the more urban counties where a significant portion of the state's African American and Asian ethnic populations reside. Additionally, supplemental samples were used to improve the sample size and precision of the estimates for specific ethnic groups. The supplemental ethnic group samples in CHIS 2001 were as follows:

- South Asian, Cambodian, Japanese, Korean, and Vietnamese;
- American Indians and Alaska Natives in urban and rural areas; and
- Latinos residing in Shasta County (a sample requested by the local health department).

1.3 Data Collection

To capture the rich diversity of the California population, interviews were conducted in six languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, Korean, and Khmer (Cambodian). These languages were chosen based on research that identified these as the languages that would cover the largest number of Californians in the CHIS sample design 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 the CHIS 2001 data collection for the CHIS project. Westat staff interviewed one randomly selected adult in each sampled household. In those households with children (under age 12) or adolescents (ages 12-17), one child and one adolescent were randomly sampled, so up to three interviews could have been completed in each sampled household. The sampled adult was interviewed, and the parent or guardian who knew the most about the health and care of the sampled child was interviewed. The sampled adolescents responded for themselves, but only after a parent or guardian gave permission for the interview. Since adolescents were not reliable sources concerning their own health insurance coverage, the parents of sampled adolescents were interviewed about this topic separately.

One criterion for the adolescent and child to be selected for the survey is that they had to be "associated" with the selected adult. This meant that in most cases the interviewed adult had to be either the parent or guardian. The CHIS 2001 sample weights adjust for this selection criterion so as not to bias estimates based on the adolescent and child surveys. Table 1-2 shows the number of completed adult, child, adolescent, and adolescents' health insurance interviews in CHIS 2001, by the type of sample (RDD or supplemental sample).

Table 1-2. Number of completed interviews by type of sample, instrument

Type of sample	Adult	Child	Adolescent	Adolescent insurance
Total RDD + supplemental cases	57,848	13,276	6,058	8,302
RDD (includes 3 cities + Solano county supplemental cases)	54,122	12,392	5,733	7,809
Santa Barbara supplemental cases	206	49	22	31
San Francisco supplemental cases	1,100	151	46	79
Total CHIS 2001 RDD file	*55,428*	*12,592*	*5,801*	*7,919*
Other supplemental samples:				
South Asian	443	158	39	65
Cambodian	126	44	37	44
Japanese	330	51	18	33
Korean	326	95	30	44
Vietnamese	540	124	34	60
American Indian/Alaska Native	351	106	51	71
Shasta Latinos	304	106	48	66

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

The interviews done in English were administered using Westat's computer-assisted telephone interviewing (CATI) system. Spanish and Vietnamese language interviews were also conducted entirely in CATI, while interviews conducted in Cantonese, Mandarin, Korean, and Khmer used English CATI screens and paper translations in tandem. The average adult interview took around 32 minutes to complete. The average child and adolescent interviews took 14 minutes and 19 minutes, respectively. Interviews in the non-English languages generally averaged longer to complete. Approximately 12 percent of the adult interviews were completed in a language other than English, as were 21 percent of all child (parent proxy) interviews and 9 percent of all adolescent interviews. Table 1-3 shows the major topic areas for each of the three survey instruments (adult, child, and adolescent).

Table 1-3. Survey topic areas by instrument

Adult interview	Child interview	Adolescent interview
Age, sex, race, ethnicity	Age, sex, race, ethnicity	Age, sex, race, ethnicity
Physical activity		Physical activity
	Bike helmet use	Bike helmet, seatbelt use
	Recent serious injury	Recent serious injury
Health status	Health status	Health status
Women's health	Child care	
Chronic health conditions	Asthma, ADD	Asthma, diabetes
Cancer history, screening		
Skin cancer prevention	Skin cancer prevention	Skin cancer prevention
Health care use and access	Health care use and access	Health care use and access
Alcohol, tobacco use		Alcohol, tobacco, drug use
Mental health		Mental health
Health insurance	Health insurance	Health insurance
Diet (fruit-vegetable intake)	General diet	General diet
Dental health	Dental health	Dental health
Employment		Employment
Gun access, training		Gun access, violence
Income		
	Family interaction	Parental involvement
	Video games, computer use	Video games, computer use
Sexual orientation		Sexual behavior, orientation
		Future plans

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

1.4 Response Rate

The overall response rate for CHIS 2001 is a composite of the ***screener completion rate*** (i.e., success in introducing the survey to a household in order to select a respondent), and the ***extended interview completion rate*** (i.e., success in getting the selected respondent to complete the full interview). For the adult survey, the screener completion rate was 59.2 percent and the extended interview completion rate was 63.7 percent. This gives an overall response rate of 37.7 percent. To maximize the survey's response rate, an advance letter (in five languages) was mailed to all sampled telephone numbers for which an address could be obtained from reverse directory services. Approximately 66 percent of the sample was mailed an advance letter. Response rates varied by sampling stratum and were slightly higher in households that received an advance letter.

To assist in achieving sample size goals, respondents that completed 80 percent of the questionnaire (i.e., through Section I on health insurance) after all followup attempts were exhausted to complete the full questionnaire were counted as "complete." This resulted in 397 "partial completes" being included in the final adult survey data. Employment and income information as well as potential public program eligibility and food insecurity information would be missing from these cases.

Proxy interviews were allowed for frail and ill persons over the age of 65. The reason is that health estimates made for elderly persons could be biased if this is not allowed. Eligible selected persons were recontacted and offered a proxy option and 316 had a proxy interview completed by either a spouse/partner or adult child. Only a subset of questions identified as appropriate for a proxy respondent were administered.

1.5 Weighting the Random Digit Dial Sample

To produce correct population estimates for the RDD CHIS results, weights are applied to the sample data to compensate for a variety of factors, some directly resulting from the design and administration of the survey. Sample weighting was carried out in CHIS 2001 to accomplish the following objectives:

- Compensate for differential probabilities of selection for households and persons (Note: households with listed addresses and thus eligible for an advance letter were assigned a probability of selection of 1.25 over unlisted households);
- Reduce biases occurring because nonrespondents may have different characteristics than respondents;
- Adjust, to the extent possible, for undercoverage 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 for the RDD samples (each stratum is an independent sample), a household weight was created for all households that completed the screener interview. This household weight is the "base weight" computed as the inverse of the probability of selection of the sample telephone number adjusted for each of the following:

- Subsampling for listed address/advance letter status;
- Unknown residential status;
- Screener interview nonresponse;
- Multiple telephone numbers; and
- Household poststratification.

A "poststratified household weight" was then used to compute a person-level weight. This person-level weight incorporates the within-household probability of selection of the sampled person and adjusts for nonresponse, plus an adjustment resulting from raking the data to person-level control totals. Each of these adjustments corresponds to a multiplicative weighting factor.

Raking can be thought of as a multidimensional poststratification procedure because the weights are basically poststratified to one set of control totals (a dimension), then these adjusted weights are poststratified to another dimension. After all dimensions were adjusted, the process was iterated until the control totals for all the dimensions were simultaneously satisfied (within a specified tolerance).

There are 11 dimensions used in CHIS 2001. The first 10 dimensions are created by combining demographic variables (age, sex, race, and ethnicity) and different geographic areas (city, county, group of counties, and state). The 11th dimension is created to adjust the weights for households without a telephone number.

The control totals used in the raking were derived from the *Census 2000 Summary File 1* (SF1). Population items in SF1 include sex, age, race, ethnicity (Latino/non-Latino), household relationships, and group quarters. The race classification in SF1 includes six groups: White, African American, American Indian/Alaska Native, Asian, Native Hawaiian/Pacific Islander, and a category of Other Race. Since a person could report multiple races, the SF1 provided counts for each of 63 possible race combinations a person could report.

One of the limitations of using the SF1 for the control totals is the inability to produce counts that exclude the fraction of the population living in "group quarters" (e.g., nursing homes, prisons) for some dimensions used in CHIS 2001. The group quarter population represented 2.4 percent of the total population in California. As a result, the number of persons living in group quarters was estimated for some of the raking dimensions, and the SF1 totals were reduced by these estimated amounts prior to raking.

1.6 Imputation Methods

Three different imputation procedures were used in CHIS 2001 to fill in missing responses that were essential for weighting the data or for such basic descriptive purposes as income categories. The first imputation technique is deterministic or non-stochastic in nature. Deterministic imputation was used to fill in the missing items for self-reported county of residence (item AH42). These imputations required no randomization because other geographic data are available that can be used to determine the respondent's county of residence with a relatively high level of probability of being correct although not with 100 percent certainty in all cases.

The second imputation technique is a completely random selection from the observed distribution. This method is used only when a very small percentage of the items are missing. For example, when imputing the missing values for self-reported age, the distributions of the responses for age by type of interview (adult, child, or adolescent) were used to randomly assign an age using probabilities associated with these distributions.

The third technique is hotdeck imputation. Hotdeck imputation was used to impute race, ethnicity, and household income in CHIS 2001. The hotdeck approach is probably the most commonly used method for assigning values for missing responses in large-scale household surveys.

With a hotdeck, 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. To carry out hotdeck imputation for CHIS 2001, the respondents to an item form a pool of donors, while the nonrespondents are a group of recipients. A recipient is matched to the subset pool of donors, with the same household structure. The recipient is then randomly imputed the same household income, ethnicity/race (depending on the items that need to be imputed) from one of the donors in the pool. Once a donor is used, it is removed from the pool of donors.

Imputation flags are used in the data file to identify all imputed values.

1.7 Methodology Report Series

A series of five methodology reports are available with more detail about the methods used in CHIS 2001:

- Report 1 - Sample Design
- Report 2 - Data Collection Methods
- Report 3 - Data Processing Procedures
- Report 4 - Response Rates
- 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 www.CHIS.ucla.edu or contact CHIS at CHIS@ucla.edu.

2. USE OF RESPONSE RATES

While the use of response rates as a single measure of the quality of a survey is unjustified, Madow et al. (1983) suggest response rates provide valuable information on the success of the survey at representing the population sampled. Keeter et al. (2000) and other researchers note that the response rate alone is not sufficient for this objective because the bias in an estimate is related to both the response rate and the characteristics of those responding. This relationship is discussed in more detail below.

The main objective of this report is to present response rates that can be used by analysts of the CHIS 2001 data to better understand how well the California population is represented. To accomplish this goal, the response rates are weighted so that the weighted response rate is an estimate of the proportion of the population responding. For example, since the sample was selected with differing sampling rates by county, the response rates are weighted so each county accounts for its appropriate fraction when the state response rate is reported.

The rationale for using weights in computing the response rate is that the bias of a simple estimate such as a mean, $\bar{y}_r$, is a function of the response rate and the difference in the means between the respondents and nonrespondents. A simple way of conceptualizing this is by assuming the population is partitioned into a stratum of respondents (*R*) and a stratum of nonrespondents (*NR*). The estimate from the survey is computed from the observations in the respondent stratum, where each observation is weighted by the inverse of its selection probability. In a probability sample, the bias of a survey estimate of a mean is

$$bias(\bar{y}_r) = (1-r)(\bar{Y}_R - \bar{Y}_{NR}), \tag{1}$$

where *r* is the appropriately weighted response rate and the quantity on the right is the difference in the means between the respondent and nonrespondent strata (Lessler and Kalsbeek 1992). This formula shows the bias increases as the response rate decreases, if the difference in characteristics between respondents and nonrespondents remains constant. If the response rates are not weighted, this relationship does not hold. Returning to the example, if the county samples are not weighted by their selection probabilities, then the response rate cannot be used in bias equation (1).

The weighted response rates are computed this way for the entire RDD sample and the geographic supplemental samples. The supplemental samples for the race and ethnic subgroups followed very different data collection procedures and are given separately in Chapter 8. For example, consider the Vietnamese supplemental sample. In this sample, households without any Vietnamese adults were ineligible and no adult was sampled to be interviewed. All of the race and ethnic supplemental samples were equal probability samples and no weights are used in the computation of rates for these. Thus, only the rates for the RDD and geographic supplemental samples are weighted.

3. DEFINING RESPONSE RATES

The first step is to define "response rate" because the term is used in so many different ways across surveys. Two publications that describe response rates in a relatively consistent manner are the *Council of American Survey Research Organizations* (CASRO 1982) and *The American Association for Public Opinion Research* (AAPOR 1998). The AAPOR report is periodically updated and is available on the organization's Uniform Resource Locator (http://www.aapor.org).

Both reports recommend that a survey response rate be defined as the ratio of completed interviews to eligible reporting units. However, the application of this recommendation is more difficult than it may appear, especially in RDD surveys. One problem in estimating response rates in RDD surveys is the determination of the eligibility of some of the sampled numbers. Some telephone numbers are never answered or are only picked up by answering machines, even after being called multiple times over a range of days. This outcome may occur for many reasons, as discussed by Shapiro et al. (1995). The eligibility of these numbers cannot be determined directly, adding ambiguity to the definition of a response rate.

We use the basic procedures described in the AAPOR (1998) report, but incorporate a relatively new method to resolve this ambiguity in determining the number of eligible telephone numbers. The AAPOR report has several different response rate definitions and the one used in this report is AAPOR's "RR4" equation (note that at the screener level RR4 is equal to AAPOR's RR3 equation since there are no partial screener interviews). Since the CHIS sample of telephone numbers was sampled with different probabilities of selection, we use the weighted number of telephone numbers rather than the unweighted number in the computation. This also compensates for the under-sampling and over-sampling that occurred in different geographic areas.

Telephone numbers with unknown eligibility are assigned to be either eligible (i.e., a residence) or ineligible based on the "survival method" (Brick, Montaquila, and Scheuren 2002). This approach is a more empirically based estimate of the percentage of the "unknown eligibility" telephone numbers that are likely to be eligible (residential). The method works by selecting a subsample of the telephone numbers with unknown eligibility and dialing these numbers additional times. In CHIS, Westat selected a sample of 7,881 telephone numbers that had unknown residential status after 14 call attempts and dialed these at least 7 more times to determine their status. The data are then analyzed using survival

analysis to predict the percentage of telephone numbers that would be residential if dialing continued indefinitely. This estimation of the proportion of eligible telephone numbers is denoted as ***'e'*** in the AAPOR RR4 equation. After the eligibility proportion is established, the response rate can be computed as the weighted ratio of the responding telephone numbers to the total of known and estimated eligible numbers. We do compute some alternative response rates using the CASRO equation (CASRO estimates 'e' as the proportion of the resolved telephone numbers that are residential) and other approaches for comparison purposes, but these are not used for most of the analysis.

The next step in computing response rates depends on the particular part of the interview process being analyzed, such as the adult interview. For example, to compute the response rate for the adult interview, the numerator of the rate is the weighted number of completed adult interviews, and the denominator is the weighted number of adults sampled in households that completed the screening interview. An overall or joint response rate can be computed by multiplying the screening and adult interview rates.

Computing a response rate for a subgroup requires that all the units in both the numerator and denominator of the rate can be classified as members of the subgroup. To do this, data must be available to classify all sampled units, not just respondents. Because the screening interview identifies if any children were in the household, extended response rates can be computed separately for households with children and without children. However, the joint rate must be computed by multiplying the extended rate for the subgroup by the overall screener response rate because data on presence of children are not available for every sampled telephone number.

At the screener level, data on the RDD sample of telephone numbers are limited; the telephone numbers can be classified by geography (county or group of counties used for sampling) and by whether there was an address for the telephone number that could be used to send an advance mailing. At the extended interview or person level, data from the screener can be used to classify households by characteristics that are known for all completed households. These data are used to compute the response rates in CHIS 2001 later in this report.

4. REVIEW OF CONTACT METHODS

Report 2 (Data Collection Methods) gives a detailed discussion of the methods used in CHIS 2001 to contact and interview persons. Here we briefly review the key procedures to provide some background on the response rates and evaluation measures presented later in this report.

As mentioned before, the survey contained both screening and extended interviews. In virtually every household, one adult was sampled for an extended interview. In households with persons under age 17 one child and one adolescent were also sampled. The screening interview took about 2 to 3 minutes to conduct, on average. The parent or guardian of the child was interviewed about the child and the adolescent was interviewed, if the parent gave permission for the sampled adolescent to be interviewed. The adult extended interview averaged about 32 minutes in length, the child interview was about 14 minutes, and the adolescent interview lasted about 20 minutes. The interviews in other languages took longer to complete. More detailed timings on the interviews are given in Report 2 (Data Collection Methods).

Shortly before the beginning of the survey, an advance or prenotification letter was sent to the sampled telephone numbers for which an address could be obtained from available sources. The intent of the letter was to inform the household that they would be called to participate in CHIS 2001, that their participation was important to the success of the survey, and that the survey was legitimate.

After the advance mailing, initial telephone calls were made to complete the screener interview with a household respondent who was at least 18 years old. Multiple attempts, at least 14 attempts if needed, were made to establish the initial contact with the household. If the household refused to participate, additional attempts were made to complete the screener after waiting at least 2 weeks after the first refusal. Prior to attempting to convert these refusals into participants, an express letter was sent to the household (if an address was available) informing them about the validity of the study and the importance of their participation. If the household refused again, a second refusal conversion telephone attempt was made at least another 2 weeks later.

A similar process was used at the extended level for the sampled adult. The sampled adult was asked to participate in the study up to three times—an initial attempt and two attempts at refusal conversion. If the adult refused, an express letter was sent (if an address was available) urging him or her

to participate. A second refusal conversion attempt for both the screener and the adult extended interview was only done for the subset of those cases where the review of interviewer reports on the previous refusals indicated an additional attempt was warranted. For child and adolescent interviews, one refusal conversion attempt was made. No express letters were sent for either the child or adolescent interview. However, if the parent refused permission for the adolescent to be interviewed, then an express letter was mailed to the parent asking him or her to reconsider. Attempts at refusal conversion were stopped at any point if the respondent expressed hostility at being called or specifically requested that they not be called again.

A variety of other methods were used to increase response rates in CHIS 2001. A very important procedure was translating and conducting the interview in Spanish, Chinese (Cantonese and Mandarin), Khmer, Korean, and Vietnamese to accommodate households that did not speak English. Another persuasion method was the use of media campaigns to encourage participation. Yet another method to increase response rates was the use of proxy interviews for adults who were over age 65 and unable to participate because of mental or physical limitations. Other adult household members knowledgeable about the sample persons' health, almost always a spouse or child of the sampled adult, completed a proxy interview in these cases.

In addition to the efforts to encourage respondents to participate, other approaches were used to increase response rates. Interviewers were trained and given refresher training on methods to avoid refusals and to convert those who had refused. Only those interviewers who had above average response rates were trained and allowed to conduct the refusal conversions. Multiple call attempts were made to contact sampled household members to complete the extended interviews. On average, 17 call attempts were made to contact an adult before a case was classified as a nonrespondent.

Later in this report, we discuss some of these methods and describe the increases in the number of interviews that resulted, where this is possible. Of course, some methods such as interviewer training cannot be assessed quantitatively without specially designed experiments.

5. RESPONSE RATE FORMULAS

This chapter describes the formulas used to compute the response rates for CHIS 2001. The types of response rates are screener response rates, extended interview or person response rates, and overall or joint response rates. We begin with the screener response rates.

A screener response rate is calculated for each sampled county or group of counties that were combined for sampling purposes. In the report we will often refer to these as counties, but we are referring to the groups of counties used in sampling as given in Table 1.1. In the tables of response rates, these groups are called sampling strata to avoid confusion. Response rates for a group of counties or the entire state can be computed in the same way. The formula for the screener response rate (rr_s) in a sample county is

$$rr_s = \frac{\sum_{i \in S_{resp}} w_i}{\sum_{i \in S_{resid}} w_i} \ , \tag{2}$$

where w_i is the weight for telephone number i in the county after adjusting for the differential sampling rates and the assignment of the numbers with unknown residential status, S_{resp} is the set of telephone numbers in the county that responded to the screening interview, and S_{resid} is the set of telephone numbers in the county that were residential. As noted earlier, the estimated residential rates were determined using the survival analysis method.

The screener response rate for the state is computed in exactly the same way, except the sum is over all the telephone numbers sampled in CHIS 2001 rather than those in the specific county. The state screener response rate is thus a weighted average of the county screener response rates with weights equal to the population in the counties. As a result, the state response rate differs from what would be obtained from the unweighted average of the response rates of the counties. Since the sample sizes by county were not that different from what would have been obtained by proportional allocation, the unweighted and weighted response rates are not very different.

The extended response rate for the adult interview in a county is the weighted percentage of the adults sampled in the screener who completed the adult extended interview. The weight in this case is

the inverse of the probability of selecting the adult within the household. Because of this weighting, adults sampled from households with more than one adult have a larger effect on the response rate than those in households with only one adult. The extended adult response rate (rr_a) is

$$rr_a = \frac{\sum_{i \in A_{resp}} w_i'}{\sum_{i \in A_{sam}} w_i'}, \tag{3}$$

where the numerator is summed over all adult respondents, and the denominator is summed over all sampled adults. The weight being summed in this case, w', is the adult weight that accounts for selecting the adult within the household. The adult response rate is conditioned on the completion of the screener interview.

The extended response rate computation for children and adolescents is similar to the adult procedure; however, the method of sampling does add some complexity. Persons under 18 years of age are not enumerated in the screener, even though the screener did ascertain whether or not there were children in the household. The full enumeration of persons under 18 is done in the adult extended interview. As a result, the child and adolescent extended response rates are computed for only those households in which the adult extended interview is completed. In other words, the child response rate is a conditional rate like the adult rate, but it is conditional on both the screener and adult interviews being completed.

The extended child response rate (rr_C) is

$$rr_c = \frac{\sum_{i \in C_{resp}} w_i''}{\sum_{i \in C_{sam}} w_i''}, \tag{4}$$

where the numerator is summed over all child respondents, and the denominator is summed over all sampled children. The weight being summed in this case, w'', is the inverse of the probability of selecting the child within the household.

Exactly the same procedure is used for the adolescent extended interview response rate (rr_t) and it is

$$rr_t = \frac{\sum_{i \in T_{resp}} w_i'''}{\sum_{i \in T_{sam}} w_i'''}, \tag{5}$$

where the numerator is summed over all adolescent respondents, and the denominator is summed over all sampled adolescents. The weight being summed in this case, w''', is the inverse of the probability of selecting the adolescent within the household.

An important source of nonresponse for the adolescent interview was the parent denying permission to conduct the interview with the adolescent. The response rate given by (5) includes the parent permission as a source of nonresponse. Another response rate of interest is the adolescent response rate conditioned on the parent giving permission to interview the adolescent. This fully conditional adolescent response rate is

$$rr_{t-p} = \frac{\sum_{i \in T_{resp}} w_i'''}{\sum_{i \in T_{sam\text{-}per}} w_i'''}, \tag{6}$$

where the only difference is that the denominator is summed over only those adolescents for whom the parents gave permission for the adolescent to be interviewed.

The response rates defined above, except for the screener response rate, are conditional rates in the sense that they depend on the household participating in a previous stage of CHIS. Overall response rates eliminate the conditioning. For example, since the adult response rate is conditioned on the completion of the screener, the product of the screener and adult response rate is an unconditional or overall adult response rate. Thus, the overall adult response is

$$orr_a = rr_s \cdot rr_a \tag{7}$$

Since the child response rate is conditioned on both the screener being completed and the adult interview in the household being completed, the overall response rate for the child is the product of the screener response rate, the adult extended response rate, and the child response rate. The overall response rate for the child is

$$orr_c = orr_a \cdot rr_c \,, \tag{8}$$

because orr_a is the product of the screener and adult response rates.

For the adolescent, overall response rate accounting for all the levels of conditioning (completion of the screener, the adult interview, and the permission request) is

$$orr_t = orr_a \cdot rr_t \,. \tag{9}$$

We do not present an overall response rate for the adolescent excluding the permission request because it is not of much interest in an overall rate.

The formulas given above are applied in the next chapter.

6. RESPONSE RATE TABLES

This chapter gives tables of response rates from the RDD and geographic supplemental samples for CHIS 2001. The first tables are the screener response rates, adult response rates, child response rates, and adolescent response rates. The overall response rates for adults, children, and adolescents are then presented. All of the rates in the tables in this chapter are weighted and use the formulas presented in the previous section.

6.1 Screener Response Rates

The screener response rates for each county (sampling stratum) are given in Table 6-1. The first column in the table gives the number of households that were interviewed and completed the screening interview. Overall, 84,051 households across the state cooperated with this first step of the CHIS 2001 interview. In all of these households at least one adult was sampled[1].

The overall screener response rate for the state is 59.2 percent. As discussed in Chapter 3, the response rate is computed using the survival method to allocate the undetermined numbers (those that on every call were not answered or only answered by an answering machine). As mentioned earlier, alternative definitions for allocating these undetermined numbers are used in some other surveys and may give different response rates. One approach used by some is to ignore the undetermined numbers in the computation of response rates. This approach really gives a ***cooperation rate***. Dropping all the undetermined numbers for CHIS 2001 gives an overall state response rate (cooperation rate) of 64.6 percent. Another approach is to use the CASRO rate. The CASRO screener response rate for the entire state is 57.5 percent, almost 2 percentage points lower than the survival method. For the remainder of the report, we use the survival method for all the response rates.

[1] About 562 households completed the screening interview but were not eligible for CHIS 2001 because they were sampled for either the San Francisco or Santa Barbara county supplemental samples but were not eligible because they were not in the county. We have not included these ineligible screeners in the counts of the completes in the tables.

Table 6-1. Number of completed screeners and response rates by sampling stratum and advance letter

	Total		Letter		No letter	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
State Total	84,051	59.2	68,229	61.1	15,822	52.4
Los Angeles	19,668	56.9	16,453	58.5	3,215	51.0
Long Beach	1,287	59.8	1,051	61.2	236	54.8
Pasadena	1,255	55.4	1,082	56.3	173	50.9
Remainder of Los Angeles	17,126	56.8	14,320	58.4	2,806	50.8
San Diego	4,055	59.9	3,434	61.3	621	53.7
Orange	3,992	59.0	3,301	61.1	691	51.8
Santa Clara	2,351	57.1	1,928	58.9	423	50.8
San Bernardino	2,350	63.7	1,838	65.6	512	58.3
Riverside	2,110	62.2	1,642	64.1	468	57.0
Alameda	2,923	57.6	2,491	59.5	432	49.8
Berkeley	1,155	62.1	985	63.0	170	58.3
Remainder of Alameda	1,768	57.2	1,506	59.2	262	49.1
Sacramento	1,810	61.3	1,508	63.0	302	54.7
Contra Costa	1,777	57.6	1,575	59.9	202	45.8
Fresno	1,658	64.0	1,308	65.2	350	60.3
San Francisco	3,686	50.7	2,987	53.5	699	41.2
Ventura	1,485	59.4	1,258	61.7	227	50.4
San Mateo	1,010	53.8	875	55.6	135	45.4
Kern	1,597	68.9	1,268	70.3	329	64.3
San Joaquin	1,609	64.7	1,303	66.8	306	58.0
Sonoma	1,095	61.3	946	65.0	149	46.2
Stanislaus	1,228	65.7	1,032	68.4	196	55.4
Santa Barbara	1,512	62.1	1,201	64.4	311	55.8
Solano	2,415	61.5	2,091	64.4	324	49.3
Tulare	1,232	67.7	963	70.4	269	60.6
Santa Cruz	1,123	57.7	896	60.9	227	49.0
Marin	972	54.7	857	59.5	115	35.5
San Luis Obispo	1,053	61.6	866	65.4	187	49.8
Placer	1,128	60.3	864	64.0	264	51.7
Merced	1,265	66.2	1,006	66.7	259	64.4
Butte	1,198	67.3	959	69.3	239	60.9
Shasta	1,170	65.7	894	68.9	276	58.0
Yolo	1,163	66.2	969	70.9	194	51.2
El Dorado	1,113	57.8	827	62.2	286	48.8
Imperial	1,225	67.0	1,012	69.3	213	58.5
Napa	1,174	59.0	1,016	63.4	158	42.0
Kings	1,249	65.5	985	66.9	264	61.2
Madera	1,207	67.8	898	72.1	309	59.0
Monterey, San Benito	1,201	60.7	968	63.2	233	53.1

Table 6-1. Number of completed screeners and response rates, by sampling stratum and advance letter (continued)

	Total		Letter		No letter	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
Del Norte, Humboldt	1,197	65.4	902	67.6	295	60.4
Lassen, Modoc, Siskiyou, Trinity	1,191	66.5	761	69.5	430	62.2
Lake, Mendocino	1,159	60.9	832	62.2	327	58.1
Colusa, Glen, Tehama	1,229	68.9	879	70.5	350	65.5
Sutter, Yuba	1,228	66.2	948	68.8	280	59.7
Plumas, Nevada, Sierra	1,130	59.5	788	63.0	342	53.3
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	1,113	58.0	700	61.2	413	53.8

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

The table shows that the screener response rates vary by county. Figure 6-1 shows the distribution. The median response for the counties is 61.4 percent, the highest response rate is 68.9 percent in both Kern county and the Colusa-Glen-Tehama stratum, and the lowest is 50.7 percent in San Francisco. As is clear in the figure, the rate in San Francisco is very much at the lower end of the scale and the county with next lowest response rate has a rate that is about 3 percentage points higher than the San Francisco rate. Part of the reason for the lower response rate may be because some of the supplemental sample for the county was added relatively late in the data collection period and there was less time to complete all the call attempts and refusal conversions during the close out of the survey. If we exclude the supplemental sample cases, the response rate for the screener for San Francisco would be about 2 percentage points higher.

Another interesting finding is that the median response rate for counties with a population of over 500,000 persons (the counties from Los Angeles through San Joaquin in the table) is 59.6 percent. This is 4 percentage points lower than the 63.8 percent median response rate for the smaller counties. Looking at the individual counties suggests that this difference may be a function of proximity to a metropolitan area or population density rather than the population size of the county. Small counties that are highly urban have rates that are similar to the rates of the more populous counties.

Table 6-1 also has tabulations on the response rates by whether an advance letter could be mailed to the household. We discuss these rates in Chapter 9 as part of the evaluation. Next, we examine the response rates for the extended adult, child, and adolescent interviews.

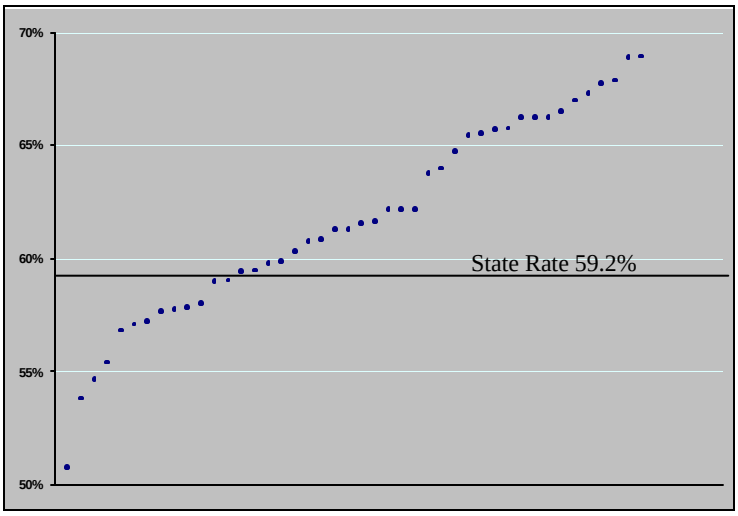

Figure 6-1. Screener response rate distribution by sampling stratum

6.2 Person Response Rates

The adult, child, and adolescent extended interview response rates for each county in CHIS 2001 are given in Table 6-2, along with the number of completed interviews for each of the instruments. A total of 55,428 adult interviews, 12,592 interviews about children, and 5,801 adolescent interviews were completed in this very comprehensive survey of the residents of California. Each of the components of CHIS 2001 is discussed, beginning with the adult interview.

Table 6-2 shows the response rate for the adult interview across the state was 63.7 percent. As with the screener response rate, the counties with larger populations tended to have lower adult extended interview response rates. The median adult response rate for the counties with a population of more than 500,000 is 63.5 percent, while for counties with less than 500,000 the median adult response rate is 67.6 percent. This difference may be attributed to a variety of reasons, including the different distribution of persons by age, education, etc. by county.

Table 6-2. Number of completed extended interviews and response rates by sampling stratum and type of interview

	Adult		Child		Adolescent	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
State Total	55,428	63.7	12,592	87.6	5,801	63.5
Los Angeles	12,215	60.0	2,824	83.7	1,123	58.5
Long Beach	819	61.1	202	87.3	62	52.7
Pasadena	814	62.8	168	81.8	54	57.0
Remainder of Los Angeles	10,582	59.7	2,454	83.5	1,007	59.0
San Diego	2,666	63.3	584	88.5	271	62.1
Orange	2,495	60.3	619	84.5	217	52.3
Santa Clara	1,514	61.2	353	92.2	138	60.1
San Bernardino	1,547	64.0	442	91.2	210	68.0
Riverside	1,386	64.7	376	90.8	157	64.8
Alameda	1,985	65.2	355	90.3	134	57.9
Berkeley	794	65.3	90	93.5	30	51.3
Remainder of Alameda	1,191	65.1	265	89.4	104	60.0
Sacramento	1,238	65.7	302	86.3	143	65.3
Contra Costa	1,199	64.9	263	88.9	128	64.1
Fresno	1,041	59.8	270	88.9	138	64.3
San Francisco	1,969	59.1	272	88.5	81	51.4
Ventura	971	63.7	225	85.4	106	60.6
San Mateo	947	60.4	157	84.5	79	65.0
Kern	1,096	66.6	323	89.2	149	66.2
San Joaquin	1,052	63.7	282	89.9	133	65.7
Sonoma	771	67.8	161	95.0	90	65.3
Stanislaus	819	64.2	198	85.8	95	60.7
Santa Barbara	1,004	66.1	223	89.7	93	63.2
Solano	1,587	63.9	403	87.0	174	65.6
Tulare	827	64.6	223	91.0	107	63.7
Santa Cruz	793	68.3	175	88.6	104	70.5
Marin	752	70.4	133	89.1	68	61.2
San Luis Obispo	799	69.7	152	93.1	67	65.0
Placer	784	68.2	179	90.5	88	70.1
Merced	832	64.0	228	86.7	115	65.2
Butte	825	67.6	169	89.6	71	64.5
Shasta	826	69.4	163	87.0	87	63.2
Yolo	834	69.3	198	95.2	91	68.8
El Dorado	780	67.6	167	92.5	96	74.2
Imperial	798	63.5	226	82.4	154	70.6
Napa	806	66.6	164	84.0	85	61.1

Table 6-2. Number of completed extended interviews and response rates by sampling stratum and type of interview (continued)

	Adult		Child		Adolescent	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
Kings	843	66.6	276	89.5	150	70.1
Madera	824	67.3	185	85.6	104	70.4
Monterey, San Benito	790	62.9	210	87.2	100	66.4
Del Norte, Humboldt	861	69.6	171	92.9	109	69.1
Lassen, Modoc, Siskiyou, Trinity	846	69.6	151	96.1	82	68.1
Lake, Mendocino	813	68.6	140	87.8	82	67.9
Colusa, Glen, Tehama	839	65.9	196	90.7	108	70.4
Sutter, Yuba	822	64.6	179	90.4	81	65.9
Plumas, Nevada, Sierra	814	70.5	149	90.0	92	78.8
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	818	72.4	126	93.7	101	75.2

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

Data collected in the screener interview about the household and the sampled adult can be used to examine the adult extended response rates since the data are available for all sampled adults. Table 6-3 shows the adult response rates by these data items from the screener[2]. The table shows that there was substantial variation in the response rates by the characteristics known in the screening interview. Women responded at a higher rate than men, older adults were more likely to respond than younger adults. The rates were not very different for households with children or without children.

A somewhat surprising finding is that adults in households with fewer adults were more likely to respond than adults in households with more adults. A concern in many RDD surveys is undercoverage of women in households with no other adults. At the extended adult interview, this does not appear to be the case with CHIS 2001[3]. Of course, it is must be noted that these rates are conditional on the screener being completed, so this expected undercoverage may still be present at that level, but there are no data to examine this.

[2] In some cases the data from the screener interview and the adult interview are not consistent. For example, the age of the adult reported by the household member in the screener may be different from the age reported by the sampled adult. All of the data used in these tabulations are the screener data because no revised data are available for the nonresponding adults.

[3] Another explanation of this finding is that in smaller households the person completing the screening interview is more likely to be the sampled adult and there is evidence that extended response rates for persons completing the screening interview are higher than for other adults in the household.

Table 6-3. Adult response rates by characteristics of the sampled adult

Characteristic	Response rate
Total	63.7
Sex	
Male	58.2
Female	68.6
Age	
18 to 30 years	58.0
31 to 45 years	62.9
46 to 65 years	66.5
Over 65 years	69.0
Type of household	
With children	63.0
Without children	64.2
Adults in household	
1	73.6
2	65.0
3 or more	57.3

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

Now, we examine the child extended interview response rates. Overall, Table 6-2 shows that the child level response rates are very high, 87.6 percent across the state. The rates do not vary much by county. The median rate in the more populous counties (88.7 percent) is only about 1 percentage point lower than the rate in smaller counties (89.6 percent).

Table 6-4 gives the child response rates by the characteristics of the child and household using data collected in the adult interview where the children were enumerated for sampling. The child rates do not show much variation by sex, age, or number of children in the household.

The last response rates presented are for the adolescent interview. An important fact to remember is that the adolescent could not be interviewed unless the parent of the adolescent verbally gave permission for conducting the interview. This requirement resulted in lower response rates for the adolescent interviews than the child instrument. Table 6-2 shows the state adolescent response rate was 63.5 percent. If we exclude the nonresponse due to parents not giving permission to interview the adolescent, the response rate rises over 20 percentage points to 84.5 percent. We discuss the differences in greater detail below.

Table 6-4. Child response rates by characteristics of the sampled child

Characteristic	Response rate
Total	87.6
Sex	
Male	87.1
Female	88.1
Age	
Less than 4 years	88.1
4 to 7 years	87.3
8 to 11 years	87.4
Children in household	
1	87.4
2	87.9
3	87.9
4 or more	86.1

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

Unlike the child instrument, we again find that there is a difference in response rates for the adolescent interviews by the size of the county. The more heavily populated counties have a median response rate of 61.3 percent and the counties with less than 500,000 persons have a median response rate of 67.1 percent.

Table 6-5 gives the adolescent response rates by the characteristics of the adolescent and household using data collected in the adult interview. The rates, like the corresponding child rates, are not very different by sex, age, and the number of adolescents in the household.

Table 6-5. Adolescent response rates, by characteristics of the sampled adolescent

Characteristic	Response rate
Total	63.5
Sex	
Male	63.5
Female	63.5
Age	
12 to 14 years	62.7
15 to 17 years	64.4
Adolescents in household	
1	62.0
2	65.1
3 or more	63.8

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

To better understand the success rate for interviewing adolescents we examine the response rates for the adolescent interview only including those adolescents the parents gave permission to interview. This response rate is more indicative of the ability of the survey operations to contact and interview the adolescents. These rates are given in Table 6-6. Table 6-6 is exactly the same as Table 6-5, but the sampled adolescents without parental permission are excluded from the denominator. Even though the response rates in Table 6-6 are about 20 percentage points greater than those in Table 6-5, the relative rates by the characteristics are consistent across the tables. One noticeable difference is the response rate by age of the adolescent. In Table 6-5 the younger group (12 to 14 years) have a response rate about 2 percentage point lower than the older group (15 to 17 years), but this is reversed in Table 6-6 when we exclude parental permission nonresponse. In Table 6-6 the younger group has higher response rate by nearly 3 percentage points. Clearly, parents were less likely to grant permission for the interview for younger adolescents than they were for older adolescents.

Differences in response rates can lead to nonresponse bias as suggested by bias equation (1). To reduce the potential for this bias, these characteristics were taken into account in the development of the estimation weights as described in Report 5 (Weighting and Variance Estimation). For example, nonresponse adjustments were done separately by county, thus accounting for the differences noted above by the size and urbanicity of the counties. In addition, the weights were also adjusted to be consistent with data from the 2000 Census totals so that other residual biases could be reduced.

Table 6-6. Adolescent response rates excluding parental permission nonresponse by characteristics of the sampled adolescent

Characteristic	Response rate
Total	84.5
Sex	
Male	84.2
Female	84.9
Age	
12 to 14 years	86.0
15 to 17 years	83.1
Adolescents in household	
1	83.3
2	85.8
3 or more	85.3

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

6.3 Overall Response Rates

This section presents the overall or unconditional response rates for the adult, child, and adolescent interviews. Table 6-7 gives these response rates for the entire state and by county. As discussed in Chapter 5, the overall rates are the product of rates. For the adult interview, the overall adult response rate is the screener response rate (from Table 6-1) multiplied by the adult response rate (from Table 6-2). The rate is computed using equation (7). For the child and adolescent it is the product of the overall rate for the adult (from Table 6-7) that accounts for both the screener and adult interview participation multiplied by the appropriate child or adolescent response rate (from Table 6-2). These are the rates given by equations (8) and (9), respectively.

Since the rates in these tables are the product of two or more rates at the interview level the issues surrounding the differences in the rates by county, type of household, and characteristic of the sampled person described in the previous sections also apply here.

Table 6-7. Overall response rates by sampling stratum and type of interview

	Adult overall response rate	Child overall response rate	Adolescent overall response rate
State Total	37.7	33.0	23.9
Los Angeles	34.1	28.6	20.0
Long Beach	36.5	31.8	19.2
Pasadena	34.8	28.4	19.8
Remainder of Los Angeles	33.9	28.3	20.0
San Diego	37.9	33.5	23.5
Orange	35.6	30.1	18.6
Santa Clara	34.9	32.2	21.0
San Bernardino	40.8	37.2	27.7
Riverside	40.2	36.5	26.0
Alameda	37.6	33.9	21.8
Berkeley	40.6	38.0	20.8
Remainder of Alameda	37.3	33.3	22.4
Sacramento	40.3	34.8	26.3
Contra Costa	37.4	33.3	24.0
Fresno	38.3	34.0	24.6
San Francisco	30.0	26.5	15.4
Ventura	37.8	32.3	22.9
San Mateo	32.5	27.5	21.1
Kern	45.8	40.9	30.3
San Joaquin	41.2	37.0	27.1
Sonoma	41.5	39.5	27.1
Stanislaus	42.2	36.2	25.6
Santa Barbara	41.1	36.8	26.0
Solano	39.3	34.2	25.8
Tulare	43.8	39.9	27.9
Santa Cruz	39.4	34.9	27.8
Marin	38.5	34.3	23.6
San Luis Obispo	42.9	40.0	27.9
Placer	41.1	37.2	28.8
Merced	42.4	36.7	27.6
Butte	45.5	40.8	29.3
Shasta	45.6	39.7	28.8
Yolo	45.9	43.7	31.6
El Dorado	39.1	36.1	29.0

Table 6-7. Overall response rates, by sampling stratum and type of interview (continued)

	Adult overall response rate	Child overall response rate	Adolescent overall response rate
Imperial	42.5	35.1	30.0
Napa	39.3	33.0	24.0
Kings	43.6	39.0	30.6
Madera	45.7	39.1	32.2
Monterey, San Benito	38.2	33.3	25.4
Del Norte, Humboldt	45.5	42.3	31.5
Lassen, Modoc, Siskiyou, Trinity	66.5	46.3	44.5
Lake, Mendocino	60.9	41.8	36.7
Colusa, Glen, Tehama	68.9	45.4	41.2
Sutter, Yuba	66.2	42.8	38.7
Plumas, Nevada, Sierra	59.5	41.9	37.7
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	58.0	42.0	39.3

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

7. COMPARISONS TO OTHER SURVEYS

The response rates for CHIS 2001 are difficult to compare to the rates from other surveys for a variety of reasons. One is the nature of the survey itself, where one adult is sampled from each household and asked to complete a relatively lengthy interview and then other household members are interviewed if there are children present. Other reasons include the size and scope of the CHIS interviews and the fact that it is done only in California.

A more generic reason for the difficulty has to do with the lack of detailed information on disposition codes available on most RDD surveys conducted in the U.S. This state of affairs is noted in several places. APPOR (1998) is a recent attempt to address this problem. Nonetheless, the currently available data to make direct comparisons are very sparse. We mention some specific comparability problems below.

Another factor affecting comparability is the recency of the data collection. Response rates in RDD surveys have been decreasing over the years, especially at the end of the decade of the 1990s. A series of articles on declines in response rates in RDD surveys was published beginning in 1998 in *Survey Research,* a newsletter of the Survey Research Laboratory at the College of Urban Planning and Public Affairs—University of Illinois at Chicago (http://www.srl.uic.edu/publist/srvrsch.htm).

Massey, O'Connor, and Krotki (1997) is the only article that reviews RDD response rates for general surveys with any specificity. Massey *et al.* identified RDD surveys conducted in the U.S. from 1990 to 1995 and attempted to document the response rates for these surveys. In the article, the authors note that the response rates are the ones reported by the surveys and were not standardized. Standardization is critical to make the rates comparable. The mean reported response rate was 62 percent with the median between 60 and 64 percent, although it is not completely clear what these rates refer to. The authors state that nearly half of the surveys excluded never-answered and answering machine-only telephone numbers in their computations of response rates. This exclusion increases the screener response rates in CHIS 2001 by 5.4 percentage points (64.6 *vs.* 59.2), as mentioned in Chapter 6. Some of the work discussed by Massey et al. has been continued and attempts made to standardize the rates, but thus far no standardized response rates have been published.

One RDD survey that can be compared to CHIS 2001 is the California Behavioral Risk Factor Surveillance System (BRFSS) survey. In this survey, one adult in each household is sampled and asked to complete an interview of about 20 minutes on health-related topics. The BRFSS interview is shorter than CHIS 2001 and does not have multiple interviews within the household, but has many other similarities to CHIS.

The Centers for Disease Control and Prevention (CDC) (2001) publishes response rates for the BRFSS. The CDC report for the survey shows detailed disposition codes, very much in the spirit of the AAPOR (1998) recommendations. Despite the detail given, it is very difficult to map the BRFSS disposition codes unambiguously to the corresponding disposition codes used in CHIS 2001 because different survey organizations use different classification schemes to create the disposition codes. The codes from both systems provide much needed information for the conduct of the operations of the survey, but they are not the same. This difference highlights the difficulty of making direct comparisons between surveys.

A response rate and a cooperation rate are reported for the 2000 BRFSS in Table 3 and Table 4 of the Quality Report. For California the two rates are 43.4 percent and 46.1 percent, respectively. We are not positive which rate is more comparable to the CHIS overall adult response rate of 37.7 percent. Nevertheless, the rates for the two surveys are not very different, given the potential discrepancies in the definitions and methods. For example, the CDC report suggests that the rates for some states do not include persons who speak a language other than English (or English and Spanish) as eligible, but does not specify the practices for California. A slightly lower rate for CHIS might be anticipated given the longer interview and the multiple interviews per household.

Another survey that can be compared to CHIS 2001 is the 1999 National Survey of America's Families (NSAF). One advantage of using this study for comparison is that Westat also conducted the 1999 NSAF, so the methods of computing response rates are the same as those used for CHIS. The 1999 NSAF also had a large enough sample size in California to provide reliable estimates of the response rates. A major difference between the two studies that has a large effect on the response rates is the use of monetary incentives. The 1999 NSAF used monetary incentives while CHIS 2001 did not. As discussed in more detail in Chapter 9, monetary incentives were very effective in raising the response rates for the NSAF. Another difference between CHIS 2001 and the 1999 NSAF is the way the sampling and interviewing was done. The 1999 NSAF used the same approach of only enumerating some persons in the screener, but children rather than adults were enumerated in the NSAF screener. This is just the

opposite of the approach used in CHIS 2001, where adults were enumerated rather than children. The difference is related to the focus of the survey. Another difference is that in the NSAF households with children and with low income were sampled at a higher rate than other households, so a substantial fraction of the households only had to complete the screener interview. In CHIS, an adult was sampled in virtually every household.

The 1999 NSAF overall response rate for adults in the California RDD sample was 51.7 percent. The NSAF rates are given in Table 5-22 in the report by Brick et al. (2000). The NSAF response rate is higher than the CHIS rate and much of the difference is probably due to the factors mentioned above, especially the use of monetary incentives in the NSAF.

The NSAF surveys also provide some additional information on bias as well as on response rates. A special follow-up study was conducted for the 1997 NSAF. See Groves and Wissoker (1999). The study used intensive methods to recontact a subsample of households that would not participate in the NSAF and compared their characteristics to the NSAF respondents. The study also included other analytic investigations of indicators of nonresponse bias. The results of the study suggested that the nonresponse did not substantially bias the estimates from the survey. Since many of the same procedures were used in CHIS 2001, the results are encouraging for CHIS. These results are consistent with the findings of Keeter et al. (2000) and Curtin, Presser, and Singer (2000).

The discussion and assessment of methods used in CHIS 2001 to increase response rates are presented in Chapter 9, after the response rates for the race and ethnic supplemental samples are given.

8. SUPPLEMENTAL RACE AND ETHNIC SAMPLE RESPONSE RATES

The supplemental race and ethnic samples had sampling and data collection protocols that were different from those used in the RDD and the geographic supplemental samples. The details on the sampling methods and data collection procedures for the race and ethnic supplemental samples are given in Report 1 (Sample Design) and Report 2 (Data Collection Methods), respectively. Below, we summarize a few differences that are related to response rates.

One of the main differences in sampling is that the race and ethnic supplemental samples were selected from special lists rather than by RDD methods. In addition, these supplemental samples were sampled at the same rate across the state rather than using different rates by county. As a result, the differential weights used for the RDD and geographic sample response rates are not appropriate for the race and ethnic supplemental samples. All the response rates for the supplemental samples are unweighted and specific to the particular supplemental sample. No aggregates over these supplemental samples are given.

The goal of the supplemental samples was to increase the sample size for the particular group to make it possible to produce reliable estimates for the group. Table 8-1 gives the number of completed interviews for each sample by the type of interview. The estimation methods used to achieve this goal are described in Report 5 (Weighting and Variance Estimation).

The response rates for the supplemental samples are given in Table 8-2 by the type of supplemental sample and interview type. Since the supplemental samples target individuals with specific race or ethnic backgrounds, questions about the race and ethnicity of the household members were asked in either the screening interview or the adult interview. If the sampled person had the required race or ethnicity, then the interview was conducted with them. If they did not, then they were designated as ineligible for the interview. The response rates given in Table 8-2 include both the completes and the ineligibles as respondents because both participated in the survey to the fullest extent requested[4]. Only the known eligible and completed extended interviews are included in Table 8-1.

[4] In some cases it is impossible to separate the eligible and ineligible respondents. For example, if the screener did not ask about race or ethnicity, then it is not possible to determine if the household was eligible unless they responded to the adult extended interview. For the supplemental sample, a household was classified as ineligible if no one in the household spoke either English or the language of the target subgroup (e.g., Korean). This practice differs from the one used in the RDD and geographic supplemental samples where language cases are counted as nonrespondents.

The response rates in Table 8-2 for the eligible adult, child, and adolescent interviews are the conditional rates similar to the ones given in Section 6.2 for the RDD and geographic supplemental samples. For the adolescents, the rates are given including and then excluding parental permission similar to those given for the RDD sample. Overall rates like those in Section 6.3 are computed by multiplying the appropriate components together and are given in Table 8-3. For example, the overall adult interview rate for the American Indian/Alaska Native supplemental sample is 49.8 percent (the product of 72.6 percent and 68.5 percent).

The response rates for the supplemental samples are generally higher than the rates for the RDD and geographic supplemental samples. The rates are higher even though the interviews took longer to complete. We suspect one reason for the higher rates is that the respondents understood the survey included special efforts to interview persons in these subgroups and this gave them a clear idea of the importance of the survey. The ability to conduct the survey in their native language and the cultural sensitivity training of the interviewers are other important reasons for the higher response rates for these subgroups.

Table 8-1. Number of completes by type of supplemental race/ethnic sample and interview type

Supplemental sample	Screener	Eligible adult	Eligible child	Eligible adolescent
South Asian	1,120	443	158	39
Cambodian	585	126	44	37
Japanese	604	330	51	18
Korean	516	326	95	30
Vietnamese	973	540	124	34
American Indian/Alaska Native	626	351	106	51
Shasta Latinos	447	304	106	48

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

Table 8-2. Response rates by type of supplemental race/ethnic sample and interview type

Supplemental sample	Screener	Eligible adult	Eligible child	Eligible adolescent	Eligible adolescent*
South Asian	59.1	66.9	85.1	57.1	81.6
Cambodian	56.6	76.4	93.6	84.8	88.6
Japanese	56.2	62.9	79.7	57.6	86.4
Korean	62.1	68.4	89.6	68.9	83.8
Vietnamese	59.1	59.8	75.2	52.9	63.2
American Indian/Alaska Native	72.6	68.5	93.0	70.8	78.5
Shasta Latinos	75.3	73.4	90.7	74.2	89.1

*Excluding parental permission nonresponse

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

Table 8-3. Overall response rates by type of supplemental race/ethnic sample and interview type

Supplemental sample	Eligible adult	Eligible child	Eligible adolescent	Eligible adolescent*
South Asian	39.5	33.6	22.6	32.3
Cambodian	43.2	40.5	36.7	38.3
Japanese	35.3	28.2	20.4	30.5
Korean	42.5	38.1	29.3	35.6
Vietnamese	35.3	26.6	18.7	22.3
American Indian/Alaska Native	49.7	46.2	35.2	39.0
Shasta Latinos	55.3	50.1	41.0	49.2

*Excluding parental permission nonresponse

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

9. METHODS OF INCREASING RESPONSE RATES

In this chapter we examine procedures used to increase response rates in CHIS 2001. The ability to evaluate the effects of the procedures is limited because no experiments were conducted especially for this purpose. The first administration of CHIS was challenging to implement within the desired schedule because so much of the survey content and structure was being developed and refined just prior to implementation. As a result, many of the decisions about procedures to encourage response were based on the existing literature of RDD surveys and specific evaluations were not included in the initial implementation.

One of the decisions made for CHIS 2001 that had important implications for the screener response rate involved the method of enumerating household members. Based on an experiment in the National Household Education Survey (NHES), only adult household members were enumerated in the CHIS 2001 screener. Household members under 18 were not enumerated as part of the screener but instead were enumerated later during the adult interview. See Brick and Collins (1997) for a description of the experiment. The NHES experiment showed that this type of procedure could significantly increase the screener response rate. We believe those results also apply to CHIS 2001 and the screening approach probably resulted in a higher response rate than enumerating all household members.

Another sampling decision with consequences on response rates is the sampling of persons within the household for interview. To limit the burden on the household and to increase response rates, only one adult was randomly selected from each household even when more were present. Nevertheless, the burden on households in CHIS is substantial, especially if there are children and adolescents present. In this case, up to three interviews could be requested from the same household. If the sampled adult is the person most knowledgeable about the sampled child, then the same adult is asked to complete two interviews.

The within-household interviewing requirement is related to response rates because it requires asking the same person to participate more than once and it increases the total interview length. Bogen (1996) found little experimental evidence on a relationship between the length of the interview and response rates, but her research did not cover multiple interviews within the same household. From an operational perspective and from monitoring the interviews during the conduct of CHIS 2001, there is wide agreement that the survey length and having multiple interviews within the household had a

negative effect on response rates. However, limiting the sample size to one adult per household was clearly useful in preventing lower response rates.

Perhaps the most effective and the most contentious method of increasing response rates in RDD surveys is giving monetary incentives to respondents. The option of offering monetary incentives was carefully considered for CHIS and it was decided not to do this. The literature on monetary incentives in RDD surveys is relatively recent. Singer, Van Hoewyk, and Maher (2000) discuss the issues associated with incentives in telephone surveys. Westat has conducted several experiments with monetary incentives in the NSAF, a large RDD survey that has many features similar to CHIS. See Cantor et al. (1998) for a discussion of the effect of these incentives. Since incentives were not offered in CHIS 2001, we do not discuss this topic further.

Another important procedure used to increase response rates that cannot be evaluated for CHIS 2001 is the interviewer training protocol. The interviewer training is discussed in detail in Report 2 (Data Collection Methods). In particular, that report describes training each interviewer to help them to avoid refusals. Interviewers who were allowed to do refusal conversions were also given special training before they were permitted to attempt to contact households or persons who previously refused to cooperate. The interviewers were also given special training to alert them to the cultural issues that might affect the response rates for key demographic groups such as American Indian/Alaska Native.

9.1 Advance Letter Mailing

Now, we turn to procedures used to improve response rates that have associated quantitative outcomes, beginning with the analysis of the advance mailing. We stress that the data presented are indicators of the effect, but they are not experimental data and many of the effects are confounded with other factors. For example, the advance mailing was sent to all households with addresses. We can examine the difference in response rate for those with an address and those without an address, but we cannot conclude the differences are due to the mailing. In fact, previous experimental research shows the actual effect is much smaller than the difference suggests (see Brick et al., 2000) because households with addresses are more likely to respond with or without an advance mailing. Despite this caveat, the results are informative.

Table 9-1 presents the interview level response rates for the screener, adult interview, child interview, and adolescent interview by whether or not an advance letter was mailed to the household. For each interview, the households that could be sent an advance letter had higher response rates. The differences in response rates are especially large for the screener and adult interviews. See Table 6-1 for the county-level differences by mailable status for the screener. As noted above, much of the difference must be due to the different propensities of the households to respond, irrespective of whether a letter is mailed. We hypothesize that the effect of the letter on respondents is to legitimize the survey. If this were the case, we would expect the differences in response rates for the extended interviews (adult, child, or adolescent) to be small because the screening interview should have much the same effect (if it is completed). Nevertheless, there is a large difference between the group that was mailed the letter and the group that was not mailed the letter for the adult interview. It is reasonable to conclude that the differences are largely due to the attributes of the households rather than the effect of the advance mailing.

Even if large increases in the response rates cannot be attributed to the advance letter mailing, the procedure is a relatively low-cost approach to increase response rates that has other desirable effects. For example, many respondents told interviewers that they read the letter and went to the UCLA web site to find out more about the survey before they were called. These respondents tended to be more willing to cooperate and this reduces the cost of data collection. Furthermore, the research noted above found small increases in response rates as a result of the advance mailing.

Table 9-1. Interview response rates by type of interview and advance letter

	Advance letter mailed		
Type	**Yes**	**No**	**Difference**
Screener	61.1	52.4	+8.7
Adult interview	64.8	57.6	+7.2
Child interview	87.9	86.0	+1.9
Adolescent interview	63.8	61.9	+1.9

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

9.2 Repeated Call Attempts

Another method used to increase response rates in CHIS 2001 was repeated attempts to contact households and sampled persons. In some surveys, the number of call attempts is limited in order

to reduce the cost of data collection and to complete the survey in a short time period. These surveys will more likely exclude some households and "hard-to-reach" persons that are not usually at home. The procedures implemented in CHIS 2001 allowed many attempts to severely limit the bias from this source of nonresponse.

Figure 9.1 shows the percentage of all completed interviews by the number of call attempts for both the screener and the adult interview. A call attempt is a telephone call placed to the sampled household. The counts of attempts in CHIS 2001 were associated with the type of interview. The first calls were to complete the screener interview. Once the screener was completed, additional calls to the household may have been required to complete the adult interview and still further attempts to complete the child and/or the adolescent interviews.

The patterns are similar for both the screener and adult interviews in Figure 9.1. Most interviews are completed within a few call attempts. The median number of call attempts for the completed screeners is three and for the adult is two. However, there is a long tail for both distributions. The 75th percentile of the number of completed interviews is not reached for the screener until the sixth attempt and for the adult it is the fifth attempt. If the number of call attempts were not extended to at least 10 attempts, then the response rate would decrease by over 10 percent for both types of interviews. Similarly, the response rates were increased by about 5 percentage points by allowing for more than 14 call attempts for both the screener and the adult interviews. However, the cost for these extra call attempts must be balanced against the gains in the number of completed interviews. The patterns for the child and adolescent interview are more compressed with fewer attempts needed to complete these interviews.

9.3 Refusal Conversion Attempts

An effective method of increasing response rates in an RDD survey is to recontact households and persons who refuse to participate in the initial interview and to ask them to reconsider and complete the interview. In CHIS 2001, these refusal conversion attempts were successful. Table 9-2 shows the cooperation rate (the ratio of the number completed to the number that were either completed or refused expressed as a percentage) by the level of effort and the type of interview. The percentages are based on the RDD and the supplemental samples. The cooperation rates are greater than the response rates presented earlier because other types of nonresponse such as not contacting the household or person

after multiple attempts are not included in the denominator of the rates. More details on the cooperation rates and other sources of nonresponse are discussed in Report 2 (Data Collection Methods).

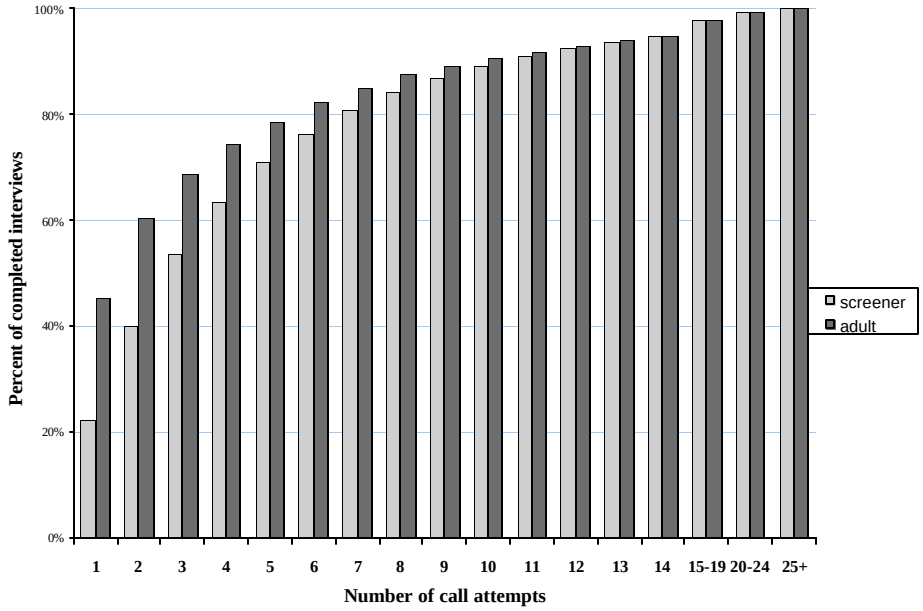

Figure 9-1. Percentage of completed screener and adult interviews by the number of call attempts

If the initial attempt to complete the screener or adult interview was not successful, then an express letter was sent prior to attempting refusal conversion. The rationale for sending the express letter is previous experimental evidence of the substantial positive effect such high visibility letters have on the conversion rates. See Brick et al. (1999). The cooperation rates in CHIS 2001 for the first refusal conversion effort for the screener interview are about 40 percent, while for the adult, child, and adolescent interview the conversion rates exceed 30 percent.

Originally, no refusal conversion attempt was planned if a parent refused permission for interviewing the sampled adolescent. Because the permission rate was relatively low, it was decided to send an express mail letter explaining the purpose and content of the adolescent interview to parents who did not give permission. It also indicated that they would be asked again if they would give permission to interview the adolescent. This conversion effort proved very successful and 26 percent of the recontacted parents gave permission.

As noted in Chapter 4, we also attempted a second refusal conversion for a subset of screener and adult interviews that were not completed in the initial or first conversion attempt and deemed eligible for another try. The results of these attempts were relatively successful. Table 9-2 shows, for example, that 70.3% of the selected adult respondents cooperated in completing their interview as a result of the initial contact. Of the balance that refused to cooperate but were eligible for a conversion attempt, 31.4% cooperated in completing their interview as a result of that one conversion attempt. Finally, of those who refused both their initial contact and their first conversion attempt, 23.2% cooperated in completing their interview as a result of a second conversion attempt. Table 9-2 shows that about one-fourth (24.7%) of the screener refusals and more than one-fifth (23.2%) of the adult interview refusals were able to be completed on a second refusal conversion attempt. This is not as high as the 39.2% and 31.4% success rates on the first attempt but certainly high enough to consider the second conversion attempt effort as productive. Although only one conversion attempt was made for the parents of children, adolescents and the permission-giving parent of adolescents, these single attempts were successful in converting 33.5% of parents of children who had refused, 44.2% of the adolescents who refused on their first contact, and 26.0% of the parents who initially refused to let us interview their adolescent.

Table 9-2. Completion rates by level of effort and type of interview

	Percent Completed at Each Level of Effort				
Level	Screener	Adult	Child	Adolescent	Adolescent permission
Initial Contact	50.7	70.3	92.5	87.0	80.4
First Conversion Attempt	39.2	31.4	33.5	44.2	26.0
Second Conversion Attempt	24.7	23.2	---	---	---

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

The refusal conversion effort greatly increased the number of completed interviews and, as a result, improved the survey response rates. Overall, considering all the interviews completed in both the RDD sample and the supplemental samples, converted interviews accounted for 26 percent of the completed screeners, 10 percent of the adult extended interviews, 2 percent of the child interviews, and 5 percent of the adolescent interviews.

As noted earlier, express letters were sent to every household and adult who refused to be interviewed if an address was available for them. The effect of the express letter can be examined by comparing the first refusal conversion rates by whether an express letter was sent, but differences cannot

be attributed to the letters because no experimental data exists. This situation is similar to the advance letter analysis because the express letter mailing is confounded with having an address.

Table 9-3 shows the first refusal conversion rates by express letter status for the RDD sample only. Overall the conversion rates for the RDD sample cases only are lower than the rates for the full sample given in Table 9-2.

The difference by express letter status for the screener interviews is about 14 percentage points. The earlier research suggests that, at most, 5 to 8 percentage points might be attributable to the express letter. Nevertheless, the express letter appears to be effective for the screener interview. The differences for the adult extended interviews is also large, nearly 10 percentage points higher for those sent an express mailing. However, the adolescent permission request rates are in the opposite direction with the loss of about 6 percentage points for those mailed an express letter. Since no previous experimental data on the effectiveness of express mailings at the extended level have been published, it is difficult to interpret the differences from the CHIS data collection.

Table 9-3. First refusal conversion rates by type of interview and express mailing status, RDD sample only

	Screener	Adult	Adolescent Permission
Express letter	42.4	33.7	21.1
No express letter	28.8	23.9	27.1

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

9.4 Proxy Reports

A special provision was made to permit other adults to respond for sampled adults who were over 65 and unable to participate because of mental or physical disabilities. No other proxy interviews were permitted in CHIS 2001 (the parents reported for children under 12 but they were designed to be the respondents who could best complete the interview). A total of 351 adult proxy interviews were done in the RDD sample. Proxy respondents had to be adult household members who were knowledgeable about the sample person's health. The proxy respondent was almost always a spouse or child of the sampled adult. While the number of interviews completed using the proxies interviews is relatively small, it does provide coverage for a group of adults with very different health characteristics that would not otherwise be included in the survey.

9.5 In-Language Interviews

A very important procedure incorporated to enhance the response rate was translating and conducting all the interviews in the language requested by the sampled person. These languages were: Spanish, Chinese (Cantonese and Mandarin), Khmer, Korean, and Vietnamese. In many cases, households that did not speak English would not have been included in CHIS had it not been for the additional languages. In some cases, the respondents would have tried to respond in English but the quality of the interviews would have been much lower if the other languages were not provided. The translation of the instruments provided a common basis for the interviewers that would not have been available otherwise.

Table 9-4 gives the number of interviews that were completed in all languages other than English. The use of languages other than English had a dramatic effect on the response rates. Over 10,000 households completed the screener using a language other than English which accounts for about 12 percent of all the completed interviews in CHIS 2001. Spanish is the most frequently used language, with about 76 percent of the non-English screeners being completed in Spanish. Vietnamese was the next most frequently used language used by respondents in the interviews, although that was primarily required for the Vietnamese supplemental sample.

Table 9-4. Number of completed interviews by language and sample type

Sample type	Non-English	Spanish	Cantonese	Mandarin	Khmer	Korean	Vietnamese
Screener							
Total	10,473	8,000	308	295	144	584	1,142
RDD	8,762	7,754	264	277	35	242	190
Santa Barbara	52	52	0	0	0	0	0
San Francisco	170	106	39	17	0	2	6
South Asian	21	21	0	0	0	0	0
Cambodian	113	0	2	0	108	0	3
Japanese	0	0	0	0	0	0	0
Korean	342	0	0	0	1	340	1
Vietnamese	943	0	0	1	0	0	942
American Indian/ Alaska Native	3	0	3	0	0	0	0
Shasta Latinos	67	67	0	0	0	0	0

Table 9-4. Number of completed interviews, by language and sample type (continued)

Sample type	Non-English	Spanish	Cantonese	Mandarin	Khmer	Korean	Vietnamese
Adult interview							
Total	6,806	5,116	232	229	123	456	650
RDD	5,770	4,959	203	217	28	223	140
Santa Barbara	36	36	0	0	0	0	0
San Francisco	111	63	29	12	0	2	5
South Asian	0	0	0	0	0	0	0
Cambodian	95	0	0	0	95	0	0
Japanese	0	0	0	0	0	0	0
Korean	231	0	0	0	0	231	0
Vietnamese	505	0	0	0	0	0	505
American Indian/ Alaska Native	0	0	0	0	0	0	0
Shasta Latinos	58	58	0	0	0	0	0
Child interview							
Total	2,789	2,358	55	42	40	126	168
RDD	2,493	2,295	51	40	7	58	42
Santa Barbara	16	16	0	0	0	0	0
San Francisco	27	16	4	2	0	1	4
South Asian	0	0	0	0	0	0	0
Cambodian	33	0	0	0	33	0	0
Japanese	0	0	0	0	0	0	0
Korean	67	0	0	0	0	67	0
Vietnamese	122	0	0	0	0	0	122
American Indian/ Alaska Native	0	0	0	0	0	0	0
Shasta Latinos	31	31	0	0	0	0	0
Adolescent interview							
Total	537	454	3	10	22	16	32
RDD	482	441	3	10	7	8	13
Santa Barbara	3	3	0	0	0	0	0
San Francisco	4	4	0	0	0	0	0
South Asian	0	0	0	0	0	0	0
Cambodian	15	0	0	0	15	0	0
Japanese	0	0	0	0	0	0	0
Korean	8	0	0	0	0	8	0
Vietnamese	19	0	0	0	0	0	19
American Indian/ Alaska Native	0	0	0	0	0	0	0
Shasta Latinos	6	6	0	0	0	0	0

Source: UCLA Center for Health Policy Research, 2001 California Health Interview Survey.

The use of a language other than English varied by the type of interview. In both the screener and the adult interviews, about 12 percent of the total completed interviews were done in languages other than English. The child interview had the largest percentage of all interviews completed in languages other than English. About 21 percent of the child interviews were done in languages other than English. The lowest percentage was for the adolescent interview, yet still nearly 9 percent of these interviews were not done in English.

The other main source of variation in the use of other languages is the type of supplemental sample. For example, 93 percent of the adult Vietnamese supplemental sample interviews were conducted in Vietnamese. (Table 9-4 shows 505 adult interviews completed in Vietnamese, and Table 1-2 shows the total of 540 adult interviews in the Vietnamese supplemental style.) Similarly, over 70 percent of the adult Korean and Cambodian adult interviews were done in those languages rather than English.

Overall the effect on the response rates of including these languages was large. As with the proxy interviews, the reduction of nonresponse bias due to the use of the multiple language interviews is probably even greater than the simple response rate computations suggest. The non-English speaking population is likely to have different characteristics with respect to health and other characteristics measured in CHIS 2001. As bias equation (1) indicates, the combination of reducing nonresponse and removing systematic differences between the respondents and nonrespondents is very effective for reducing bias.

REFERENCES

The American Association for Public Opinion Research (AAPOR). (1998). Standard definitions: Final disposition of case codes and outcome rates for RDD telephone surveys and in-person household surveys. Ann Arbor, MI: *American Association for Public Opinion Research*. (1998), http://www.aapor.org.

Council of American Survey Research Organizations (CASRO). (1983). The report of the CASRO task force on response rates, in *Improving Data Quality in a Sample Survey*, edited by F. Wiseman, Cambridge, MA: Marketing Science Institute.

Bogen, K. (1996) The effect of questionnaire length on response rates—A review of the literature. *Proceedings of the Survey Research Methods Section of the American Statistical Association*: Washington D.C, 1020-1025.

Brick, J.M., Broene, P., Cantor, D., Ferraro, D., Hankins, T., Rauch, C., and Strickler, T. (2000). *Report No. 8: 1999 NSAF Response Rates and Methods Evaluation*, Urban Institute: Washington D.C.

Brick, J.M., and Collins, M. (1997). *An Experiment in Random-Digit-Dial Screening*. U.S. Department of Education, Office of Educational Research and Improvement, NCES 98-255.

Brick, J.M., Flores-Cervantes, I., and Cantor, D. (1999). Report *No. 8: 1997 NSAF Response Rates and Methods Evaluation*, Urban Institute: Washington D.C.

Brick, J. M., Montaquila, J., and Scheuren F. (2002). Estimating Residency Rates for Undetermined Telephone Numbers, *Public Opinion Quarterly*, 66, 18-39.

Centers for Disease Control (CDC) (2001). BRFSS Summary Data Quality Report: published at www.cdc.gov/nccdphp/brfss.

Cantor, D., Allen, B., Cunningham, P., Brick, J.M., Slobasky, R., Giambo, P. and Kenney, G. (1998). Promised incentives on a RDD Survey. *Proceedings of the Eight International Workshop on Household Survey Nonresponse*, 219-228.

Curtin, R., Presser, S. and Singer, E. (2000). The Effects of Response Rate Changes on the Index of Consumer Sentiment, *Public Opinion Quarterly*, 64, 413-428.

Groves, R., and Wissoker, D. (1999). *Report No. 7: 1997 Early NSAF Nonresponse Studies*, Urban Institute: Washington D.C.

Keeter, S., Miller, C., Kohut, A. Groves, R., and Presser, S. (2000). Consequences of Reducing Nonresponse in a National Telephone Survey, *Public Opinion Quarterly*, 64, 125-148.

Lessler, J., and Kalsbeek, W. (1992). *Nonsampling Errors in Surveys*, John Wiley and Sons.

Madow, W., Nisselson, H., and Olkin, I. (1983). Incomplete Data on Sample Surveys, Volume 1 – Report and Case Studies, Academic Press.

Massey, J., O'Connor, D., and Krotki, K. (1997). Response Rates in Random Digit Dialing Telephone (RDD) Surveys, *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 707-712.

Shapiro, G., Battaglia, M., Camburn, D., Massey, J., and Tompkins, L. (1995). Calling local tele phone company business offices to determine the residential status of a wide class of unresovled telephone numbers in a random-digit-dialing sample. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 975-980.

Singer, E., Van Hoewyk, J., and Maher, M. (2000). Experiments with Incentives in Telephone Surveys. *Public Opinion Quarterly*, 64, 171-188.

The California Health Interview Survey (CHIS) is a collaboration of:

UCLA Center for
Health Policy Research

California Department
of Health Services

Public Health
Institute

California Health Interview Survey (CHIS)
UCLA Center for Health Policy Research
10911 Weyburn Avenue, Suite 300
Los Angeles, California 90024-2887
Phone 310-794-0925
Toll Free 1-866-275-2447
Fax 310-794-2686
chis@ucla.edu
www.chis.ucla.edu