

**CALIFORNIA HEALTH INTERVIEW SURVEY**

**CHIS 2001 METHODOLOGY SERIES**

**TECHNICAL REPORT**

**REVISED CHIS 2001 WEIGHTS**

**January 2005**

*This report was prepared for the California Health Interview Survey by Ismael Flores-Cervantes, Brandon Barret and J. Michael Brick of Westat, Inc.*



*www.chis.ucla.edu*

This report provides analysts with information about the revised CHIS 2001 weights. .

**Suggested citation:**

California Health Interview Survey. *CHIS 2001 Methodology Series: Revised CHIS 2001 Weights*. Los Angeles, CA: UCLA Center for Health Policy Research, 2005.

Copyright © 2005 by the Regents of the University of California.

The California Health Interview Survey is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Funding for CHIS 2003 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

*Revised CHIS 2001 Weights* is the last in a series of methodological reports describing the 2001 California Health Interview Survey (CHIS 2001).

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. The 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.

This report describes the procedures used to revise the CHIS 2001 weights. The CHIS 2001 weights were revised to be consistent with the methodology used for weighting CHIS 2003. The main difference in the procedures used in the two surveys was the source and nature of the population control totals used in raking. The CHIS 2003 weights were raked to control totals derived from the population projections from the California Department of Finance (DOF). The original CHIS 2001 weights were raked to control totals derived from Census 2000 data. The main benefit of the revision of the weights for 2001 will be in comparing estimates from the 2003 and 2001 surveys. The estimates of differences between the estimates for the two years will be more accurate because the weighting methods for the two are consistent. Naturally, some 2001 estimates using the revised weights will differ from the estimates computed using the original weights.

The primary purpose of this report is to provide analysts information about the methodology used to revise the CHIS 2001 weights. This report also provides a discussion on differences in the estimates using the revised and original weights.

## TABLE OF CONTENTS

	<u>Page</u>
REVISED CHIS 2001 WEIGHTS .....	1
1. Original CHIS 2001 weights .....	1
2. Reweighting the CHIS 2001 sample.....	1
3. Weighting Scheme.....	7
4. Imputation procedures .....	11
5. Control totals .....	11
6. Notes about the revised CHIS 2001 weights .....	12
Appendix A .....	13

## REVISED CHIS 2001 WEIGHTS

This technical report describes the methods used to produce revised weights for CHIS 2001. One difference between the original and revised weights is that the revised weights include additional observations that were not included with the original weights. The revised weights include the Korean and Vietnamese list sample cases that originally were not combined with the Random Digit Dialing (RDD) sample. In 2003 the list samples of Koreans and Vietnamese were included in the weighting procedure with the RDD sample cases and the revision of the 2001 weights accomplishes this same goal. The second difference between the original and revised weights is the one that has the greatest effect on the estimates themselves. The source of the raking population control totals was changed along with some of the definitions of the control totals. These issues are discussed in more detail in this report.

### 1. Original CHIS 2001 weights

The original CHIS 2001 delivery files included eight sets of weights for adult, child and adolescent interviews. The set with the largest number of records corresponds to the RDD sample<sup>1</sup>. The remaining sets of weights were created for specific race-ethnic groups (American Indian/Alaska Native, Cambodian, South Asian, Japanese, Korean, Vietnamese, and Shasta Latinos) for whom the sample size was increased by sampling persons from lists. As a result, these sets are sometimes called the list samples. However, in this report we refer to these weights as RDD-LIST weights because they include records from the RDD and list samples for persons of the specified race-ethnicity (some records in the RDD-LIST were sampled from the RDD and appear in the RDD file but with different weights than in the RDD-LIST file).

The CHIS 2001 RDD weights are used to produce estimates for the total population of California. The RDD weights could also be used to produce estimates for the race-ethnic groups mentioned above, but these estimates have a larger variation compared to the estimates computed using the corresponding RDD-LIST weights. The RDD-LIST weights were specifically created for these race-ethnic groups.

The methodology used to create the weights for the RDD and the RDD-LIST sample is described in the CHIS 2001 Methodology Report 5 (Weighting and Variance Estimation). The RDD

---

<sup>1</sup> The RDD sample included the supplemental geographic samples in San Francisco County and Santa Barbara County.

sample was raked to 14 dimensions for the total population in California, four of which include the total number of persons in some race-ethnic groups in California. The control totals were derived using the Census 2000 files.

## **2. Reweighting the CHIS 2001 sample**

Due to significant differences in the weighting methodology for CHIS 2003 and the original CHIS 2001 sample, it was decided that the weights for CHIS 2001 should be revised to follow the CHIS 2003 weighting methodology as closely as possible. The revised CHIS 2001 sample permits comparing estimates over time with the CHIS 2003 data, where appropriate, more accurately than is possible with the original weights. When the original weights are used to compare estimates between 2001 and 2003, a part of the difference may be attributed to the weighting methodology. The revised weights eliminate this difference.

There are three differences in the methodology used to create the weights for CHIS 2003 and the original CHIS 2001 that may affect comparisons between these surveys:

- The source of the control totals,
- The definition of raking dimensions, and
- The creation of a single set of weights for the RDD and list samples.

The main source for the control totals used for CHIS 2003 was the 2003 California Department of Finance (DOF) Population Projections. The DOF provides population estimates at the county level by race, ethnicity, gender and single age for each year from 2000 to 2050. The DOF uses a baseline cohort-component method to project population estimates based on fertility/mortality rates and life expectancy by different race-ethnic groups and age cohorts. The most significant advantage of using the DOF files is that they provide the most up-to-date estimates, and they can be used not only as the source of control totals for CHIS 2003 but also for future rounds of CHIS. A disadvantage is that the population living in group quarters—not eligible to participate in the CHIS survey—has to be estimated and removed from the population counts. As noted earlier, the original 2001 weights were based on Census 2000.

The second difference between the original CHIS 2001 and CHIS 2003 weighting procedures is the definition of the raking dimensions. The 11 dimensions used in CHIS 2001 are shown in Table A-1 in Appendix A, and Table A-2 shows the raking dimensions used in CHIS 2003. Although many of the CHIS 2001 dimensions are similar to the dimensions used in CHIS 2003, there are significant differences, especially in the definitions of the cells that use race. The first 8 dimensions in CHIS 2003 (Table 7-2) were created by combining demographic variables (age, sex, race, and ethnicity) and different geographic areas (city, county, region or group of counties, and state). The 9<sup>th</sup>, 10<sup>th</sup>, and 11<sup>th</sup> dimensions used socio-economic variables. The 11<sup>th</sup> dimension in both CHIS 2001 and 2003 was specifically created to adjust the weights for households without a landline telephone.

Although some of the dimensions were defined in the same way in 2001 and 2003, some dimensions used additional variables and others were removed in CHIS 2003. Dimensions defined using age and gender in CHIS 2001 matched those used in CHIS 2003 (i.e., the 1<sup>st</sup>, 2<sup>nd</sup>, and 6<sup>th</sup> in CHIS 2001 correspond to the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> dimensions in CHIS 2003). In CHIS 2003, socio-economic variables were used in several dimensions: adult's education attainment for the 9<sup>th</sup> dimension, household structure (number of adults living in the household) for the 10<sup>th</sup> dimension, and household tenure in dimension 11. In CHIS 2001, the socio-economic variables (AFDC participation, public housing assistance, and household structure, i.e., number of children and number of adults in the households) were used in combination in one dimension (11<sup>th</sup> dimension or nontelephone dimension). Several analyses of the CHIS 2001 data suggested that the estimates could be improved if additional raking dimensions for more socio-economic variables were used to control the population estimates by these variables. In CHIS 2003, household tenure (own or rent home) was identified as a potentially important auxiliary variable. Household tenure was also the most significant for adjusting for the presence or absence of a landline telephone. To avoid redundancy in the definition of the dimensions, household tenure was included only in the nontelephone dimension (11<sup>th</sup> dimension) in 2003. In CHIS 2001, the question about household tenure was not asked. As a result, this variable cannot be included in the nontelephone dimension or as a separate dimension in reweighting 2001. Therefore, the original CHIS 2001 nontelephone dimension was used. In other words, one of the socioeconomic dimensions—household tenure—used in CHIS 2003 was not included as a raking dimension in the revised weights for CHIS 2001. Table 1 shows the nontelephone dimension used in CHIS 2001.

Table 1 Nontelephone adjustment cell definition for CHIS 2001

Person type	Cell	In household with AFDC participant or receiving public housing assistance	Number of children in the household	Number of adults in the household	Race/ethnicity	
Adult	4	Yes	0 or 1		Latino or Black non-Latino Other non-Latino	
Adult	5	Yes	2 or more			
Adult	10	No				
Adult	11	No				
Child	1	Yes	1 or 2			
Child	2	Yes	3 or more			
Child	6	No				Latino
Child	7	No				Non-Latino
Teen	3	Yes				
Teen	8	No				0 or 1
Teen	9	No				2 or more

The 3<sup>rd</sup> (American Indian/Alaska Native), 4<sup>th</sup> (Asian), 5<sup>th</sup> (Latino), 6<sup>th</sup> (African American), 7<sup>th</sup> (White), 8<sup>th</sup> (Other race), and 9<sup>th</sup> (Pacific Islander or Native Hawaiian) dimensions in CHIS 2001 were replaced by the 5<sup>th</sup> dimension in CHIS 2003 with the race-ethnicity levels shown in Table 2. The race levels used in CHIS 2003 were based on the Office of Management and Budget (OMB) race definition, also known as “modified” race. The main difference is that there are no separate population counts for “other” race. The DOF population projections comply with the U.S. Office of Management and Budget 1997 revised standards for collection, tabulation and presentation of federal data on race and ethnicity<sup>2</sup>. The revised OMB standards identify only five main racial categories (White, Black or African American, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander) and combinations of these categories. However, in CHIS (both in 2001 and 2003), respondents who could not identify themselves as any of the five OMB race categories could answer a sixth race category “some other race.” This approach is consistent with the 2000 Census method and the counts from the 2000 Census include “other” race as a separate group.

<sup>2</sup> Office of Management and Budget (1997). “Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity”, Vol. 62, No. 210, Thursday, October 30, 1997.



In order to use the DOF projections, any sampled person who reported other race (alone or in combination with another race) had to be recoded into one or more of the OMB categories. This implied the imputation of a OMB race category for the 12.2 percent of adult respondents who self-reported “other race” only. As an alternative, a variable that combines Latino ethnicity and the OMB race variable was proposed and approved for CHIS 2003 and this approach reduces the number of imputations needed. The recoding includes an additional level that groups Latinos of any race as shown in Table 2. Because most respondents who self-reported other race only were Latinos, the number of imputed records was reduced significantly to the 0.2 percent of the sample who self-reported as non-Latino and other race. The advantage of this classification is that it matches the categories of the population projections available in the DOF files.

Table 2. Description of the variable using the OMB race definition

OMBSRREO	Definition
1	Latino
2	White non-Latino
3	African American non-Latino
4	American Indian Alaskan Native non-Latino
5	Asian non-Latino
6	Pacific Islander Native Hawaiian non-Latino
7	Two or more races non-Latino

CHIS 2003 included a dimension (4<sup>th</sup> dimension) defined by Special Planning Areas (SPAs) in Los Angeles, and cities that were oversampled in Alameda County. There was no special dimension defined at the SPA level in CHIS 2001. However, raking dimensions defined at the self-reported stratum level included separate levels for cities that were oversampled in CHIS 2001 (i.e., Long Beach and Pasadena in Los Angeles County and Berkeley in Alameda County).

A third difference in the weighting methodologies was the creation of one set of weights for the combined RDD and Korean and Vietnamese surname list samples in CHIS 2003. Only Korean and Vietnamese surname list samples were included in CHIS 2003. In the weighting procedures for CHIS 2001, separate weights were created for all the surname list groups, including the Korean and Vietnamese lists.

### 3. Revised Weighting Approaches

After the CHIS 2003 weights were completed, two options for reweighting the CHIS 2001 data were considered. These weighting options shared the following features:

- The use of the same raking dimensions as in CHIS 2003, to the extent possible. The control totals would be derived using the 2001 DOF files using the procedures used for the CHIS 2003 totals.
- The use of OMB races for persons who self-reported non-Hispanic "other" race. The new dimensions include OMB race groups for non-Hispanics. As a result, an OMB race had to be imputed for approximately 130 records in CHIS 2001.
- The inclusion of records from the RDD, Korean (KR) and Vietnamese (VT) samples for the file containing the revised CHIS 2001 weights. Other CHIS 2001 race-ethnic list samples were not included.

In the first option the same weighting steps used in CHIS 2003 would be repeated for the creation of the CHIS 2001 revised weights. The implementation of this option depended on the availability of additional information from the list frames from the sampling vendor, Genesys. This information was required in order to compute the multiple probabilities of selection of the telephones in the RDD, Korean, and Vietnamese frames. Because this option required repeating all weighting adjustments (including the creation of a new base weight), it was relatively expensive and time-consuming.

The second option was to instead further adjust the 2001 weights that had already been created originally. This option was less expensive, did not require additional data for the selection probabilities, and could be implemented more quickly. This option was selected for creating the revised CHIS 2001 weights.

The first step in the approved approach was to combine the RDD samples with the Korean and Vietnamese list samples by replacing the records (including weights) identified as Korean (KR) and Vietnamese (VT) in the RDD sample with those from the combined KR/VT RDD-LIST files. We explored three ways of defining the groups to be replaced: (1) self reported KR/VT alone, (2) self reported KR/VT alone or in combination with another race, or (3) the subset of (2) who reported single

race or most identified as KR/VT. Table 2 shows the number of records available for the creation of the revised CHIS 2001 data files using these three definitions.

The CHIS 2001 delivery files for the Korean and Vietnamese RDD-LIST samples included only records for persons who primarily self-reported Korean (or Vietnamese), i.e., variable BESTRACE= KR or VT. The original files used to create the Korean and Vietnamese RDD-LIST weights also contained records for all persons who self-reported Korean (or Vietnamese) alone or combined with another race. These original Korean and Vietnamese files were used as the starting point for the creation of the new weights. Sums of weights or population estimates were computed for the total number of persons who reported Korean (or Vietnamese) alone, and total number of persons who reported Korean (or Vietnamese) alone or in combination with other race. These estimates were compared to the same estimates produced using the RDD sample to assess differences in the estimates. Because the 7<sup>th</sup> raking dimension in CHIS 2003 was defined as Asian ethnic groups alone (including Korean and Vietnamese alone), it was decided to use the records for person who self-reported as Korean (or Vietnamese) alone for the replacement.

Table 2. Number of records in the CHIS 2001 Samples

CHIS 2001 Type	Race-ethnic	Sample Size
Original records delivered to UCLA		
RDD	All	73,824
RDD-LIST	Korean (Best race)	1,064
	RDD- Korean (Best race)	614
	LIST- Korean (Best race)	450
RDD-LIST	Vietnamese (Best race)	1,121
	RDD- Vietnamese (Best race)	426
	LIST- Vietnamese (Best race)	695
Available record for CHIS 2001 reweighting		
RDD	All	73,824
RDD-LIST	Korean alone or combined	1,091
	RDD- Korean alone or combined	641
	LIST- Korean alone or combined	450
RDD-LIST	Vietnamese alone or combined	1,169
	RDD- Vietnamese alone or combined	472

	LIST- Vietnamese alone or combined	697
Processed records for CHIS 2001 reweighting		
RDD	All	73,820 <sup>a</sup>
RDD-LIST	Korean alone race	1,056
	RDD- Korean alone	616
	LIST- Korean alone	446 <sup>b</sup>
RDD-LIST	Vietnamese alone	1,062
	RDD- Vietnamese alone	401
	LIST- Vietnamese alone	668 <sup>c</sup>

<sup>a</sup> Four households in the RDD sample were replaced by the households in the RDD-LIST sample. However, two children and two adolescents in these households did not have RDD-LIST weights. As a result, these records were excluded from the new CHIS 2001 weight file.

<sup>b</sup> Includes 6 records for children/adolescents who reported Korean combined

<sup>c</sup> Includes 7 records for children/adolescents who reported Vietnamese combined

### 3. Weighting Scheme

Define  $\hat{T}$  as the estimate of the total population in California as

$$\hat{T} = \sum_{h=1}^H \sum_{p=1}^{h_p} w_{hp} ,$$

where  $w_{hp}$  is the RDD weight for person  $p$  in household  $h$ . The sum,  $\hat{T}$ , can be split into two totals as

$$\hat{T} = \hat{T}^G + \hat{T}^{\bar{G}} = \sum_{h \in G} \sum_{p=1}^{h_p} w_{hp} + \sum_{h \in \bar{G}} \sum_{p=1}^{h_p} w_{hp} ,$$

where  $\hat{T}^G$  is the total number of persons in households where the sample adult is a member of race group  $G$ ,  $\hat{T}^{\bar{G}}$  is the total number of persons in households where the sample adult is not a member of the race group  $G$ .

The RDD-LIST weights can be used to estimate the total number of persons in households where the adult is a member of group  $G$  as

$$\hat{T}^{*G} = \sum_{h \in G} \sum_{p=1}^{h_p} w_{hp}^*,$$

where  $w_{hp}^*$  is the RDD-LIST weight for person  $p$  in household  $h$ . Note that  $\hat{T}^{*G}$  and  $\hat{T}^G$  estimate the same quantity (number of persons in household where the sample adult belongs to group  $G$ ).

This total can be split into two totals depending on the source of the record as

$$\hat{T}^{*G} = \hat{T}_{RDD}^{*G} + \hat{T}_{LIST}^{*G} = \sum_{h \in RDD} \sum_{p=1}^{h_p} w_{hp}^* + \sum_{h \in LIST} \sum_{p=1}^{h_p} w_{hp}^*,$$

where  $\hat{T}_{RDD}^{*G}$  includes only the records from the RDD and  $\hat{T}_{LIST}^{*G}$  include the records from the list sample. Note that the records in  $\hat{T}_{RDD}^{*G}$  are the same records used in  $\hat{T}^G$ . Substituting these totals

$$\hat{T} = \hat{T}^G + \hat{T}^{*G} = \hat{T}^G + \hat{T}_{RDD}^{*G} + \hat{T}_{LIST}^{*G} = \sum_{h \in RDD} \sum_{p=1}^{h_p} \hat{w}_{hp} + \sum_{h \in LIST} \sum_{p=1}^{h_p} w_{hp}^*,$$

where the weight  $\hat{w}_{hp}$  is defined as

$$\hat{w}_{hp} = (1 - \delta_h(G))w_{hp} + \delta_h(G) \cdot w_{hp}^*,$$

and  $\delta_h(G) = 1$  if the sample adult in household  $h$  belongs to group  $G$ , and zero otherwise. This result can be extended to include additional non-overlapping groups  $G$ , (i.e., Korean and Vietnamese).

The replacement of records was done at the primary sampling unit (PSU) or household level. Households where the adult self-reported Korean (or Vietnamese) alone were flagged in both the RDD and RDD-LIST (Korean and Vietnamese) samples. Weights of members of the households flagged in the RDD sample were replaced by the corresponding weights from the RDD-LIST samples. Records and weights of members that were not found in the RDD sample were flagged in the RDD-LIST samples and then physically added to the file. After replacing and adding the list records, the file contained 74,934 records.

In order to create the revised CHIS 2001 weights, these records were raked to population control totals to produce estimates consistent with the 2001 California Department of Finance (DOF)

Population Projections. As mentioned before, raking was used in both the RDD sample and RDD-LIST samples in CHIS 2001. The new raked weight,  $RAKEDW_i$ , can be expressed as

$$RAKEDW_i = WGT_i \cdot \prod_{k=1}^K RAKEDF_{k_l},$$

where  $RAKEDF_{k_l}$  is the raking factor for dimension  $k$ , level  $l$  to which the person  $i$  is assigned. For example, if dimension  $k=4$  is sex with two levels (male  $l=1$  and female  $l=2$ ), then the raking factor for this dimension is  $RAKEDF_{4_1}$  if the person  $i$  is male. The raking factors are derived such that the following relationship holds for every raking dimension  $k$ , and level  $l$ ,

$$CNT_{k_l} = \sum_i \delta(k_l)_i \cdot RAKEDW_i,$$

where  $CNT_{k_l}$  is the control total, and  $\delta(k_l)_i = 1$  if the adult  $i$  is in level  $l$  of dimension  $k$  and zero otherwise. Marginal raking dimension cells with less than 50 respondents were collapsed to nearby cells.

To determine which set of weights ( $WGT_i$ ) would be raked to the new control totals, we explored two options. In both of these options, the original CHIS 2001 RDD-LIST design-based raked weights were used for the combined sample cases<sup>3</sup>. For RDD cases, we tested the original person nonresponse adjusted weights (before raking) and the final 2001 raked weight. We found that there was less variability in the revised CHIS 2001 weights if the weights before raking were used. When the original final 2001 RDD weights were raked to DOF totals, some very large raking factors were found due to original large raking factors being multiplied by new large raking factors using the new definitions of the raking dimensions.

After raking, we examined the distribution of the weights to determine if there were very large weights that could have a large effect on either the estimates or the variances of the estimates. When observations with large weights were found, the weights for these cases were trimmed. Trimming reduces the weight and the influence of the observation on the estimates and their variances. Trimming factors  $TFACT_i$  computed as

---

<sup>3</sup> In CHIS 2001, two weights were produced for the RDD-LIST samples. The first weight, the design based weight, was created using the probability of selection of the household. The second weight, or model based weight, ignored this probability of selection.

$$TFACT_i = \begin{cases} 1 & \text{if the weight } i \text{ is not trimmed} \\ t_i & \text{otherwise} \end{cases}$$

where  $0 < t_i < 1$ . The before-raking weights were multiplied by the trimming factors. The trimmed weights were then re-raked and their distribution was reexamined. If additional trimming was needed, the trimmed and raked weights were discarded and new trimming and weighting were undertaken. Candidates for trimming were identified separately by type of interview (adult, child, and adolescent) using the same methods in CHIS 2003 (See Section 4.3 of the CHIS 2003 Methodology Report 5: Weighting and Variance Estimation).

#### **4. Imputation procedures**

In order to create the raking dimensions for the revised CHIS 2001 weights, imputations beyond those already done for CHIS 2001 had to be undertaken. Specifically, the three variables for adult's education attainment (variable SREDUC), OMB race-ethnicity (variable OMBSRREO) and Asian ethnic groups (variable OMBSRASO) had missing values and were imputed.

The procedures used to impute these variables are the same to those used in CHIS 2003. Section 8.3 of the CHIS 2003 Method Report 5 give the details for the creation and imputation of the self-reported education variable SREDUC.

The DOF control totals are defined in terms of OMB race categories for raking dimensions 5, 6, and 8. Persons who reported themselves as Latino "some other race" were assigned an OMB race following procedures similar to those used by the Census Bureau. Since the OMB assignment is done using the imputed regular race variables when imputing for the variable OMBSRREO, all sampled persons had nonmissing race values (the race variables are SRW, SRAA, SRAI, SRAS, SRPI, and SRO). All these variables were previously imputed for the original CHIS 2001 weights. Section 8.4.2 of the CHIS 2003 Method Report 5 gives the additional details for creating and imputing the variable OMBSRREO.

The person weights were raked using a dimension defined for the Asian groups (dimension 7). The variable OMBSRASO with the Asian ethnic groups was imputed for all persons where the variable OMBSRREO=4 (non-Latino Asian alone) using five flags indicating the Asian ethnic groups of

the respondent. See Section 8.4.3 of the CHIS 2003 Method Report 5 for details on the creation and imputation of the variable OMBSRASO.

## **5. Control totals**

The control totals for the revised CHIS 2001 weighting were derived following the same procedures used in CHIS 2003. The only difference is that the source file is the 2001 DOF file. See Section 7.2 CHIS 2003 Method Report 5 for details for the creation of dimensions 1, 2, 3, 5, 6, and 8.

The creation of dimension 4, defined by SPA in Los Angeles County and Berkeley in Alameda County, used information from the 2000 Census Summary File 1 (SF1). The Los Angeles County Department of Health (LACDH) produced a listing of Census tracts by SPAs. The 2000 SF1 file was used to compute the percentages of the population in the SPAs by aggregating population counts in the Census tracts. These percentages were applied to the total 2001 DOF population total (after the population in group quarters was excluded) to produce the controls for dimension 4. For Berkeley, the same SF1 file was used to compute the percentages of 2000 population of Alameda County in Berkeley. The percentages were applied to the 2001 DOF Alameda County population total (excluding group quarters). As in the previous dimension, it was assumed that there were no changes in the percentage of the population living in the SPAs and Berkeley in Alameda County between 2000 and 2001.

For dimensions 9 (adult's education attainment) and 10 (number of adults in the household), the percentages of the population were computed using the 2001 American Community Survey public use microdata file and then applied to the 2001 DOF population total (excluding group quarters). The assumption was that there were no changes in the distribution of the population between 2002 and 2001 for this variable.

## **6. Computing estimates using the revised weights**

The primary goal of revising the weights for CHIS 2001 was to make estimates of change between 2001 and 2003 more precise. A direct consequence of this approach is that it is now possible to produce two estimates from the 2001 CHIS data files, one using the original weights and another using the revised weights. Estimates computed using the different weights and data sets will differ. In this section we provide some guidance on the magnitude of the differences that can be expected using the



different data sets and weights, and the circumstances in which these differences are likely to be relatively large. Because the estimates depend on both the analysis variable and the populations being analyzed, it is difficult to make sweeping generalizations about the differences. However, in most cases the differences are the result of changes of the definition of the raking dimension and the control totals and this can be used to predict the types of estimates that are likely to differ substantial.

In general, the differences between the two data sets and weights are small for estimates of proportion and somewhat larger for estimates of totals. Small differences are expected for characteristics correlated to raking dimensions that were common to the original and revised weights procedures (i.e., age, gender, self-reported strata). On the other hand, larger differences may be expected for characteristics correlated to raking dimensions not used in the original CHIS 2001 but incorporated in the revised weights. In particular, estimates of the number of persons by race are those that are most likely to be different because this raking variable was very different in the original and revised weighting procedures.

A full evaluation of the relative accuracy of the estimates using the original and revised weights for particular CHIS data items would require having an external source of the estimates. No such evaluation has been conducted at this time. Thus, no claims are presented indicating one or the other set of estimates is “better” than the other in this sense, although for estimates of change between 2001 and 2003 the revised 2001 weights are clearly better. Below, we compare estimates of proportions and totals for the entire state for 2001 produced from the two data sets and weights for that year. We then consider estimates by race where larger differences are expected and conclude with some estimates for specific geographic areas that might also have large differences because the raking dimensions for the two weights are different for the areas.

For the general estimates of proportions and totals, we computed 29 estimates for selected categorical variables (data items) from the adult, adolescent, and child interviews. These estimates and their standard errors are listed in Table B-1 in Appendix B. For the whole population, the average of the difference of the estimated proportions computed using the revised and original weights is -0.04 percentage points. Since the average might be small due to large positive and negative differences, we also computed the average of the absolute difference between the proportions. The average of the absolute differences is 0.42 percentage points. These averages suggested that proportions computed using the revised weights for these 29 estimates are very similar to those produced using the original weights. The average of the ratios of the standard errors of the estimates was 1.015, i.e., the revised estimates tended to

have a slightly larger standard error. This increase was expected due to the inclusion of the Korean and Vietnamese list records in the RDD file (the estimates for Koreans and Vietnamese is decreased by this inclusion, but the estimates over all races are slightly less precise).

The comparison for estimates for totals for the data items show larger differences, as expected because the original CHIS 2001 weights were raked to the Census 2000 population (2001 totals were not available) and the revised weights used the 2001 DOF projections. Thus, the revised weights reflect the population growth in 2001 as estimated by the DOF. The average of the ratios of totals computed using the revised weight to totals computed using the original weight was 1.033. The differences between the original and revised weight are almost entirely due to the population growth that was not captured in the original weights. The average of the ratios of the standard errors of totals for the 29 estimates was 1.043, consistent with the larger variation found in the proportion estimates with the revised weights.

As noted several times, the most significant contributor to the differences in estimates between the original and revised weights is the change in the definition of the race and ethnicity raking dimensions. As described before, control totals for non-OMB race groups alone or combined (i.e., White alone or in combination with other race, etc.) were controlled separately in 6 of the raking dimensions (the 3<sup>rd</sup>, 4<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup>, 8<sup>th</sup>, and 9<sup>th</sup> dimensions) in the original weights. In contrast, ethnicity and OMB race groups are partially controlled (i.e., non-Latino white alone, etc) in 3 of the raking dimensions (the 5<sup>th</sup>, 6<sup>th</sup>, and 8<sup>th</sup> dimensions) in the revised weights. To evaluate the effect of the new definitions and dimensions, several race estimates were computed and compared. For these tabulations, we include Census 2000 population totals and proportions for reference. Table 3 and 4 shows these estimates of totals and proportions respectively by non-OMB race groups (alone or combined with other race) using the original and revised weights. Because the original raking dimensions were defined for totals of race alone or in combination with one or more other races, these tables show the effect on the estimates for these groups of collapsing the 5 race and one ethnicity dimensions into 3 dimensions with combined race and ethnicity levels.

The largest difference for the estimates in the tables is for persons who self-report as American Indian or Alaska Native (AIAN) alone or in combination with other race. The estimate using the revised weights is more than twice the estimate using the original CHIS 2001 weights. The estimate using the original weights for the number of AIAN is very similar to the Census 2000 total. However, the correspondence was achieved in the original weights by allowing an average overall adult raking adjustment for AIAN was 0.38, by far the lowest raking factor for any large group. The factor was

necessitated because a large number of persons self-reported as AIAN in CHIS 2001. Although there are various reasons that might explain this outcome at least in part, it is likely that an important reason is that the way the race question was asked in CHIS (prompting for "any other race?") encouraged more respondents to report multiple races in addition to AIAN. Because the original before-raking weight was used in the revised weights (the weight prior to the average factor of 0.38 for the AIAN adults); the sum of weights for the AIAN used in the revised weights were much larger than the original final raked weights. In the revised raking, the raking dimension only controlled for the non-Latino AIAN total, and the Latino AIAN were not treated as AIAN (there were in the Latino, any race group). Since a large fraction of the adults who reported being AIAN also reported being Latino, the overall raking adjustment for all adults who were AIAN (i.e., Latino and non-Latino AIAN) was much larger than in the original weighting. The tables show the consequences for this group.

The tables show that the differences for the other race groups are much smaller, but still substantial in some cases. The total for White alone or combined with one or more other races decreases by 3.1 percentage points due to the revision. At least part of this change is likely to be the result of the DOF projection for the number of Whites in 2001 in California. Small changes are also present for Native Hawaiian or Other Pacific Islander (NHPI) alone or combined with one or more other races and for the other races alone or combined with one or more other races.

Table 3. Census 2000, original CHIS 2001 and revised CHIS 2001 estimates of totals for race alone or in combination with one or more other races

Race	Census 2000*	Original CHIS 2001		Revised CHIS 2001	
	Total	Total	SE Total	Total	SE Total
White alone or combined	21,490,973	20,984,429	3,036	20,496,664	15,470
African American alone or combined	2,513,041	2,367,887	1,463	2,430,541	16,975
American Indian/Alaska Native alone or combined	627,562	610,168	1,694	1,307,698	6,962
Asian alone or combined	4,155,685	4,090,447	1,622	4,267,773	17,764
Native Hawaiian or Other Pacific Islander alone or combined	221,458	213,637	1,455	296,781	6,356
Other races alone or combined	6,575,625	6,447,099	2,345	6,244,409	9,407

\* Total includes population in group quarters

Table 4. Census 2000, original CHIS 2001 and revised CHIS 2001 estimates of percentages for race alone or in combination with one or more other races

Race	Census 2000*	Original CHIS 2001		Revised CHIS 2001	
	Percentage	Percentage	SE Percent	Percentage	SE Percent
White alone or combined	63.4	63.5	0.009	60.4	0.188
African American alone or combined	7.4	7.2	0.004	7.3	0.058
American Indian/Alaska Native alone or combined	1.9	1.8	0.005	3.8	0.102
Asian alone or combined	12.3	12.4	0.005	12.5	0.049
Native Hawaiian or Other Pacific Islander alone or combined	0.7	0.6	0.004	0.8	0.039
Other race alone or combined	19.4	19.5	0.007	18.5	0.179

\* Total includes population in group quarters

Estimates of totals and percentages were also computed for non-OMB race groups defined by single race (i.e. White alone, African American alone, etc.). These definitions for the race groups do not correspond directly to the revised 2001 raking dimensions, but the definitions are commonly used for comparisons with other surveys and other data sources. Tables 5 and 6 show the estimates and percentages by race groups defined by single race. These tables show that the original AIAN alone total is about half of the Census 2000 total, despite the fact that the original total for AIAN alone or combined with other race is close to the Census figure. In other words, in the original CHIS 2001 weighting the raking adjustment under-adjusted the total for AIAN alone. In contrast, the revised 2001 weights give an estimate that is more than twice the 2000 Census total.

Tables 5 and 6 also show differences for other races. The revised proportion of persons who self-reported as White alone is 1.7 percentage points lower than the estimate from the original weights. Because one of the main objectives of the revised weights was to include the Korean and Vietnamese list respondents in the file, we also computed estimates of totals and proportions for these groups. The revised estimates are closer to the totals from the Census 2000 with much smaller standard errors.

Table 5. Census 2000, original CHIS 2001 and revised CHIS 2001 estimates of totals for race alone or in combination with one or more other races

Race	Census 2000* Total	Original CHIS 2001		Revised CHIS 2001	
		Total	Standard Error Total	Total	Standard Error -Total
White alone	20,170,059	19,630,224	15,470	19,551,054	63,075
African American Alone	2,263,882	1,897,310	16,975	2,279,601	14,662
American Indian/Alaska Native alone	333,346	173,503	6,962	703,857	27,571
Asian Alone	3,697,513	3,340,857	17,764	3,912,073	11,051
Korean	345,882	389,833	18,600	363,787	3,844
Vietnamese	447,032	378,453	22,633	468,445	1,487
Native Hawaiian or Other Pacific Islander alone	116,961	92,570	6,356	186,234	9,755
Other Alone	5,682,241	6,373,003	9,407	6,215,818	59,922
Two or more races	1,607,646	1,544,428	10,395	1,064,584	22,817
Total	33,871,648	33,051,894	0	33,913,222	0

\* Total includes population in group quarters

Table 6. Census 2000, Original CHIS 2001 and Revised CHIS 2001 estimates of totals for race alone

Race	Census 2000* Percentage	Original CHIS 2001		Revised CHIS 2001	
		Percentage	Standard Error Percent	Percentage	Standard Error Percent
White alone	59.5%	59.4	0.0468	57.7	0.1860
African American Alone	6.7%	5.7	0.0514	6.7	0.0432
American Indian/Alaska Native alone	1.0%	0.5	0.0211	2.1	0.0813
Asian Alone	10.9%	10.1	0.0537	11.5	0.0326
Korean	1.0%	1.2	0.0563	1.1	0.0113
Vietnamese	1.3%	1.1	0.0685	1.4	0.0044
Native Hawaiian or Other Pacific Islander alone	0.3%	0.3	0.0192	0.5	0.0288
Other Alone	16.8%	19.3	0.0285	18.3	0.1767
Two or more races	4.7%	4.7	0.0315	3.1	0.0673
Total	100.0%	100.0	0.0000	100.0	0.0000

\* Total includes population in group quarters

Since the revised weights are controlled to DOF population projections, Tables 7 and 8 compare the estimates from the original and revised weights to the 2001 DOF totals by OMB race-ethnic groups. These tables show that the revised estimates are closer to the DOF totals than the original estimates. This was expected because the DOF total and the OMB race-ethnic groups were used to produce the revised weights. A comparison of estimates of total and proportions by OMB race totals is not possible because all Latinos can not be assigned into the 5 OMB race groups in the CHIS 2001 data files.

Table 7. DOF 2001, Original CHIS 2001 and Revised CHIS 2001 estimates of totals OMB race-ethnicity alone

OMB Race-ethnicity	DOF 2001*	Original CHIS 2001		Revised CHIS 2001	
	Total	Total	Standard Error Total	Total	Standard Error Total
Latino	11,557,941.0	10,774,044	4,413	11,380,586	0
White non-Latino alone	16,029,331	15,931,533	28,173	15,603,435	0
African American non-Latino alone	2,275,847	1,770,685	18,974	2,141,059	75
American Indian/Alaska Native non-Latino alone	217,004	53,325	3,622	210,296	13
Asian non-Latino alone	3,860,687	3,233,339	19,007	3,802,558	52
Native Hawaiian or Other Pacific Islander non-Latino alone	116,337	62,261	5,435	113,932	2
Two or more non-Latino	672,593	1,223,094	21,432	661,355	12
<b>Total</b>	<b>34,729,740</b>	<b>33,048,282</b>	<b>1,534</b>	<b>33,913,222</b>	<b>0</b>

\* Total includes population in group quarters

Table 8. DOF 2001, Original CHIS 2001 and Revised CHIS 2001 estimates of totals for race alone

OMB	DOF 2001*	Original CHIS 2001		Revised CHIS 2001	
	Percentage	Percentage	Standard	Percentage	Standard

<b>Race-ethnicity</b>			<b>Error Percent</b>		<b>Error Percent</b>
Latino	33.3	32.6	0.013	33.6	0.000
White non-Latino alone	46.2	48.2	0.085	46.0	0.000
African American non-Latino alone	6.6	5.4	0.057	6.3	0.000
American Indian/Alaska Native non-Latino alone	0.6	0.2	0.011	0.6	0.000
Asian non-Latino alone	11.1	9.8	0.058	11.2	0.000
Native Hawaiian or Other Pacific Islander non-Latino alone	0.3	0.2	0.016	0.3	0.000
Two or more non-Latino	1.9	3.7	0.065	2.0	0.000
<b>Total</b>	<b>100</b>	<b>100.0</b>	<b>0.000</b>	<b>100.0</b>	<b>0.000</b>

\* Total includes population in group quarters

The last group of estimates that might be expected to have large differences is that for totals for the oversampled cities (Long Beach, Pasadena and Berkeley) in CHIS 2001. Table 9 shows the estimates of total population of the oversampled cities compared to the Census 2000. The raking dimensions defined by sampling stratum in the original weights included separate cells for the cities. In the revised weights, the 4<sup>th</sup> raking dimension included a cell for Berkeley and the SPAs in Los Angeles County. As shown in Table 9, the original estimates for totals for Long Beach and Pasadena are closer to the Census totals than the estimates from the revised weights. On the other hand, Table 10 shows the total estimates for the SPAs using the revised weights are closer to the totals from the Census 2000.

Table 9. Estimates of totals for oversampled cities

City	<b>Census 2000*</b>	<b>Original CHIS 2001</b>		<b>Revised CHIS 2001</b>	
	<b>Total</b>	<b>Total</b>	<b>Standard Error Total</b>	<b>Total</b>	<b>Standard Error Total</b>
Long Beach	461,522	451,341	6,496	562,357	19,323
Pasadena	133,936	130,418	6,562	162,293	7,595
Berkeley	102,743	96,921	17	99,153	0

\* Total includes population in group quarters

Table 10. Estimates of totals for Los Angeles SPAs

<b>Census 2000*</b>	<b>Original CHIS 2001</b>	<b>Revised CHIS 2001</b>
---------------------	---------------------------	--------------------------

SPA	Total	Total	Standard Error Total	Total	Standard Error Total
Antelope Valley <sup>a</sup>	298,539	384,126	22,950	299,818	9
San Fernando	2,027,276	2,120,612	46,512	2,035,957	8
San Gabriel	1,703,636	1,700,065	41,866	1,710,930	17
Metro	1,103,293	972,682	30,902	1,108,017	7
West	685,217	647,986	23,428	688,151	15
South <sup>b</sup>	805,009	854,584	36,453	808,457	20
East	1,279,849	1,222,301	39,265	1,285,329	4
South Bay	1,605,998	1,440,539	31,119	1,612,875	7

\* Total includes population in group quarters

In summary, the comparisons show that for most estimates of proportions the differences between the estimates using the two sets of weights are very small. For totals, the estimates differ more, largely due to the growth in the population that is taken into account in the revised weights but is not in the original weights. Estimates of totals by race and ethnicity from the original and revised weights do have differences and those differences are very large for estimates of totals for AIAN. Although differences for proportions for the AIAN are likely to be much smaller, any estimate that involves AIAN should be carefully considered. The estimates for totals for some of the cities that were oversampled in 2001 also have some substantial differences. The differences in the estimates for totals for the AIAN and the oversampled cities are due largely to the definitions used in raking dimensions. These findings make one conclusion clear; any estimates of differences between the years should use the revised weights from CHIS 2001 rather than the original weights.



## APPENDIX A

Table A-1. Original definitions of the dimensions used in raking in CHIS 2001

Dimension	Level	Description	Categories
1	Stratum	Large age groups (3) × sex (2)	11 Under 12 years, males
			12 Under 12 years, females
			21 12 to 17 years, males
			22 12 to 17 years, females
			31 18 years or older, males
			32 18 years or older, females
2	Stratum	Small age groups (9)	1 Under 5 years
			2 6 to 11 years
			3 12 to 17 years
			4 18 to 24 years
			5 25 to 29 years
			6 30 to 39 years
			7 40 to 49 years
			8 50 to 64 years
			9 65 years or older
366	State	American Indian/Alaska Native indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older
4	State	Asian indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older
5	Collapsed stratum	Latino indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older
6	Collapsed stratum	African American indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older
7	Collapsed stratum	White indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older

Dimension	Level	Description	Categories
8	State	Other indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older
9	State	Pacific Islander or Native Hawaiian indicator (2) × large age groups (2)	11 Yes, under 18 years old
			12 Yes, 18 years or older
			21 No, under 18 years old
			22 No, 18 years or older
10	State	Small age groups (13) × sex (2)	11 0 to 3 years, male
			12 0 to 3 years, female
			21 4 to 7 years, male
			22 4 to 7 years, female
			31 8 to 11 years, male
			32 8 to 11 years, female
			41 12 to 14 years, male
			42 12 to 14 years, female
			51 15 to 17 years, male
			52 15 to 17 years, female
			61 18 to 25 years, male
			62 18 to 25 years, female
			71 26 to 30 years, male
			72 26 to 30 years, female
			81 31 to 37 years, male
			82 31 to 37 years, female
			91 38 to 45 years, male
			92 38 to 45 years, female
			101 46 to 53 years, male
			102 46 to 53 years, female
111 54 to 64 years, male			
112 54 to 64 years, female			
121 65 to 77 years, male			
122 65 to 77 years, female			
131 78 years or older, male			
132 78 years or older, female			
11	State	Nontelephone adjustment cells	See Table 1

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

Table A-2. Definitions of the dimensions used in raking in CHIS 2003

Dimension	Level	Description		Categories
1	Stratum (collapsed where necessary)	Age groups (3) x Sex (2)	11	Under 12 years, male
			12	Under 12 years, female
			21	12 to 17 years, male
			22	12 to 17 years, female
			31	18 years or older, male
			32	18 years or older, female
2	Stratum (collapsed where necessary)	Age groups (9)	1	Under 6 years
			2	6 to 11 years
			3	12 to 17 years
			4	18 to 24 years
			5	25 to 29 years
			6	30 to 39 years
			7	40 to 49 years
			8	50 to 64 years
			9	65 years or older
3	State	Age groups (13) x Sex (2)	11	Under 4 years, male
			12	Under 4 years, female
			21	4 to 7 years, male
			22	4 to 7 years, female
			31	8 to 11 years, male
			32	8 to 11 years, female
			41	12 to 14 years, male
			42	12 to 14 years, female
			51	15 to 17 years, male
			52	15 to 17 years, female
			61	18 to 24 years, male
			62	18 to 24 years, female
			71	25 to 30 years, male
			72	25 to 30 years, female
			81	31 to 37 years, male
			82	31 to 37 years, female
			91	38 to 45 years, male
			92	38 to 45 years, female
			101	46 to 53 years, male
			102	46 to 53 years, female
111	54 to 64 years, male			
112	54 to 64 years, female			
121	65 to 77 years, male			
122	65 to 77 years, female			
131	78 years or older, male			
132	78 years or older, female			
4	SPAs in Los Angeles Co., Alameda County,	SPAs (8), Alameda Co. (3), Remainder of CA (1)	1	SPA 1
			2	SPA 2
			...	...
			7	SPA 7
			8	SPA 8
9	Hayward Census Place,			

	Remainder of CA		10 11 12	Oakland Census Place, Remainder of Alameda Co. Remainder of CA
5	Region (collapsed where necessary)	Race (6)	1 2 3 4 5 6 7	Latino, White non-Latino, African American non-Latino, American Indian non-Latino, Asian non-Latino, Native Hawaiian non-Latino Two or more races non-Latino
6	State	Gender (2) x Race (6) x Age groups (4) (collapsed where necessary)	111 112 113 114 121 122 123 124 131 132 133 134 141 142 143 144 151 152 153 154 161 162 163 164 171 172 173 174 111 212 213 214 221 222 223 224 231 232 233 234 241	Male, Latino, under 12 years, Male, Latino, 12 to 17 years, Male, Latino 18 to 64 years, Male, Latino 65 years or older, Male, White non-Latino, under 12 years, Male, White non-Latino, 12 to 17 years, Male, White non-Latino, 18 to 64 years, Male, White non-Latino, 65 years or older, Male, African American non-Latino, under 12 years, Male, African American non-Latino, 12 to 17 years, Male, African American non-Latino, 18 to 64 years, Male, African American non-Latino, 65 years or older, Male, American Indian non-Latino, under 12 years, Male, American Indian non-Latino, 12 to 17 years, Male, American Indian non-Latino, 18 to 64 years, Male, American Indian non-Latino, 65 years or older, Male, Asian non-Latino, under 12 years, Male, Asian non-Latino, 12 to 17 years, Male, Asian non-Latino, 18 to 64 years, Male, Asian non-Latino, 65 years or older, Male, Native Hawaiian non-Latino, under 12 years, Male, Native Hawaiian non-Latino, 12 to 17 years, Male, Native Hawaiian non-Latino, 18 to 64 years, Male, Native Hawaiian non-Latino, 65 years or older, Male, Two or more races non-Latino, under 12 years, Male, Two or more races non-Latino, 12 to 17 years, Male, Two or more races non-Latino, 18 to 64 years, Male, Two or more races non-Latino, 65 years or older Male, Latino, under 12 years, Female, Latino, 12 to 17 years, Female, Latino 18 to 64 years, Female, Latino 65 years or older, Female, White non-Latino, under 12 years, Female, White non-Latino, 12 to 17 years, Female, White non-Latino, 18 to 64 years, Female, White non-Latino, 65 years or older, Female, African American non-Latino, under 12 years, Female, African American non-Latino, 12 to 17 years, Female, African American non-Latino, 18 to 64 years, Female, African American non-Latino, 65 years or older, Female, American Indian non-Latino, under 12 years,

			242	Female, American Indian non-Latino, 12 to 17 years,
			243	Female, American Indian non-Latino, 18 to 64 years,
			244	Female, American Indian non-Latino, 65 years or older,
			251	Female, Asian non-Latino, under 12 years,
			252	Female, Asian non-Latino, 12 to 17 years,
			253	Female, Asian non-Latino, 18 to 64 years,
			254	Female, Asian non-Latino, 65 years or older,
			261	Female, Native Hawaiian non-Latino, under 12 years,
			262	Female, Native Hawaiian non-Latino, 12 to 17 years,
			263	Female, Native Hawaiian non-Latino, 18 to 64 years,
			264	Female, Native Hawaiian non-Latino, 65 years or older,
			271	Female, Two or more races non-Latino, under 12 years,
			272	Female, Two or more races non-Latino, 12 to 17 years,
			273	Female, Two or more races non-Latino, 18 to 64 years,
			274	Female, Two or more races non-Latino, 65 years or older
7	State	Asian groups (5) x Age groups (4) (collapsed where necessary)	11	Non-Latino Chinese only, under 12 years,
			12	Non-Latino Chinese only, 12 to 17 years,
			13	Non-Latino Chinese only, 18 to 64 years,
			14	Non-Latino Chinese only, 65 years or older,
			21	Non-Latino Korean only, under 12 years,
			22	Non-Latino Korean only, 12 to 17 years,
			23	Non-Latino Korean only, 18 to 64 years,
			24	Non-Latino Korean only, 65 years or older,
			31	Non-Latino Filipino only, under 12 years,
			32	Non-Latino Filipino only, 12 to 17 years,
			33	Non-Latino Filipino only, 18 to 64 years,
			34	Non-Latino Filipino only, 65 years or older,
			41	Non-Latino Vietnamese only, under 12 years,
			42	Non-Latino Vietnamese only, 12 to 17 years,
			43	Non-Latino Vietnamese only, 18 to 64 years,
			44	Non-Latino Vietnamese only, 65 years or older,
			51	Other or non-Asian only, under 12 years,
			52	Other or non-Asian only, 12 to 17 years,
			53	Other or non-Asian only, 18 to 64 years,
			54	Other or non-Asian only, 65 years or older,
8	Stratum (collapsed where necessary)	Race (2) x Age groups (4)	11	Latino, under 12 years,
			12	Latino, 12 to 17 years,
			13	Latino 18 to 64 years,
			14	Latino 65 years or older,
			21	White non-Latino, under 12 years,
			22	White non-Latino, 12 to 17 years,
			23	White non-Latino, 18 to 64 years,
			24	White non-Latino, 65 years or older,
			31	Non-White non-Latino, under 12 years,
			32	Non-White non-Latino, 12 to 17 years,
			33	Non-White non-Latino, 18 to 64 years,
			34	Non-White non-Latino, 65 years or older,
9	State	Education (4)	1	Not applicable (age < 25 years)
			2	Less than High School,

			3 4	High School grad or GED recipient, At least some college
10	State	# Adults in HH (3)	1 2 3	0 or 1 adult, 2 adults, 3 or more adults
11	State	Non- telephone	1 2 3 4 5 6	Homeowner, Renter, 1 adult in household, less than 25 years old Renter, 1 adult in household, 25 years old or older, less than HS Renter, 1 adult in household, 25 years old or older, High School grad or GED recipient Renter, 1 adult in household, 25 years old or older, At least some college Renter, 2 or more adults in household

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey.

Table A-3. Revised raking dimensions used in CHIS 2001

Dimension	Level	Description	Categories
1	Stratum (collapsed where necessary)	Age groups (3) x Sex (2)	11 Under 12 years, male
			12 Under 12 years, female
			21 12 to 17 years, male
			22 12 to 17 years, female
			31 18 years or older, male
			32 18 years or older, female
2	Stratum (collapsed where necessary)	Age groups (9)	1 Under 6 years
			2 6 to 11 years
			3 12 to 17 years
			4 18 to 24 years
			5 25 to 29 years
			6 30 to 39 years
			7 40 to 49 years
			8 50 to 64 years
			9 65 years or older
3	State	Age groups (13) x Sex (2)	11 Under 4 years, male
			12 Under 4 years, female
			21 4 to 7 years, male
			22 4 to 7 years, female
			31 8 to 11 years, male
			32 8 to 11 years, female
			41 12 to 14 years, male
			42 12 to 14 years, female
			51 15 to 17 years, male
			52 15 to 17 years, female
			61 18 to 24 years, male
			62 18 to 24 years, female
			71 25 to 30 years, male
			72 25 to 30 years, female
			81 31 to 37 years, male
			82 31 to 37 years, female
			91 38 to 45 years, male
			92 38 to 45 years, female
101 46 to 53 years, male			
102 46 to 53 years, female			
111 54 to 64 years, male			
112 54 to 64 years, female			
121 65 to 77 years, male			
122 65 to 77 years, female			
131 78 years or older, male			
132 78 years or older, female			
4	SPAs in Los Angeles Co.,	SPAs (8), Alameda Co. (3),	1 SPA 1 2 SPA 2 ... ..

Dimension	Level	Description	Categories
	Alameda County, Remainder of CA	Remainder of CA (1)	7 SPA 7 8 SPA 8 9 Berkeley Census Place, 10 Remainder of Alameda Co. 11 Remainder of CA
5	Region (collapsed where necessary)	Race (6)	1 Latino, 2 White non-Latino, 3 African American non-Latino, 4 American Indian non-Latino, 5 Asian non-Latino, 6 Native Hawaiian non-Latino 7 Two or more races non-Latino
6	State	Gender (2) x Race (6) x Age groups (4) (collapsed where necessary)	111 Male, Latino, under 12 years, 112 Male, Latino, 12 to 17 years, 113 Male, Latino 18 to 64 years, 114 Male, Latino 65 years or older, 121 Male, White non-Latino, under 12 years, 122 Male, White non-Latino, 12 to 17 years, 123 Male, White non-Latino, 18 to 64 years, 124 Male, White non-Latino, 65 years or older, 131 Male, African American non-Latino, under 12 years, 132 Male, African American non-Latino, 12 to 17 years, 133 Male, African American non-Latino, 18 to 64 years, 134 Male, African American non-Latino, 65 years or older, 141 Male, American Indian non-Latino, under 12 years, 142 Male, American Indian non-Latino, 12 to 17 years, 143 Male, American Indian non-Latino, 18 to 64 years, 144 Male, American Indian non-Latino, 65 years or older, 151 Male, Asian non-Latino, under 12 years, 152 Male, Asian non-Latino, 12 to 17 years, 153 Male, Asian non-Latino, 18 to 64 years, 154 Male, Asian non-Latino, 65 years or older, 161 Male, Native Hawaiian non-Latino, under 12 years, 162 Male, Native Hawaiian non-Latino, 12 to 17 years, 163 Male, Native Hawaiian non-Latino, 18 to 64 years, 164 Male, Native Hawaiian non-Latino, 65 years or older, 171 Male, Two or more races non-Latino, under 12 years, 172 Male, Two or more races non-Latino, 12 to 17 years, 173 Male, Two or more races non-Latino, 18 to 64 years, 174 Male, Two or more races non-Latino, 65 years or older 111 Male, Latino, under 12 years, 212 Female, Latino, 12 to 17 years, 213 Female, Latino 18 to 64 years, 214 Female, Latino 65 years or older, 221 Female, White non-Latino, under 12 years, 222 Female, White non-Latino, 12 to 17 years, 223 Female, White non-Latino, 18 to 64 years, 224 Female, White non-Latino, 65 years or older, 231 Female, African American non-Latino, under 12 years, 232 Female, African American non-Latino, 12 to 17 years,



Dimension	Level	Description	Categories
			233 Female, African American non-Latino, 18 to 64 years, 234 Female, African American non-Latino, 65 years or older, 241 Female, American Indian non-Latino, under 12 years, 242 Female, American Indian non-Latino, 12 to 17 years, 243 Female, American Indian non-Latino, 18 to 64 years, 244 Female, American Indian non-Latino, 65 years or older, 251 Female, Asian non-Latino, under 12 years, 252 Female, Asian non-Latino, 12 to 17 years, 253 Female, Asian non-Latino, 18 to 64 years, 254 Female, Asian non-Latino, 65 years or older, 261 Female, Native Hawaiian non-Latino, under 12 years, 262 Female, Native Hawaiian non-Latino, 12 to 17 years, 263 Female, Native Hawaiian non-Latino, 18 to 64 years, 264 Female, Native Hawaiian non-Latino, 65 years or older, 271 Female, Two or more races non-Latino, under 12 years, 272 Female, Two or more races non-Latino, 12 to 17 years, 273 Female, Two or more races non-Latino, 18 to 64 years, 274 Female, Two or more races non-Latino, 65 years or older
7	State	Asian groups (5) x Age groups (4) (collapsed where necessary)	11 Non-Latino Chinese only, under 12 years, 12 Non-Latino Chinese only, 12 to 17 years, 13 Non-Latino Chinese only, 18 to 64 years, 14 Non-Latino Chinese only, 65 years or older, 21 Non-Latino Korean only, under 12 years, 22 Non-Latino Korean only, 12 to 17 years, 23 Non-Latino Korean only, 18 to 64 years, 24 Non-Latino Korean only, 65 years or older, 31 Non-Latino Filipino only, under 12 years, 32 Non-Latino Filipino only, 12 to 17 years, 33 Non-Latino Filipino only, 18 to 64 years, 34 Non-Latino Filipino only, 65 years or older, 41 Non-Latino Vietnamese only, under 12 years, 42 Non-Latino Vietnamese only, 12 to 17 years, 43 Non-Latino Vietnamese only, 18 to 64 years, 44 Non-Latino Vietnamese only, 65 years or older, 51 Other or non-Asian only, under 12 years, 52 Other or non-Asian only, 12 to 17 years, 53 Other or non-Asian only, 18 to 64 years, 54 Other or non-Asian only, 65 years or older,
8	Stratum (collapsed where necessary)	Race (2) x Age groups (4)	11 Latino, under 12 years, 12 Latino, 12 to 17 years, 13 Latino 18 to 64 years, 14 Latino 65 years or older, 21 White non-Latino, under 12 years, 22 White non-Latino, 12 to 17 years, 23 White non-Latino, 18 to 64 years, 24 White non-Latino, 65 years or older, 31 Non-White non-Latino, under 12 years, 32 Non-White non-Latino, 12 to 17 years, 33 Non-White non-Latino, 18 to 64 years, 34 Non-White non-Latino, 65 years or older,

<b>Dimension</b>	<b>Level</b>	<b>Description</b>	<b>Categories</b>
Dimension	Level	Description	Categories
9	State	Education (4)	1 Not applicable (age < 25 years) 2 Less than High School, 3 High School grad or GED recipient, 4 At least some college
10	State	# Adults in HH (3)	1 0 or 1 adult, 2 2 adults, 3 3 or more adults
11	State	Nontelephone dimension	See Table 2

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey.

**Appendix B**

Table B-1. Estimates of totals and proportions computed using the original and revised CHIS 2001 weights.

Variable	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage
AB1										
General health rating										
1 Excellent	10,897	4,639,092	53,706	19.5	0.23	10,956	4,657,909	50,707	18.9	0.21
2 Very Good	18,229	7,674,828	60,433	32.2	0.25	18,386	7,656,340	67,651	31.1	0.28
3 Good	16,135	7,263,424	62,685	30.5	0.26	16,455	7,508,807	68,227	30.5	0.28
4 Fair	7,677	3,406,252	50,439	14.3	0.21	7,880	3,791,812	52,904	15.4	0.21
5 Poor	2,447	846,624	24,464	3.6	0.10	2,548	970,736	31,065	3.9	0.13
Total	55,385	23,830,220	3,512	100.0	0.00	56,225	24,585,603	4,175	100.0	0.00
AB29										
Doctor diagnosed high blood pressure										
1 Yes	14,308	5,195,612	49,066	21.8	0.21	14,489	5,449,532	51,171	22.2	0.21
2 No	41,016	18,613,565	48,866	78.2	0.21	41,671	19,112,385	50,725	77.8	0.21
Total	55,324	23,809,178	5,282	100.0	0.00	56,160	24,561,917	6,056	100.0	0.00
AH1										
Have usual source of healthcare										
1 Yes	36,753	15,564,300	63,078	82.3	0.24	36,964	15,911,194	66,915	81.7	0.25
2 No	6,163	3,355,183	46,964	17.7	0.24	6,348	3,566,136	51,642	18.3	0.25
Total	42,916	18,919,483	56,849	100.0	0.00	43,312	19,477,330	61,489	100.0	0.00
AF1										
Ever diagnosed with any cancer										

Variable	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage
1 Yes	5,540	1,782,943	28,996	7.5	0.12	5,553	1,789,963	29,736	7.3	0.12
2 No	49,805	22,038,420	28,486	92.5	0.12	50,632	22,785,005	29,797	92.7	0.12
Total	55,345	23,821,363	3,349	100.0	0.00	56,185	24,574,969	4,179	100.0	0.00
AH43										
Marital status										
1 Married	28,367	13,173,915	55,730	55.4	0.23	28,917	13,515,223	55,010	55.1	0.22
2 Living with partner	3,372	1,770,301	31,771	7.4	0.13	3,385	1,833,527	32,240	7.5	0.13
3 Widowed	5,242	1,285,320	21,757	5.4	0.09	5,301	1,394,365	24,677	5.7	0.10
4 Divorced	7,117	1,879,848	28,162	7.9	0.12	7,154	1,871,407	29,208	7.6	0.12
5 Separated	1,756	602,952	21,351	2.5	0.09	1,775	665,447	24,448	2.7	0.10
6 Never married	9,411	5,066,841	45,325	21.3	0.19	9,572	5,259,370	42,551	21.4	0.17
Total	55,265	23,779,177	7,180	100.0	0.00	56,104	24,539,339	6,525	100.0	0.00
AH39										
Is respondent citizen of united states										
1 Yes	6,055	3,240,912	48,654	44.2	0.56	6,563	3,519,010	52,763	43.3	0.58
2 No	6,283	4,001,776	52,381	54.5	0.58	6,567	4,496,194	56,523	55.4	0.58
3 Application pending	159	97,501	9,388	1.3	0.13	181	102,686	8,347	1.3	0.10
Total	12,497	7,340,189	57,843	100.0	0.00	13,311	8,117,889	56,759	100.0	0.00
A11										
Covered by Medicare										
1 Yes	12,512	4,062,656	27,555	17.2	0.12	12,690	4,239,680	31,838	17.4	0.13
2 No	42,599	19,587,847	29,382	82.8	0.12	43,261	20,157,377	34,071	82.6	0.13
Total	55,111	23,650,502	14,312	100.0	0.00	55,951	24,397,058	17,102	100.0	0.00
AL8A										
Phys/mntl impairment kept from working.										
1 Yes	3,830	1,261,507	26,957	10.4	0.22	3,927	1,394,756	31,828	10.7	0.24
2 No	23,971	10,842,893	57,103	89.6	0.22	24,453	11,597,796	63,962	89.3	0.24
Total	27,801	12,104,400	54,905	100.0	0.00	28,380	12,992,552	59,620	100.0	0.00

Variable	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage
CA6										
Rating of child's health										
1 Excellent	6,373	3,025,204	39,011	48.4	0.62	6,433	2,988,544	35,387	47.8	0.57
2 Very Good	3,208	1,556,820	33,780	24.9	0.54	3,264	1,559,954	33,378	24.9	0.53
3 Good	2,274	1,228,638	33,970	19.7	0.54	2,355	1,274,296	34,245	20.4	0.55
4 Fair	655	390,449	20,707	6.2	0.33	664	379,445	20,825	6.1	0.33
5 Poor	79	48,418	8,291	0.8	0.13	83	51,277	8,775	0.8	0.14
Total	12,589	6,249,529	1,585	100.0	0.00	12,799	6,253,516	1,302	100.0	0.00
CA12										
Dr ever told you child has asthma										
1 Yes	1,508	706,431	19,707	12.3	0.34	1,525	731,632	22,584	12.7	0.39
2 No	10,104	5,049,442	26,493	87.7	0.34	10,276	5,023,823	28,818	87.3	0.39
Total	11,612	5,755,873	20,357	100.0	0.00	11,801	5,755,455	21,403	100.0	0.00
CB1										
Total										
Child injured seriously in past 12 mos.										
1 Yes	1,193	516,768	20,589	8.3	0.33	1,205	496,129	19,776	7.9	0.32
2 No	11,389	5,729,912	20,704	91.7	0.33	11,587	5,754,038	20,034	92.1	0.32
Total	12,582	6,246,680	2,108	100.0	0.00	12,792	6,250,166	1,944	100.0	0.00
CG1										
Have regular childcare arrangement										
1 Yes	4,487	2,109,270	34,943	33.8	0.56	4,516	2,108,416	34,077	33.8	0.55
2 No	8,092	4,132,298	34,685	66.2	0.56	8,273	4,137,255	34,124	66.2	0.55
Total	12,579	6,241,567	3,306	100.0	0.00	12,789	6,245,671	3,000	100.0	0.00
TB1										
General health										
1 Excellent	892	465,443	22,601	15.8	0.77	906	487,290	22,745	16.0	0.75
2 Very Good	2,258	1,057,746	26,285	35.9	0.89	2,267	1,112,528	27,912	36.5	0.91

Variable	Sample Size	Weighted Size	Standard Error Weighted	Total Percentage	Standard Error Total Percentage	Sample Size	Weighted Size	Standard Error Weighted	Total Percent age	Stand ard Error Total Percent age
3 Good	2,049	1,053,909	25,768	35.8	0.87	2,076	1,087,888	27,569	35.7	0.91
4 Fair	552	340,295	18,411	11.5	0.62	556	328,770	17,852	10.8	0.59
5 Poor	43	29,739	6,452	1.0	0.22	46	29,824	6,747	1.0	0.22
Total	5,794	2,947,131	1,894	100.0	0.00	5,851	3,046,301	2,122	100.0	0.00
TA4										
Currently attend school										
1 Yes	5,688	2,891,008	7,409	98.0	0.25	5,746	2,991,218	7,167	98.1	0.23
2 No	110	59,505	7,369	2.0	0.25	110	58,942	7,099	1.9	0.23
Total	5,798	2,950,513	1,004	100.0	0.00	5,856	3,050,159	861	100.0	0.00
TE22										
Ever had more than few sips of alcohol										
1 Yes	2,037	946,842	25,423	32.1	0.86	2,048	948,301	26,166	31.1	0.86
2 No	3,758	2,002,337	25,262	67.9	0.86	3,805	2,100,206	25,988	68.9	0.86
Total	5,795	2,949,179	1,324	100.0	0.00	5,853	3,048,507	1,456	100.0	0.00