MANAGING LONGITUDINAL RESEARCH STUDIES:

Issues and Decisions to be made on Collecting, Coding and Reporting Race and Ethnicity for Public Health Indicators

By

Linda L Remy, MSW PhD
Ted Clay, MS
Geraldine Oliva M.D. MPH

UCSF Family Health Outcomes Project
Geraldine Oliva, MD MPH, Director
Jennifer Rienks, PhD, Associate Director

500 Parnassus Ave. Room MU-337
San Francisco, California 94143-0900
Phone: 415-476-5283
Fax: 415-476-6051
Web: https://fhop.ucsf.edu/

September 2011
Updated January 2017

TABLE OF CONTENTS

Introduction **1**

Race/Ethnic Categories **2**

- Available Groups 2
- Race Groups with Self-contained Denominators 4

Population Denominators **5**

- California Department of Finance 5
- NCHS Population Estimates 6
- Available Alternatives 8

Grouping Race **9**

- Overview of Race Processing Decisions 9
- Assigning Race Text to Race Codes 11
- Hawaiian “Trump” and Duplicate Codes 11
- Derive Race Groups from Race Codes 12
- Simplify Other Race 14
- Federal Single Race Bridging Guidelines for Trends 14

Allocating Bridged Race Categories **15**

- Bridging Methods 15
- Bridging the Numerator 19
- Bridging the Denominator 20

Summary **21**

Endnotes **22**

TABLE OF FIGURES

Figure 1.	Annual average projection error and actual growth rates	8
Figure 2.	Summary of bridging steps with options	10

TABLE OF TABLES

Table 1.	Available numerator and denominator options for race/ethnic reporting	3
Table 2.	Proposed longitudinal layout	4
Table 3.	DOF population estimates compared with NCHS/Census, 2009 and 2010	7
Table 4.	Race groups for California Vital Statistics files, and for bridging	13
Table 5.	Allocation methods	16
Table 6.	Percent distribution of single-race assignment after application of the NHIS-regression method to bridge multiple-race counts to single-race categories: public-use Census Modified Race Summary file, United States, 2000	17
Table 7.	Bridged multi-race versus first race choice (%): California births, 2000-2009	17
Table 8.	Bridging results by method - California births, 2000-2009	18

Suggested Citation

Remy LL, Clay T, Oliva G. (2011). Issues and Decisions to be made on Collecting, Coding and Reporting Race and Ethnicity for Public Health Indicators. Family Health Outcomes Project, University of California, San Francisco. Updated January 2017. Available at: http://fhop.ucsf.edu/data-management-methods.

ACRONYMS

AIAN	American Indian/Alaska Native
API	Asian/Pacific Islander
ASC	Ambulatory Surgery Center
CDC	Centers for Disease Control
CDPH	California Department of Public Health
DOF	Department of Finance
DRU	Demographic Research Unit/DOF
ED	Emergency Department
FHOP	Family Health Outcomes Project
MCAH	Maternal Child and Adolescent Health
NCHS	National Center for Health Statistics
NCVS	National Center for Vital Statistics
NHIS	National Health Interview Survey
NHOPI	Native Hawaiian/Other Pacific Islander
OHIR	Office of Health Informatics Research
OMB	Office of Management and Budget
OSHPD	Office of Statewide Health Planning and Development
PDD	Patient Discharge Data
UCSF	University of California, San Francisco

Issues and Decisions to be made on Collecting, Coding and Reporting Race and Ethnicity for Public Health Indicators

INTRODUCTION

The “Race/Ethnicity Guidelines”, approved in 2003 by the California Directors of Public Health (CDPH) and Health and Human Services (CHHS) for use by all programs, explicitly did not address how to handle multi-race coding for trend analysis [1]. We did not have enough experience using these variables at the time. Further, the National Center for Health Statistics (NCHS) had not yet provided guidance on what to do when the same groups are not available over time or there is a mismatch between groups in the numerator and denominator.

This document discusses issues related to developing a standardized approach to coding and reporting race and ethnicity for data sets maintained by CDPH. The focus is using these to explore race/ethnic differences in indicators of health status and outcomes over time.

The first step is to propose an approach to defining race/ethnic categories, particularly where more than one race is available on California Vital Statistics files. The second is to propose a method for bridging race categories over time, to allow trend analysis. There are two principal concerns in accomplishing these tasks: accuracy and comparability. This document addresses bridging issues related to both concerns. The purpose is to make groups comparable, across indicators and time, between numerators and denominators. Bridging continues until consistent numerators and denominators are available over the period of interest.

Officially, race bridging is defined as “making data collected using one set of race categories consistent with data collected using a different set of race categories, to permit estimation and comparison of race-specific statistics at a point in time or over time. More specifically, race bridging is a method used to make multiple-race and single-race data collection systems sufficiently comparable to permit estimation and analysis of race-specific statistics [2].” The goal of bridging is to approximate the size of single-race groups rather than to approximate how each individual would have responded to the traditional single-race question [3]. These definitions treat race separately from Hispanic ethnicity, and this document has the same focus.

Background. The Family Health Outcomes Project (FHOP) at the University of California, San Francisco (UCSF), makes data products and does research for units in the CDPH and other funders. Most of this work involves longitudinal analysis of Vital Statistics data (birth, death, fetal death) and medical data from the Office of Statewide Health Planning and Development (OSHPD) (inpatient discharges (PDD), emergency department (ED) visits, ambulatory surgery center (ASC) visits), often supplemented with longitudinal data from other sources.

Most relevant to this discussion is that we developed and maintain a product called DataBooks, distributed to all California health jurisdictions. These Excel files provide 12 years of data on selected health indicators. They contain rate tables and trend graphs by race/ethnic categories for longitudinal monitoring. DataBook source files lack time-consistent numerators and denominators, and this circumstance will continue for some years. Different groups have been available at different times in data sets we use for numerators (Vital Statistics and OSHPD). A similar problem exists for denominators (California Department of Finance (DOF) or National Center for Health Statistics (NCHS)). In making DataBooks, we partially implement Federal race bridging guidelines for trends [4,5].

Questions. Below are a set of questions to be answered in making determinations of how to provide guidance to longitudinally monitor of race and ethnic differences over time. We also propose answers to these questions, discussed in more detail later in this document.

1. Should we change categories used to monitor health status given the availability of multi-race fields? Given longitudinal limitations in available numerators and denominators it would be unwise at this time. However, we propose an alternative way to display available data for smaller race/ethnic groups.

2. Should we change the data source for denominators for population-based indicators? CDPH branches and programs traditionally have used DOF population estimates but they sometimes are not able to provide these data in a timely fashion and in the detail that we need. NCHS now provides an alternative data source that we can use. Using Federal estimates has many advantages.

3. How should we calculate multi-race categories or collapse multi-race into single race categories? It seems prudent to follow Federal methods for bridging race for trends [6].

4. Should we change the bridging method CDPH programs currently use? We recommend the method NCHS uses for vital statistics reporting [7]. If that is not desirable, we recommend using Race 1 in Vital Statistics files bridged according to NCHS rules given the fact that OSHPD data has only one race variable.

RACE/ETHNIC CATEGORIES

Available Groups

Table 1. shows race groups available for numerator and denominator datasets. Vital Statistics categories are pre-and post-2000. OSHPD pre-1995 covers the period 1985 forward. They had yet another definition prior to 1985.

The RACE column shows whether the group is defined (Y) or must be calculated (C). In Vital Statistics data, API, Asian, Native Hawaiian/Other Pacific Islander (NHOPI), and AIAN groups are aggregates of multiple categories. Blank cells indicate the option is not available. While Vital Statistics introduced different codes at different times, all codes aggregate to these groups.

Table 1. Available numerator and denominator options for race/ethnic reporting

	Numerator								Denominator							
	Vital Statistics				OSHPD				DOF				NCHS			
	Pre-2000		Post-2000		Pre-1995		Post-1995		Pre-2000		Post-2000		1990-1999		Post-2000	
Group	RACE	HISP	RACE	HISP	RACE	HISP	RACE	HISP	RACE	HISP	RACE	HISP	RACE	HISP	RACE	HISP
AIAN	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y		Y	Y	Y	Y
API	C	Y	C	Y	Y		Y	Y	Y	Y	C		Y	Y	Y	Y
Asian	C	Y	C	Y							Y					
Black	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y		Y	Y	Y	Y
NHOPI	C	Y	C	Y							Y					
White	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y		Y	Y	Y	Y
Multi		Y	C *	Y							Y *					
Other	Y *	Y	Y *	Y	Y *		Y *	Y								
Unknown	Y *	Y	Y *	Y	Y *		Y *	Y								
Hispanic	C		C		Y		C		C		Y		C		C	

Note that multi-race is not included in the Vital Statistics files before 2000 and is not in OSHPD files. The asterisk (*) indicates the group must be bridged to provide consistent longitudinal race classifications. Asterisked groups are multi-race, other, and unknown (which includes missing and declined to state). In this document, we refer to them collectively as the Bridge Group.

The HISP column reflects whether the department that generates the data for each data set separately reports ethnicity (Hispanic and Non-Hispanic), which Federal rules prefer. Table 1. highlights that DOF and OSHPD groups limit longitudinal options.

For longitudinal analyses, we suggest adding rows for API sub-groups and for Bridge Groups that display the numerators and denominators (if available), but not calculating rates or trends for those groups. Table 1. makes clear that, while we could have this layout for many indicators, we would not always have numerator and denominator data. Table 2. is an example of a template for reporting race and ethnic categories over time.

Table 2. Proposed longitudinal layout

Race/Ethnicity	Numerator											
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
0 Total	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
1 White	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
2 Black	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
3 Hispanic	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
4 Asian/PI	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
Asian	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
Native Hawaiian/Other PI	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
5 Amer Indian/Alaska Native	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
Bridge groups	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
Multi-race	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
Other	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999
Unknown	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999	9,999

This layout allows the display of numerators and denominators for detailed race groups. It would allow local health jurisdictions to compute annual rates but would not allow longitudinal monitoring of comparable groups for enough years to calculate a trend. Bridge Groups are apportioned to primary races, and their rates would not be independent.

Race Groups with Self-contained Denominators

Some indicators use numerators and denominators from the same data set such as the total number of births, with variations from indicator to indicator on cases to include. For such indicators it is possible to display data for more subgroups. However, this creates a situation with different groups from indicator to indicator, and makes it impossible to compare groups across indicators. Increasing the number of subgroups also creates many situations with very small numerators with their associated statistical difficulties.

API is the main group affected by these circumstances. As discussed above, we propose to continue calculating rates of total API so there will be stable numbers to calculate rates and trends. We suggest adding rows in indicator tables for NHOPI and Asian sub-categories including, if desired, Southeast Asians. This would give jurisdictions enough information to calculate rates for those subgroups if needed, with the caveat that these would not be independent from the API group.

Thus, while we could change the basic decision of common bridged race groups across indicators, we do not believe it advisable. We recommend that we continue trending only the basic groups AIAN, API, Black, Hispanic, and White. This is consistent with Federal guidelines.

In the next section, we discuss the limitation in available sources for denominators used to calculate population-based indicators. Then we turn to assigning multi-race. The last section discusses alternative bridging methods.

POPULATION DENOMINATORS

The Demographic Research Unit (DRU) in the DOF is the official source of demographic data for California planning and budgeting [8]. Recent Census data and NCHS population estimates indicate that DOF estimates are off target. DOF is aware of the problem but lacks staff to re-estimate 2001-2009 population in the near future. This requires deciding whether to develop methods to address problems with DOF data or to use available NCHS files for denominators.

California Department of Finance

DOF projections are about 5% higher on average than what the recent Census reported. The Census released 1-year and 5-year age, sex, and race numbers beginning May 2011. California's population data were released in July 2011. Due to staff constraints, DOF has no plans in the near future to re-estimate 2001-2009 numbers to synchronize with the 2000-2010 Census. This directly affects rates in all products the State branches and programs or LHJs produce that rely on DOF denominator data.

Since 2004, DOF groups Hispanic all race, then breaks out non-Hispanic as White, Black, Asian, NHOPI, AIAN, and multi-race (7-group race/ethnicity). DOF no longer classifies race by Hispanic and non-Hispanic ethnicity [9]. From the mid-1990's until 2004, DOF data allowed analysts to make ethnicity by race tables consistent with Federal guidelines although we lacked sufficient years for trends. In 2004, DOF introduced 7-group race/ethnicity and its multi-race allocation table, and discontinued population by race and ethnicity [10]. Analysts apply the table after data summary.

Mary Heim, former Director of the DOF Demographic Research Unit (DRU) and the prime contact for census data, recently retired from her position. Linda Remy spoke with John Malson, who assumed most of her duties [11]. Linda also spoke at length with Melanie Martindale [12], who is responsible for population projections but not estimates. John and Melanie both said they are short-staffed and have to complete mandatory responsibilities before they update intercensal population estimates and make new projections.

By late summer, they hoped to release new county-level numbers for 7-group race/ethnicity, but not by age or sex. They reported no near term plan to update estimates for 2001-2010. The next projections will be 2010-2020 and it will be quite some time before these are available. Even previously, when DOF had no staffing issues, they did not release multi-race population estimates by age and sex, and the allocation table, until June 2004 for the 2000 census.

NCHS Population Estimates

NCHS publishes 1-year age group population numbers for Hispanic and non-Hispanic bridged race groups (AIAN, API, Black, White) [13] from 1990 forward. It uses the most sophisticated algorithms to produce these estimates. Federal agencies use these files as denominators to calculate national bridged statistics, including those that monitor vital statistics and progress toward Healthy People objectives. For comparability, NCHS recommends states use the same publicly available files.

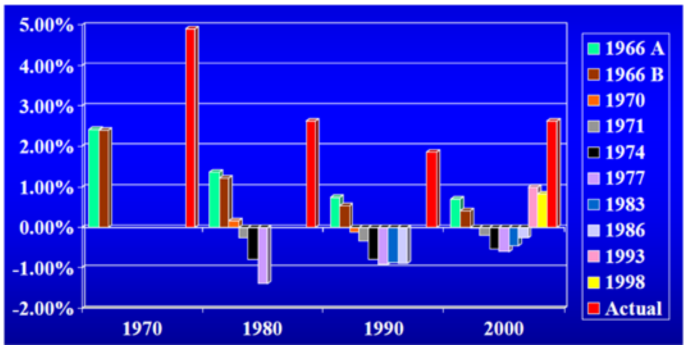
Table 3. compares county-level estimates from DOF with county-level estimates from NCHS and the Census for 2009 and 2010. The 2010 numbers are from Appendix A (Melanie Martindale's email to Linda Remy and the Excel file she forwarded).

Comparing 2009 DOF estimates we have been using to NCHS population estimates, DOF was 4.5% higher statewide (county average, 5.6% higher). The 2009 range is from negative percent (Placer, San Francisco) to 8% or more (Alpine, Colusa, Inyo, Imperial, Lassen, Merced, Modoc, Mono, San Benito, Sierra). The US Census reported California had a 2010 population of 37,253,956. For 2010, DOF estimated 39,023,830, a difference of 1,769,874 (4.5% higher statewide, county average, 5.1% higher). At the county level, DOF differences range from negative percent (Humboldt, Placer, San Luis Obispo) to 8% or more (Alpine, Colusa, Modoc, Plumas, San Benito, Sierra, Trinity, Yuba).

Table 3. DOF population estimates compared with NCHS/Census, 2009 and 2010

	2009 Population				2010 Population			
Area	DOF	NCHS	Difference	Diff %	DOF	Census	Difference	Diff %
0 California	38,688,293	36,961,664	1,726,629	4.5	39,023,830	37,253,956	1,769,874	4.5
1 Alameda	1,540,499	1,491,482	49,017	3.2	1,547,725	1,510,271	37,454	2.4
2 Alpine	1,358	1,041	317	23.3	1,366	1,175	191	14.0
3 Amador	39,867	37,876	1,991	5.0	40,220	38,091	2,129	5.3
4 Butte	226,819	220,577	6,242	2.8	229,292	220,000	9,292	4.1
5 Calaveras	47,197	46,731	466	1.0	47,612	45,578	2,034	4.3
6 Colusa	23,305	21,321	1,984	8.5	23,667	21,419	2,248	9.5
7 Contra Costa	1,064,755	1,041,274	23,481	2.2	1,073,137	1,049,025	24,112	2.2
8 Del Norte	30,636	29,114	1,522	5.0	30,896	28,610	2,286	7.4
9 El Dorado	186,336	178,447	7,889	4.2	188,565	181,058	7,507	4.0
10 Fresno	964,755	915,267	49,488	5.1	978,797	930,450	48,347	4.9
11 Glenn	30,411	28,299	2,112	6.9	30,763	28,122	2,641	8.6
12 Humboldt	134,024	129,623	4,401	3.3	134,595	134,623	(28)	(0.0)
13 Imperial	184,704	166,874	17,830	9.7	188,432	174,528	13,904	7.4
14 Inyo	19,088	17,293	1,795	9.4	19,159	18,546	613	3.2
15 Kern	853,225	807,407	45,818	5.4	867,102	839,631	27,471	3.2
16 Kings	161,030	148,764	12,266	7.6	163,659	152,982	10,677	6.5
17 Lake	66,727	65,279	1,448	2.2	67,329	64,665	2,664	4.0
18 Lassen	37,570	34,473	3,097	8.2	37,831	34,895	2,936	7.8
19 Los Angeles	10,449,155	9,848,011	601,144	5.8	10,498,286	9,818,605	679,681	6.5
20 Madera	158,253	148,632	9,621	6.1	161,149	150,865	10,284	6.4
21 Marin	253,517	250,750	2,767	1.1	253,641	252,409	1,232	0.5
22 Mariposa	18,936	17,792	1,144	6.0	19,065	18,251	814	4.3
23 Mendocino	92,466	86,040	6,426	6.9	92,991	87,841	5,150	5.5
24 Merced	267,699	245,321	22,378	8.4	272,376	255,793	16,583	6.1
25 Modoc	10,684	9,107	1,577	14.8	10,778	9,686	1,092	10.1
26 Mono	14,589	12,927	1,662	11.4	14,772	14,202	570	3.9
27 Monterey	430,418	410,370	20,048	4.7	432,567	415,057	17,510	4.0
28 Napa	140,834	134,650	6,184	4.4	142,284	136,484	5,800	4.1
29 Nevada	101,822	97,751	4,071	4.0	102,442	98,764	3,678	3.6
30 Orange	3,190,126	3,026,786	163,340	5.1	3,218,409	3,010,232	208,177	6.5
31 Placer	340,705	348,552	(7,847)	(2.3)	345,834	348,432	(2,599)	(0.8)
32 Plumas	21,744	20,122	1,622	7.5	21,804	20,007	1,797	8.2
33 Riverside	2,178,729	2,125,440	53,289	2.4	2,223,972	2,189,641	34,331	1.5
34 Sacramento	1,437,311	1,400,949	36,362	2.5	1,448,227	1,418,788	29,439	2.0
35 San Benito	62,436	55,058	7,378	11.8	63,782	55,269	8,513	13.3
36 San Bernardino	2,136,425	2,017,673	118,752	5.6	2,167,303	2,035,210	132,093	6.1
37 San Diego	3,169,126	3,053,793	115,333	3.6	3,192,061	3,095,313	96,748	3.0
38 San Francisco	814,225	815,358	(1,133)	(0.1)	817,179	805,235	11,944	1.5
39 San Joaquin	723,964	674,860	49,104	6.8	737,054	685,306	51,748	7.0
40 San Luis Obispo	267,958	266,971	987	0.4	269,290	269,637	(347)	(0.1)
41 San Mateo	734,230	718,989	15,241	2.1	736,058	718,451	17,607	2.4
42 Santa Barbara	430,756	407,057	23,699	5.5	433,562	423,895	9,667	2.2
43 Santa Clara	1,823,759	1,784,642	39,117	2.1	1,833,961	1,781,642	52,319	2.9
44 Santa Cruz	266,776	256,218	10,558	4.0	267,706	262,382	5,324	2.0
45 Shasta	189,109	181,099	8,010	4.2	191,069	177,223	13,846	7.2
46 Sierra	3,644	3,174	470	12.9	3,632	3,240	392	10.8
47 Siskiyou	46,853	44,634	2,219	4.7	47,045	44,900	2,145	4.6
48 Solano	436,254	407,234	29,020	6.7	439,859	413,344	26,515	6.0
49 Sonoma	491,415	472,102	19,313	3.9	494,413	483,878	10,535	2.1
50 Stanislaus	549,408	510,385	39,023	7.1	557,133	514,453	42,680	7.7
51 Sutter	100,044	92,614	7,430	7.4	101,756	94,737	7,019	6.9
52 Tehama	64,632	61,138	3,494	5.4	65,353	63,463	1,890	2.9
53 Trinity	15,005	14,165	840	5.6	15,130	13,786	1,344	8.9
54 Tulare	456,605	429,668	26,937	5.9	464,321	442,179	22,142	4.8
55 Tuolumne	58,435	55,175	3,260	5.6	58,650	55,365	3,285	5.6
56 Ventura	846,802	802,983	43,819	5.2	853,608	823,318	30,290	3.5
57 Yolo	202,673	199,407	3,266	1.6	205,243	200,849	4,394	2.1
58 Yuba	78,465	72,925	5,540	7.1	79,925	72,155	7,770	9.7
Average Difference				5.6				5.1

Figure 1. Annual average projection error and actual growth rates



In 2007, DOF issued a report on population trends [14]. Figure 1 from that report focuses on population projection error through the 2000 Census. The present projection error is greater than any period since the mid-1960's. Traditionally, the DOF underestimated growth. For the last ten years, they overestimated growth.

Overall, NCHS estimates appear more accurate than DOF in that they are closer to the Census. NCHS probably will remain more accurate because of the sophisticated methods the agency uses. This is one reason some demographers believe we no longer need the Census.

Available Alternatives

DOF will not release corrected sex by age by 7-race group population estimates for some time. New detailed numbers for 2000 forward will not be available for another 3 years at the earliest. If history is an example, we may not see revised estimates by age and sex before 2014. Our current bridge period begins before 2000.

It may be time to revisit the use of DOF population data. Continuing to use DOF estimates post-2000 is problematic. Given differences between NCHS and DOF, rates we have been producing have been imprecise. If we continue to use DOF, we will have to recalculate indicators when they release new estimates. Whether we change now or later, some areas/groups/indicators will have different rates than we now report.

In conclusion, available DOF estimates are not comparable with Census or NCHS. Given sizeable discrepancies with the Census we recommend that we not use existing DOF estimates, and, given DOF staffing issues, that we not wait for new DOF estimates.

Should we use NCHS estimates? NCHS is producing bridged population data closer to Census numbers. The Federal government has made a strong commitment to produce this. If we use NCHS estimates, we would better approximate Federal vital statistics reports for California.

Either way we will have to explain the change. We could explain the change in the context of the Census, while it is a current topic. We think it would be better to make the explanation now rather than later.

GROUPING RACE

Overview of Race Processing Decisions

This section is mainly about the three race variables in the Vital Statistics files. These files differ from OSHPD files, which have only one race field (See Table 1.).

When filling out Vital Statistics forms, people are allowed to self-identify up to three race groups. When making master files, Vital Statistics employs an algorithm to convert these into one variable with a new value indicating multi-race if it applies. The multi-race algorithm complies with Federal guidelines for annual vital statistics reporting [4], but it does not follow Federal guidelines for bridging race longitudinally [5,7].

Bridging race is a multi-step process in Vital Statistics files. It begins with assigning text strings to race codes, then race codes to groups, followed by alternative ways of combining race groups.

Figure 2 shows a series of actions and decisions based on Federal race bridging guidelines for trends [6,7]. It also identifies various bridging methods the Federal government investigated before coming to a decision as to which method to use. Understanding this process and the alternatives examined will give the best opportunity to understand why they picked the method they did. Following their method will make it possible for California to approximate rates that NCHS and other federal agencies report.

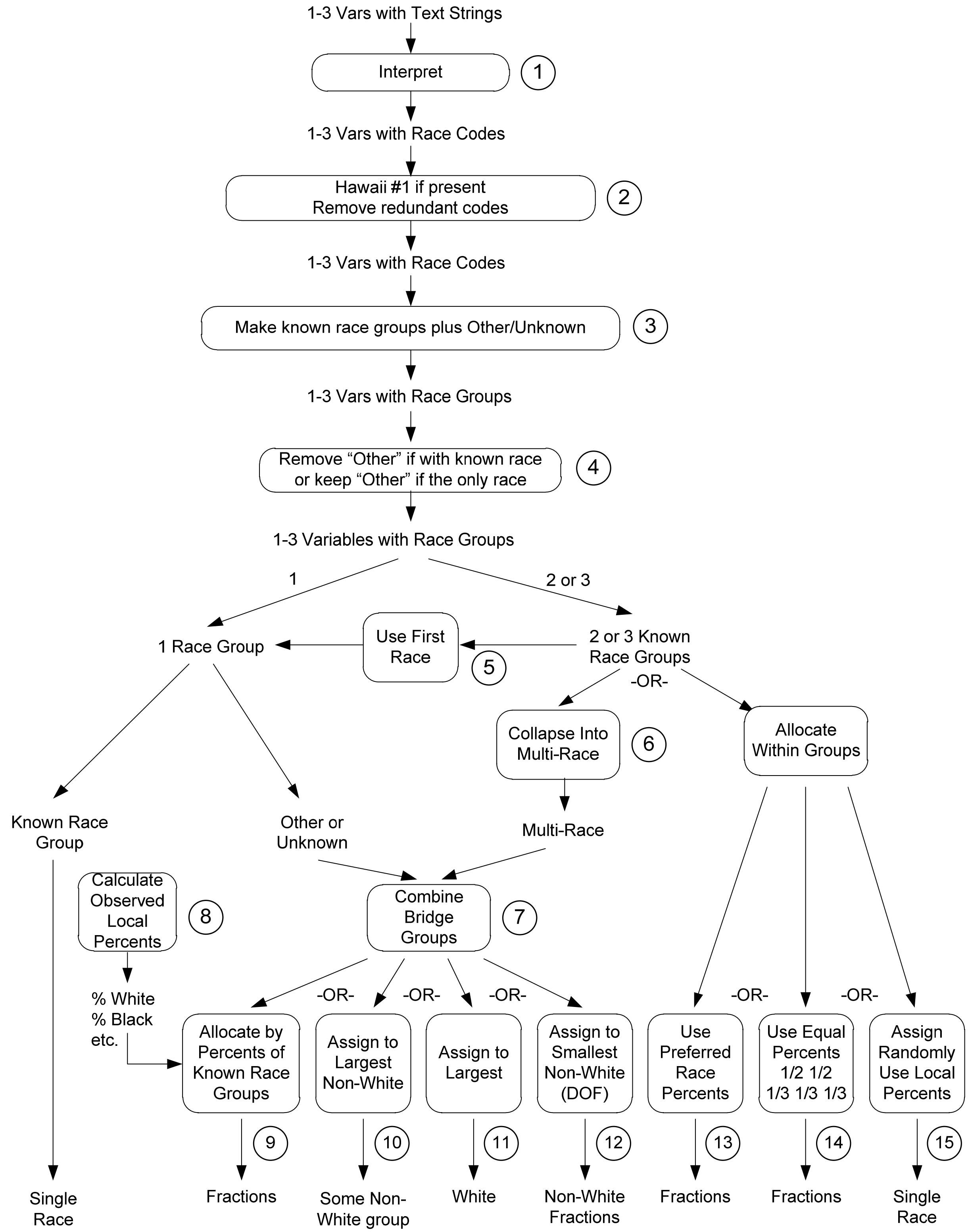
Steps ① through ④ in Figure 2 apply to both Vital Statistics and OSHPD files. Everything below ④ involves decisions on allocation methods for bridging.

Before turning to the more detailed description pertaining primarily to Vital Statistics, we want to discuss OSHPD files. The path of their race variable is to the left, after ④, because multi-race is not available. Then we must choose a method to bridge Other/Unknown. DOF advised [15] that they use the NCHS race allocation file to bridge Other/Unknown. We discuss this file later.

Linda Remy had a series of phone conversations with Debra Ingram, Brady Hamilton, Dave Johnson, and others at the NCHS to be sure she understood the process. They provided information and various resources, which we describe as the issues arise in the following pages.

Immediately after Figure 2, we begin to discuss the individual steps and decisions for bridging race longitudinally. Each step of the explanation is keyed to numbers circled in Figure 2.

Figure 2. Summary of bridging steps with options



Assigning Race Text to Race Codes

① The Confidential Vital Statistics files we use have three coded race variables and three race text variables (2004 forward). When people complete a birth or death certificate, they write the race(s) the person reports. These text strings, with many spelling errors, are the basis for the coded variables. For example, “Navajo” and “Navaho” should be assigned to code 30, “American Indian”, and “East Indian” should be 52 “Indian (Asian)”.

Under Federal rules, people of Eskimo and Aleut ancestry are to be described as Alaska Native, Central and South American Indians should be classified as American Indians, and the definition of the "American Indian or Alaska Native" category should be modified to include original peoples from North, Central, and South America [16].

Through our work with CDPH-OHIR, we identified that identical race strings were being assigned to different codes, in the same file and across files and time [17]. AIAN is under-counted because lookup tables did not include many variations of tribal spellings and were assigned to “Other”. Variations of “Mongolian” were assigned to eight different race group codes. Staff making different files used different rules to assign race, and lookup tables were not coordinated across files. The largest impact was in birth certificate files.

OHIR confirmed our findings and convened an agency-wide committee that decided to follow Federal coding rules. The agency has been improving computer algorithms. They are incorporating an agency-wide lookup table NCHS prepares from previous years of California vital statistics data. NCHS updates the file annually. Given a text string in any Vital Statistics dataset, the lookup table assigns the same race code. The 2010 files probably will be the first prepared under the new procedures. The main effect will be a small increase in the number of AIAN, reducing Other. There also will be a small reduction in people classified multi-race. Implementing these rules will make a small bump in race groups and trends.

With text strings from 2004 forward, we think we should incorporate re-assigning race groups based on the master CDC lookup table into our race macro. We would make the macro and lookup table publicly available. The lookup table would be extremely helpful to anyone facing a conversion of race text strings to codes for any purpose. Jurisdictions get their Vital Statistics data before the agency releases master files. Making these resources available will help local coding immensely. It also will move us sooner toward our goal of improving data quality.

Hawaiian “Trump” and Duplicate Codes

② By Federal regulation, if Hawaiian (but not other Pacific Islanders) appears in second or third position in the array of race variables, it moves to first position and the others shift accordingly. Note that this has an effect only when “Use First Race” is selected. ⑤

The assignment of codes to strings often results in redundant codes. For example, Serbian, French, and Irish resolve to one race: White. Set redundant codes to blank.

Derive Race Groups from Race Codes

③ Larger numbers of race groups result in larger numbers of multi-race persons. Over the years, we have received five algorithms (that we can find) to calculate single race. Three are from Vital Statistics [18]:

- UniqRace collapses multiple race codes into unique race combinations. For example, the White, Black, White combination is changed to White and Black, with position three empty. If Hawaiian is second or third position, it moves to first. It may be thought of as a data cleaning step, perhaps a precursor to the other programs. See ② above.

- OMB_Race incorporates the UniqRace algorithm, recodes each of three existing race codes into three temporary variables with the OMB minimum categories plus Other and “unknown/withheld,” and finally assigns multi-race. *Vital Statistics uses this algorithm to calculate multi-race variables for mothers, fathers, and decedents.* This meets Federal revised guidelines for race/ethnicity [2], but does not follow bridged race guidelines [4, 5].

- DHS_Race follows the same process as OMB_Race to make a single race variable that complies with DHS/FHOP guidelines [1]. *Note that DHS/FHOP guidelines specifically deferred multi-race coding decisions until more data were available.*

Table 4. is from the paper on race coding that Vital Statistics provides with file documentation. We added a column “Bridging for Trends” on the right side showing the NCHS Bridge Groups. In reviewing this table, keep in mind that a good number of American Indians of North, Central, and South American origins have been classified “Other.”

Table 4. Race groups for California Vital Statistics files, and for bridging

Race Origin from Birth Certificate	Race Codes 1, 2 & 3	FHOP/DHS Minimum Category*	DHS Race Code[1]	OMB Minimum Category	Multi- Race Code	Bridging for Trends
White	10	White	1	White	1	White
Black	20	Black, African American, or Negro only	2	Black or African American only	2	Black
American Indian	30	American Indian or Alaska Native only	3	American Indian or Alaska Native only	3	AIAN
Eskimo	57					
Aleut	58					
Asian – Unspecified	40	Asian only	4	Asian only	4	API
Asian – Specified	41					
Asian – Chinese	42					
Asian – Japanese	43					
Asian – Korean	44					
Indian (Asian)	52					
Asian – Vietnamese	45	Southeast Asian Only	5			
Asian – Cambodian	46					
Asian – Thai	47					
Asian – Laotian	48					
Asian – Hmong	49					
Filipino	53	Filipino only	6			
Hawaiian	54	Native Hawaiian or Pacific Islander only	7	Native Hawaiian or Other Pacific Islander only	5	
Guamanian	55					
Samoan	56					
Other Pacific Islander	59					
Other – Specified	51	Other race only	9	Other/Unknown** (or may be combined with White)	6	Bridge
Withheld	98				9	
Unknown	99					
Two or more unique races		More than one race	8	Multiple Race	7	

* All races in this column are non-Hispanic. All Spanish/Hispanic/Latino Origin persons of one or more races are to be combined on the category "Spanish/Hispanic/Latino Origin."

** OMB does not specify "Other/Unknown" as a category, since census records are imputed to some other category. For maximum compatibility with the census, California records with Other or Unknown race are combined with White.

1 This field is not included on the birth files, but can be generated by collapsing races in the manner of the SAS program DHS race.sas.

Larger numbers of race groups result in larger numbers of multi-race. Relative to Federal bridging guidelines, OMB_Race over-estimates multi-race by grouping various combinations of NHOPI, Asian, and Other as multi-race. DHS_Race breaks out cultural groups within API, further over-estimating multi-race. With Federal bridging, NHOPI and Asian group as API before looking for multi-race. The combination Black/Other groups as Black, under the rule that specific trumps non-specific. Then multi-race is calculated.

In addition to these algorithms, the following are from MCAH:

- RACEETH RECODE.SAS (Carrie Florez, 16-Apr-2002). This program uses Race1 and resolves to White (which includes other, unknown, missing), Black, AIAN, API, Spanish. It does not include code for Hawaiian to trump. Otherwise, this is the same as ⑤. This program essentially reflects the DOF denominator files of the time.

- STEPB.SAS (Mike Curtis 22-Nov-2002). This was for breastfeeding. The program resolves to Black, API, Hispanic, AIAN, White, Other (which includes multi-race), and Missing. It essentially resolves to ⑦.

Simplify Other Race

④ When Other race is with a known race, Other is removed. Other remains when it is the only race recorded.

Federal Single Race Bridging Guidelines for Trends

The following summarizes Federal guidelines where more than one race variable is available, race is collected after ethnicity, and results will be used for trends [6].

Move Hawaiian to first position. Remove multiples of the same code.

Next, assign each race variable to one of four mutually exclusive categories: AIAN, API, Black, or White. Someone who is Chinese, Vietnamese, and Hawaiian will be API in all fields. Because codes are identical across the array, the single-race variable will be API, not multi-race. Another rule is that specific trumps nonspecific. Assign API/Other to API. The result is a list of one or more race categories. If multiple race categories remain at the end of this process, the individual is multi-race.

Sometimes data are available for a dyad such as mother/father as in birth certificates. If race or ethnicity is known for one but not the other of the pair, use known race and/or ethnicity. In other words, if mother's race or ethnicity is unknown, look to father and assign from his if known [19].

When following Federal guidelines to classify California births, multi-race occurs for less than 2% of parents (mother and father) in the birth statistical master file [17]. Of these, 61% are White plus some other group.

In the end, some number of the Bridge Group (multi-race, other/unknown) must be apportioned. Over the period 2000-2009, using race assigned by Vital Statistics and applying NCHS bridging rules to California birth certificates, 3.9% of mother's records are in the Bridge Group. Of these, 1.9% are multi-race and 2.0% Other.

ALLOCATING BRIDGED RACE CATEGORIES

This section describes bridging methods, where the Bridge Group is assigned to main groups. Multi-race information is handled in two ways, either: (1) collapsed into a single multi-race category ⑥ and sent to the combined Bridge Group ⑦; or (2) the explicit list of multiple race categories (e.g. "White/Black") is allocated within combinations. Combined Bridge Groups are allocated by one of methods ⑧ through ⑫. Explicit multi-race information permits allocation sensitive to specific groups as in ⑬ through ⑮. Notice that ⑮ uses information derived from ⑧.

Steps "Combine" ⑦ and "Calculate Observed Percents" ⑧ require obtaining totals for each known race within the group of interest. The group of interest may be females aged 15-19, etc. Within that group, some have known race and others are in the Bridge Group. Per ⑧, if known race at the geographic level of interest was distributed as, say, 60% White, 20% Black, and 30% API, and the Bridge method is ⑨, the Bridge Group is allocated to those groups in those proportions. Circles ⑨ through ⑮ describe various bridging methods discussed below.

Bridging Methods

In 2004, NCHS released a report evaluating various bridging methods using the 2000 Census modified race summary data and data from the National Health Interview Study (NHIS) [6]. Previous OMB studies identified the following as the most likely allocation methods [4]. Table 5. describes the methods evaluated.

Under equal fractions ⑭, equal proportions of each multiple-race group are allocated to the component races. AIAN/API is assigned 50% to AIAN and 50% to API. AIAN/API/White is assigned 33% to each group.

Table 5. Allocation methods

Method	Type	Group	Description
Deterministic	Whole Allocation	Largest group other than White	Assign to the category, other than White, with the largest single-race count.
		Largest group	Assign to the category with the largest single-race count.
		Smallest group	Assign to category, other than White, with the smallest single-race count.
		Plurality	Assign based on the National Health Interview Survey (NHIS) plurality.
		Proportional fractions	Assign randomly using local non-ordered proportions. NCHS/CDC adopted this for Vital Statistics reporting and trending [7].
	Fractional allocation	Observed fractions	Assign by fractions to each category, where fractions equal percents of known race groups.
		Equal fractions	Assign in equal fractions to each of the possible categories.
		NHIS Proportional fractions	Assign by fractions to each category identified, where fractions equal NHIS proportions based on preferred race (Table 6 below . This method is not available for Vital Statistics or OSHPD data.
Probabilistic	Regression	NHIS logistic regression	Randomly assign based on NHIS fraction. The CDC does not recommend this method for local uses, and preferred race is not available for Vital Statistics or OSHPD data.

With NHIS fractions ⑬, assignment is based on preference fractions in the population, and all groups receive a varying proportional assignment. Table 6 shows these fractions, based on national estimates from regression.

Table 6. is from a related NCHS study to understand single race preferences among the multi-race [3]. The table uses NHIS data, which captures three race groups plus a question on preferred race. Of 11 multi-race groups, seven included White, and proportional single-race preference was predominantly to White in six. Most Black combinations predominantly preferred Black and AIAN/API preferred AIAN. This method is not available for DataBooks.

Table 6. Percent distribution of single-race assignment after application of the NHIS-regression method to bridge multiple-race counts to single-race categories: public-use Census Modified Race Summary file, United States, 2000

	Single-race assignment			
Multiple-race response	AIAN	API	Black	White
AIAN+API	63.3	36.7		
AIAN+Black	15.9		84.1	
AIAN+White	22.4			77.6
API+Black		41.4	58.6	
API+White		40.9		59.1
Black+White			62.9	37.1
AIAN+API+Black	26.8	25.4	47.8	
AIAN+API+White	2.2	8.7		89.1
AIAN+Black+White	18.7		57.4	23.9
API+Black+White		12.0	11.9	76.1

The plurality method ⑪ uses information similar to this to assign single race. AIAN/API is assigned to AIAN because of the 63.3% preference. Most groups with White are assigned to White, and most groups with Black are assigned to Black. API does not receive any assignments.

For this report, we expanded our earlier CHS study [20] to cover the period 2000-2009. Table 7. summarizes results for California resident women whose children were born in California. As before, we classified multi-race and compared results to first race.

Table 7. Bridged multi-race versus first race choice (%): California births, 2000-2009

	Total		Race 1 Percent			
Multi-race group	Count	Percent	AIAN	API	Black	White
Total	105,056	100.0	5.7	18.7	21.5	54.2
12 AIAN+API	936	0.9	43.5	56.5		
13 AIAN+Black	2,785	2.7	14.1		85.9	
14 AIAN+White	26,238	25.0	18.4			81.6
23 API+Black	4,231	4.0		26.0	74.0	
24 API+White	41,860	39.8		41.8		58.2
34 Black+White	23,883	22.7			62.0	38.0
123 AIAN+API+Black	252	0.2	8.3	18.7	73.0	
124 AIAN+API+White	1,053	1.0	12.1	25.5		62.5
134 AIAN+Black+White	2,806	2.7	7.0		56.5	36.6
234 API+Black+White	1,012	1.0		17.3	44.1	38.6

We found 105,056 multi-race mothers (1.9%) with 54.2% White as first race. As in Table 6. Black is the plurality in groups with this race and White is the plurality in White multi-race groups not including Black. Expressed relative to first race rather than preferred, Table 7. results are not remarkably different from Table 6.

In general population studies, multi-race occurs more often among younger than older groups. We evaluated this for mothers in the birth certificate files. Of non-Hispanic women age less than

25, 4.5% were multi-race compared with 1.6% of non-Hispanic women age 35 or older. Of Hispanic women age less than 25, 1.3% were multi-race compared with 1% of Hispanic women age 35 or older.

Table 8. summarizes results of applying various methods to bridge race data for California births over the period 2000-2009. These methods begin with NCHS race recodes (Figure 2, ① to ④). The first row shows NCHS recoded Race 1 ⑤. The next randomly allocates the NCHS Race 1 group Other records, using the lag function within single-race records ⑮ [21]. From ⑬, Table 6 is the basis for the NCHS fractions and Table 7 for the California fractions. The CDC/NCHS method ⑮ randomly assigns to one race using the State File Number, using known county-level distributions of race, sex, age, and Hispanic ethnicity [22]. With each method in Table 2, the number of records assigned to a given group varies from method to method, but total numbers remain constant.

Table 8. Bridging results by method - California births, 2000-2009

		NCHS Race				Bridge Groups	
Bridge Method	Total	AIAN	API	Black	White	Other	Multi
NCHS Race 1	5,418,909	31,923	670,257	326,298	4,284,177	106,254	
Row %	1.00	0.01	0.12	0.06	0.79	0.020	
(15) Other reassigned	106,254	480	15,949	6,088	83,737		
NCHS Race 1 Bridged	5,418,909	32,403	686,206	332,386	4,367,914		
Row %	1.00	0.01	0.13	0.06	0.81		
NCHS Race	5,418,909	25,941	650,632	303,746	4,227,280	106,254	105,056
Row %	1.00	0.00	0.12	0.06	0.78	0.020	0.019
(14) Multi-race assigned	105,056	16,350	24,286	16,806	47,614		
Equal Fractions	5,418,908	42,771	690,867	326,640	4,358,631		
Row %	1.00	0.01	0.13	0.06	0.80		
Race 1 Bridge Difference	0	10,368	4,661	(5,746)	(9,283)		
(13) Multi-race assigned	105,056	7,528	19,493	21,695	56,340		
NCHS Fractions	5,418,908	33,949	686,074	331,529	4,367,356		
Row %	1.00	0.01	0.13	0.06	0.81		
Race 1 Bridge Difference	0	1,546	(132)	(857)	(558)		
(13) Multi-race assigned	105,056	6,000	19,619	22,538	56,898		
California Fractions	5,418,908	32,421	686,200	332,372	4,367,914		
Row %	1.00	0.01	0.13	0.06	0.81		
Race1 Bridge Difference	0	18	(6)	(14)	0		
(15) Multi-race assigned	105,056	7,749	18,262	23,946	55,099		
CDC/NCHS	5,418,909	34,170	684,843	333,780	4,366,116		
Row %	1.00	0.01	0.13	0.06	0.81		
Race 1 Bridge Difference	0	1,767	(1,363)	1,394	(1,798)		

Starting with NCHS Race 1, 106,254 Other records are randomly reassigned to a known single race. This increases the AIAN number by 480, for a new AIAN total of 32,403.

Next, we examine the result of apportioning multi-race records using methods NCHS tested. With equal fraction assignment, AIAN increases 16,350, returning a new total of 42,771. This is calculated by summing ⑮ Other reassigned, NCHS Race, and ⑭ Multi-race assigned.

Allocations are evaluated relative to the simplest method, using NCHS bridged Race 1. Relative to Race 1, the equal fraction method returns 10,368 new AIAN cases. It also increases API, but decreases Black and White. Using either NCHS or California fractions, few cases are gained or lost in any group. The CDC/NCHS method results in small increases for AIAN and Black, and small decreases for API and White.

Evaluating results relative to Race 1 implicitly assumes that Race 1 is preferred, an assumption the NCHS/CDC method ⑮ does not make. The OMB emphasized the equal fractions method. This equally respects all heritages and significantly increases AIAN and API. The important observation is that after bridging Other, the net change in number of people reassigned to a given race is very small relative to Race 1 for all but the equal fractions method.

These results are based on 10 years of California birth certificates. Adopting a method other than using Race 1 requires considerable work for very little change overall, even with equal fractions. At the county level, results would be imperceptible annually.

Bridging the Numerator

We need to determine which of the described bridging alternatives to recommend. The new method has to be simple to describe and understand. It would be helpful to make a SAS macro that could be available for all state agencies and the public so jurisdictions could replicate state rates by race/ethnic categories.

We have seen the problems with whole assignment, whether to smallest or largest non-white, or to largest. Smallest and largest non-White under-estimate White, and largest over-estimates White. The OMB reported that smallest and largest non-White results in the highest misclassification and is the least reliable, while largest group whole assignment (typically White) can be used in a wide variety of applications and has the fewest reliability problems [4, p. 97]. **Note:** FHOP used largest whole group assignment for DataBooks, as did DOF before 2004.

The main issue with smallest and largest non-White group assignment is that White receives no cases, when most multi-race people with White ancestry prefer that group (Tables 6 and 7). Plurality ⑬ (Tables 6 or 7) may approximate preferred single race, but API does not receive any assignments.

Observed ⑨ and proportional ⑭ fraction methods are easy to apply and understand. They assure that all groups receive cases.

Proportional fraction methods need a reference (Table 6 or Table 7). Table 6 shows results of the regression method NCHS uses to estimate bridged population. We used California data to apply a variation on the NCHS method (Table 7).

In NCHS studies, preferred race tended to be the first race reported. This suggests another simple alternative: use first race, after the Hawaiian trump, the base comparison in Table 8. In this way, all groups are represented with no change in general perception of population characteristics. Using first race also most closely approximates the single-race variable in OSHPD files. In any alternative, Other must be bridged because no denominators include that group.

Linda Remy spoke with Debra Ingram of the CDC [23]. She is one of the authors on most of the OMB/CDC/NCHS bridge studies. Debra said whole assignment methods ⑩ through ⑫ were the least desirable. Debra felt that Table 6. was not as sensitive as needed for our purposes, because preferences vary by age, the same problem as the DOF allocation table. It also is a national distribution and may not reflect local preference distributions [4, p. 96, 23]. The OMB considers the regression method too difficult to implement and describe [4]. Debra agrees with this, and we do, too.

Equal fraction assignment ⑭ seems to be the OMB's preferred method. Each category receives an equal fraction, where the sum of assignments equals one [4, p. 89]. This method tabulates people without accounting for preference and gives equal weight to component groups. It increases counts for traditionally under-counted AIAN and API populations. From a policy perspective, equal fraction assignment is an attractive option.

In the end, the NCHS/CDC decided to assign single race randomly, using local distributions for Hispanic, sex, and age ⑮ [7]. Linda Remy spoke with David P. Johnson of CDC/NCHS, who sent the file NCHS uses to bridge race for all published NCHS reports. DOF uses this file to bridge Other/Unknown for their population estimates. Decision rules are clearly described and easily implemented. We used this file to calculate the CDC/NCHS results.

If the decision is not to use ⑮, then our next recommended alternative is to use NCHS defined Race 1 and bridge Other ⑨. This most closely approximates the race variable in OSHPD files. The method is easy to implement and describe, and, according to the CDC, has good reliability.

Bridging the Denominator

The major decision is whether we continue to use DOF estimates or move to NCHS bridge estimates [24]. Using NCHS estimates would bring California rates into better agreement with Federal reports. This is the simple solution, and it is consistent longitudinally from 1990 forward.

Pre-2004, using whole allocation to largest group, DOF assigned the Bridge Group to White [11]. Although it did not include the Hawaiian trump, RACEETH_RECODE.SAS from MCAH used these same rules. From 2004 forward, DOF allocated the Bridge Group based on the smallest non-white group [12]. AIAN/White is assigned to AIAN. AIAN/Black goes to the group with the smaller proportion of those in the county. We earlier discussed the statistical problems with both methods. While previous DOF rules assigned too many to White, its current rules assign too few. NCHS determined that this is the least valid bridging method.

DOF recommended using their allocation table of county-specific proportions to apportion the Bridge Group to 2000 population data forward [10]. FHOP has been using this to bridge denominator population for DataBooks. The table is not age sensitive. It ignores that younger populations (such as MCAH) are more likely to be multi-race than older populations [6].

Because DOF population estimates do not jibe with Federal estimates, rates and trends will change no matter what denominator is used. Total population differences between DOF and NCHS range from negative percent to 8% or more, and average more than 5% in a given year. Since the mid-1960s, disagreement between DOF estimates and the Census has never been so extreme.

These differences have major implications for calculating rates. Regardless of whether the Federal population estimates are used now, or denominators are changed in several years when DOF releases new bridging estimates, the increase will be apparent at some point. Given the size of the discrepancies in population estimates it is probably better to move now to NCHS estimates, when the changes can be discussed in the context of the newly released Census. It is the simplest solution, the easiest to explain, and will produce rates closer to Federal estimates.

SUMMARY

The document reviewed methods to calculate race/ethnicity, affecting CDPH data products. We advise that CDPH calculate rates and trends for Total plus the five standard OMB groups (AIAN, API, Black, Hispanic, White), but modify longitudinal displays of numerators and denominators for API and Bridge sub-groups. For indicators needing external denominators, it seems advisable to use the NCHS population denominator files. Like the OMB, it is advisable to allocate Bridge Groups using whole assignment based on a random function and local sex, age, and Hispanic ethnicity [15]. It is easy to explain and understand. If the decision is not to use whole assignment, then the next recommended alternative is to use NCHS defined Race 1 and randomly bridge Other [9].

We recommend a short time-line to making the necessary decisions and implementing all changes now. It is the most cost-effective strategy. We also recommend adopting these decisions agency-wide, making public the macros to implement these decisions, and mandating agency-wide use of these macros for reports based on trends. This would allow analysts at the state and local levels to changes their programs at the same time and avoid publishing data with disparate rates.

ENDNOTES

1 FHOP/CHS,CDHS (2003) Guidelines on Race/Ethnicity Data Collection, Coding and Reporting. Last accessed 23-Jan-2017 at: http://fhop.ucsf.edu/sites/fhop.ucsf.edu/files/wysiwyg/raceEth2003.pdf

2 See National Vital Statistics Statistics System, last accessed 23-Jan-2017at: http://www.cdc.gov/nchs/nvss/bridged_race.htm.

3 Parker JD, Schenker N, Ingram DJ, Weed JA, Heck KE, Madans JH (2004) Bridging between two standards for collecting information on race and ethnicity: An application to Census 2000 and vital rates. Public Health Reports, 119:Mar-Apr, 192-205. Last accessed 23-Jan-2017 at: https://wonder.cdc.gov/wonder/help/populations/bridged-race/PublicHealthReports119-2-p192.pdf.

4 Office of Management and Budget (2000). Provisional guidance on the implementation of the 1997 standards for federal data on race and ethnicity. Last accessed 12-Apr-2011 at: http://www.whitehouse.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/re_guidance2000update.pdf.

5 Office of Management and Budget (2000). Appendix C: The Bridge Report: Tabulation Options for Trend Analysis. Provisional guidance on the implementation of the 1997 standards for federal data on race and ethnicity. Last accessed 12-Apr-2011 at: http://www.whitehouse.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/re_app-ctables.pdf.

6 Ingram DD, Parker JD, Schenker N, Weed JA, Hamilton B, Arias E, Madans JH. United States Census 2000 population with bridged race categories. National Center for Health Statistics. Vital Health Stat 2(135). 2003. Last accessed 23-Jan-2017 at: http://wonder.cdc.gov/wonder/help/populations/bridged-race/VitalHealthStatistics-Series2No135.pdf.

7 National Center for Health Statistics (2004). NCHS Procedures for Multiple-Race and Hispanic Origin Data: Collection, Coding, Editing, and Transmitting. Division of Vital Statistics, National Center for Health Statistics, Centers for Disease Control and Prevention. May 7, 2004. Last accessed 23-Jan-2017 at: http://www.cdc.gov/nchs/data/dvs/Multiple_race_docu_5-10-04.pdf.

8 See: http://www.dof.ca.gov/Forecasting/Demographics/. Last accessed 23-Jan-2017

9 Phone conversation, Linda Remy and Mary Heim of DOF, 24-Feb-2010.

10 State of California, Department of Finance (2004). Suggested Allocations of the Multirace Category for Use with Population Projections by Race/Ethnicity for California and It's Counties 2000-2050. Sacramento, California, June 2004. Last accessed 24-Apr-2011 at: http://www.dof.ca.gov/research/demographic/reports/projections/allocations-multirace_category/.

11 Phone conversation, Linda Remy and John Molson of DOF, 15-Apr-2011.

12 Phone conversation, Linda Remy and Melanie Martindale, DRU/DOF, 18-Apr-2011. See Appendix A.

13 National Center for Health Statistics (2010). Postcensal estimates of the resident population of the United States for July 1, 2000-July 1, 2009, by year, county, age, bridged race, Hispanic origin, and sex (Vintage 2009). Prepared under a collaborative arrangement with the U.S. Census Bureau; released June 20, 2010. Available from: www.cdc.gov/nchs/nvss/bridged_race.htm as of 23-Jan-2017.

14 Heim M (2007). Future population growth trends in California. Demographic Research Unit, Department of Finance. Last accessed of 23-Jan-2017at: http://www2.bren.ucsb.edu/~keller/energy-water/1-2%20Mary%20Heim.pdf.

15 Email, Mary Heim (DOF) to Linda Remy, 24-Feb-2010.

16 See: http://www.whitehouse.gov/omb/fedreg_1997standards. Last accessed 29-Apr-2011.

17 Remy LL, Oliva G, Fujimoto S, McKendry J, Doebbert G (2010). Creating consistent coding for race and ethnicity for birth and death certificates in California.

18 McKendry CJ (2001). Race variables, multiple races. CDPH-OHIR. Provided as part of documentation for Vital Statistics files.

19 Martin JA, Hamilton BE, Sutton PD, Ventura SJ, et al. Births: Final data for 2006. National vital statistics reports; 57:7. Hyattsville, MD: National Center for Health Statistics. 2009.

20 Remy LL, Oliva G, Fujimoto S, McKendry J (2010). Multi-race coding in California birth and death certificates. UCSF Family Health Outcomes Project.

21 Email, David P Johnson of the CDC, to Linda Remy, 31-May-2011.

22 Johnson DP (2008). Using the Multiple Race Bridging Matrix. Documentation and NCHS race bridging file emailed to Linda Remy on 27-May-2011 from Dave Johnson, Survey Statistician, DHHS/PHS/CDC/CCHIS/NCHS/DVS/SPSRB. CDC/NCHS uses this file to bridge multi-race for all Federal vital statistics publications. Per Dave Johnson, CDC revisited whether to update the file in 2010, determined it did not need to be updated.

23 Phone conversation, Linda Remy with Debra D. Ingram of the CDC, 06-May-2011.

24 See Population Studies Center, last accessed 23-Jan-2017 at: http://www.psc.isr.umich.edu/dis/census/subject/race.html.