



Technical Guide for Using Title V 5-Year Needs Assessment Indicators Databooks 1994-2005

■ Linda L. Remy, PhD
Ted Clay, MS
Geraldine Oliva, MD MPH

December 2007



Family Health Outcomes Project
University of California, San Francisco
www.ucsf.edu/FHOP

Funding provided by the Maternal Child and Adolescent Health Branch,
California Department of Public Health

TABLE OF CONTENTS

Layout	1
Calculating Annual Rates.....	2
Example Data Quality Tab	4
Example Rate Tab	5
Displaying and Testing Trends	6
Appendix A: Other Statistical Issues.....	11

TABLE OF TABLES AND FIGURES

Table 1: Example Kotelchuck Index Data Quality Tab	3
Table 2: Fertility Rate Per 1,000 Female Population Age 15 to 17 by Race/Ethnicity	5
Figure 1: Fertility Rate Trend Per 1,000 Female Population Age 15 to 17.....	6
Table 3. Fertility Rate Trend Regression Results Per 1,000 Female Population Age 15 to 17....	7
Table 4. Fertility Rate Trend Regression Results Per 1,000 Female Population Age 40 to 44....	8

Suggested Citation:

Remy LL, Clay T, Oliva G. (2007) Technical Guide for Using Title V 5-Year Needs Assessment Indicators: Databooks 1994-2005. San Francisco, CA: University of California, San Francisco, Family Health Outcomes Project. Available at: <http://www.ucsf.edu/fhop>.

TITLE V 5-YEAR NEEDS ASSESSMENT INDICATORS

The layout for the set of Title V Indicator Databooks has changed significantly. The purpose of this document is to provide an overview of the changes and to introduce implications for using the Databooks to inform local monitoring and planning activities.

LAYOUT

The revised databook contains a minimum of four types of tabs: Data, Data Quality, Rates, and Graphs. A few databooks also have a Definition tab.

Data Tabs. As in previous years, the Data Tabs are a 2-tab set of 12 years of local jurisdiction and state data summarizing required and optional indicators. Each Data Tab presents the numerators with the appropriate denominators to calculate rates. Data presented here are the source for calculating Rates and Graphs tab for each indicator.

Definition Tab. A few indicators include a Definitions Tab to provide more specific information as to how they were calculated. These are mainly indicators using a population subset, e.g., births to mothers age 15 to 44, singleton births. In these instances, the total reported will be less than the total for the source data, e.g., all births vs. singleton births. Other indicators require adding cases from multiple data sources, e.g., births and fetal deaths. For these indicators the total number typically reported will be greater than either file independently.

The Data Quality Tab focuses on the last year in the 12-year trend. It identifies key data quality issues that may impact reliability of information used to calculate the indicator. Jurisdictions are advised to review this tab carefully to understand if underlying quality issues affect their data to such an extent that the validity of their local statistics may be compromised. Jurisdictions with proportionally more exclusions may have distorted rates due to the smaller numbers of cases used. In these cases it is difficult to know if the county truly is an outlier for the indicator or if the results are affected by a problem with the underlying data. If data quality appears to be compromised, jurisdictions are urged to be extremely careful in reporting their data.

The Rate Tab incorporates information previously available in the FHOP data templates. This table is preset to print in three pages.

Page 1 defines the indicator and its risk factors. Below this is the rate table for total cases. It may have as many as 12 or as few as no data rows, depending on the number of cases available each year. Rules for determining the number of rows are in the next section of this document. Below the rate table on page 1, we identify the data sources and additional analyses you might wish to undertake to understand your jurisdiction's performance on the indicator.

Page 2 of the rate table summarizes rates for White Non-Hispanic and Hispanic All-Race, and Page 3 summarizes results for Non-Hispanic African-American and Asian populations. We used the same rules to calculate these tables as we used for the Total table.

The Graph Tab new this year. This summarizes results of trend tests for data on the Rate Tab. As with the Rate Tab, page 1 summarizes results for all data, page 2 for White Non-Hispanic and Hispanic All-Race, and page 3 for Non-Hispanic African American and Asians. Methods for calculating trend statistics are described in Appendix A.

CALCULATING ANNUAL RATES

MINIMUM NUMBER OF EVENTS

If the minimum number of events over all years was greater than or equal to 10 in each year, the 12 years of indicator data was left as given.¹ If the minimum number of events in any period was less than 10, we aggregated the data into six 2-year periods. If the minimum still was less than ten, we aggregated into four 3-year periods. When all four 3-year periods still did not meet the minimum of 10, we declared the number of cases not big enough and did not calculate rates. These are shown on the Rates tab by the phrase "Rates not calculated." Otherwise, the Rates Tab tables show the periods, numerators, rates, and rate confidence intervals.

Given the final level of aggregation for the local data, the same aggregation was performed on the corresponding California data. Then the local and state data were merged for side-by-side presentation in the Rate Tab.

DENOMINATORS

We used two different types of denominators: Annual county-level population estimates from the California Department of Finance (DOF), and counts based on qualifying records from the birth certificates and/or fetal death certificates.

The DOF one-year age categories were summarized as needed for the indicators that use state population as their denominators. Some indicators use only certain ages in the female population for denominators, others use all population or all population in given age categories.

Some indicators have as their denominator the total number of records in a given category and file, for example from the birth certificates file, the number of women age 15 to 44 delivering a live born infant. Others need refinement: women age 15 to 44 delivering a live born singleton infant. Before deciding if a record meets the minimum qualifying condition(s), our macros check to be sure we are excluding invalid or unlikely cases (e.g., age missing, 92 year old mother, 27-month gestational age, 5 gram birthweight). We excluded cases outside allowable ranges and report their frequency on the Data Quality Table described in a later section.

BRIDGING RACE/ETHNICITY

The various data sources used had different definitions of race/ethnicity over time. After the 2000 Census, the federal government issued bridging guidelines.² They recommend that researchers use their groupings longitudinally until sufficient years permit more detailed analyses. Allocations for longitudinal or trend studies differ from civil rights monitoring.³

-
- 1 Note that the minimum numbers we used are smaller than those suggested in FHOP's "Small Numbers Guidelines." Here the focus is on longitudinal trends rather than one-year comparisons, and the statistical tests are based on counts rather than rates.
 - 2 Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity. Executive Office of the President, Office of Management and Budget, Washington, D.C. 20503. December 15, 2000. Last accessed 27-Dec-2007 at: <http://www.ohrc.gov/templates/browse.aspx?lvl=2&lvid=172>.
 - 3 Office of Management and Budget. 09-Mar-2000. OMB BULLETIN NO. 00-02: Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement. Last accessed 27-Dec-2007 at: <http://www.whitehouse.gov/omb/bulletins/b00-02.html>.

California requires state-funded researchers to use Department of Finance (DOF) population estimates.⁴ The DOF provides county-level estimates by sex and race/ethnicity, with age in 1-year intervals. Through 1999, race/ethnicity was categorized as White, Black, Hispanic (all races), Asian and Pacific Islander, and American Indian.⁵ In coding race/ethnicity, FHOP follows DOF rules. First, cases are assigned to Hispanic all-race. The remainder are assigned to White, Black, Asian/Pacific Islander, and American Indian. Groups that do not fit these classifications are assigned to White race/ethnicity. By using these categories, jurisdictions will be able to calculate population rates using DOF-calculated population estimates. FHOP does not recommend calculating statistics for the non-Hispanic American Indian/Native American. DOF rules diminish their numbers so significantly that they are not reliable.

For 2000 and later, DOF classifies race as White, Black, Hispanic, Asian, Pacific Islander, American Indian, and Multi-race.⁶ To make race/ethnic classifications compatible longitudinally, we combined the 2000-and-later race categories as follows. First, we combined Asian with Pacific Islander. Then we broke apart the Multi-race category using the DOF multi-race allocation table, which gives the percentage to allocate to each race, within each county separately.⁷ Note that the Hispanic category had no multi-race allocation because the DOF made assignment to this category before assigning to other single- or multi-race categories. Also note that the DOF allocation table was based on the 2000 census, and we use it for years later than 2000 as well, as the DOF guidelines recommend.

We did other bridging reclassifications for records in the birth, death, and fetal death certificates, and for records from the hospital discharge data in order to attain compatibility across numerator and denominator sources. This was accomplished using SAS macros. The result is five race/ethnic groups: White, Black, Hispanic All-Race, Asian, American Indian. In calculating total population rates, we include all cases, but we do not calculate rates for American Indians. Their numbers are small and we believe unreliable because of definitional issues.

CALCULATING RATES AND CONFIDENCE INTERVALS

Rates were calculated overall and for each race/ethnic group by dividing the numerator (number of events) by its appropriate denominator (population, births, deaths, etc), and multiplying the result times the appropriate factor (e.g. rate per 1000 births, etc.). The confidence interval for the rate was calculated using the Wilson score without continuity correction.^{8 9}

4 See: http://www.dof.ca.gov/html/Demograp/DRU_datafiles/DRU_datafiles.htm. Last accessed 27-Dec-2007.

5 State of California, Department of Finance, Race/Ethnic Population with Age and Sex Detail, 1990–1999. Sacramento, CA, May 2004. Last accessed 27-Dec-2007 at: http://www.dof.ca.gov/html/Demograp/DRU_datafiles/Race/RaceData_90-99.htm

6 State of California, Department of Finance, Race/Ethnic Population with Age and Sex Detail, 2000–2050. Sacramento, CA, May 2004. Last accessed 27-Dec-2007 at: http://www.dof.ca.gov/html/Demograp/DRU_datafiles/Race/RaceData_2000-2050.htm

7 State of California, Department of Finance, 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 27 Dec 2007 at: <http://www.dof.ca.gov/HTML/DEMOGRAP/ReportsPapers/Projections/Allocations/MultiraceAllctns2000-2050.php>.

8 Newcombe RG. Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Statist. Med.* 17, 857-872 (1998). Note that this is different from that recommended in FHOP's "Small Numbers Guidelines," which focuses on easy-to-calculate statistics.

EXAMPLE DATA QUALITY TAB

The Data Quality Tab focuses on the numbers of cases set aside in the last year of the 12-year period for each grouped set of indicators. Table 1 presents an example of a Data Quality Tab. The top part identifies the indicator that has been pre-screened for data quality, in this case the Kotelchuck Index for Adequacy of Prenatal Care. Then it describes the implications in terms of whether a county can rely on the reasonableness of the resulting index. This table shows only those counties on the first page of the table. The column percentages do not add to 100.

Table 1: Example Kotelchuck Index Data Quality Tab

Kotelchuck Index. In 2005, 546,443 births were recorded to California resident women age 15 to 44. Statewide, 7,783 birth certificate records (1.4%) were missing one or more data elements needed to calculate the Kotelchuck Index. At the county level, an average of 1.4% of birth records (0.7% median) were missing one or more elements of the Kotelchuck Index.

Implications. Study this table carefully. If more than 0.7% (the county-level median) of your area's birth certificates were missing one or more data elements needed to calculate the Kotelchuck Index, rates will be increasingly inaccurate. Further research may be needed to find out how your area can improve its data quality. Until this is resolved, be increasingly skeptical of Kotelchuck Adequacy of Prenatal Care rates as the Area percent missing increases.

Local	Total Births		Age 15 to 44		Missing		
	Number	Percent	Number	Percent	Number	Area %	Miss %
State	548,700	100.00	546,443	100.00	7,783	1.42	100.00
1 Alameda	20,902	3.81	20,803	3.81	79	0.38	1.02
2 Alpine	15	0.00	15	0.00	0	-	-
3 Amador	288	0.05	286	0.05	1	0.35	0.01
4 Butte	2,451	0.45	2,445	0.45	23	0.94	0.30
5 Calaveras	371	0.07	370	0.07	2	0.54	0.03
6 Colusa	381	0.07	380	0.07	2	0.53	0.03
7 Contra Costa	13,143	2.40	13,093	2.40	205	1.57	2.63
8 Del Norte	327	0.06	327	0.06	2	0.61	0.03
9 El Dorado	1,930	0.35	1,922	0.35	22	1.14	0.28
10 Fresno	15,936	2.90	15,879	2.91	92	0.58	1.18
11 Glenn	431	0.08	429	0.08	7	1.63	0.09
12 Humboldt	1,598	0.29	1,595	0.29	50	3.13	0.64
13 Imperial	3,058	0.56	3,054	0.56	82	2.69	1.05
14 Inyo	205	0.04	205	0.04	0	-	-
15 Kern	14,022	2.56	13,979	2.56	1,500	10.73	19.27
16 Kings	2,554	0.47	2,548	0.47	11	0.43	0.14
17 Lake	728	0.13	726	0.13	9	1.24	0.12
18 Lassen	289	0.05	288	0.05	2	0.69	0.03
19 Los Angeles	150,377	27.41	149,677	27.39	1,449	0.97	18.62
20 Madera	2,349	0.43	2,336	0.43	20	0.86	0.26
21 Marin	2,785	0.51	2,759	0.50	1	0.04	0.01

Counties with few records set aside can be more comfortable with their results than counties with more records set aside. For example, reviewing the column Missing Area %, Marin County can be relatively certain that their scores on the Kotelchuk Index are reliable. A county with more than 0.7% of their records set aside faces increasing uncertainty about their results.

9 A copy of the SAS programs we used to calculate rates and confidence intervals is available upon request.

EXAMPLE RATE TAB

Table 2 is an example of the Rate Tab for one indicator, in this case the number of live births per 1,000 women age 15 to 17. The top section includes definitions of the indicator, numerator, denominator, the Healthy People 2010 Objective, and risk factors associated with this indicator.

Table 2: Fertility Rate per 1,000 Female Population Age 15 to 17 by Race/Ethnicity

DEFINITION:	The number of live births per 1,000 women age 15-17						
NUMERATOR:	The number of live births to women age 15-17, by place of residence in a calendar year						
DENOMINATOR:	Population of women age 15-17, by place of residence in a calendar year						
HP 2010 OBJECTIVE	9-7: Reduce pregnancies among adolescents to 43 pregnancies per 1,000 (Baseline: 68 per 1,000 in 1996)						
RISK FACTORS:	Not applicable						

Year	TOTAL POPULATION							
	California				Local			
	Births	Rate	Lower C.L.	Upper C.L.	Births	Rate	Lower C.L.	Upper C.L.
1994	26,378	42.3	41.8	42.8	853	65.9	61.7	70.3
1995	25,821	40.4	39.9	40.9	901	68.1	64.0	72.6
1996	24,047	36.4	36.0	36.9	798	58.6	54.8	62.7
1997	23,064	33.8	33.4	34.2	782	54.4	50.8	58.2
1998	21,630	31.1	30.6	31.5	766	50.2	46.8	53.8
1999	20,209	28.8	28.4	29.2	747	46.3	43.1	49.6
2000	18,887	26.6	26.2	26.9	716	42.9	39.9	46.1
2001	17,307	23.8	23.5	24.2	591	34.5	31.9	37.3
2002	16,660	22.4	22.0	22.7	636	36.2	33.6	39.1
2003	16,193	21.1	20.8	21.4	678	37.8	35.1	40.7
2004	16,263	20.6	20.3	20.9	661	36.3	33.7	39.1
2005	16,740	20.3	20.0	20.7	671	35.9	33.3	38.7

Sources: [Definition:](http://www.cdc.gov/nchs/datawh/nchsdefs/rates.htm#birth) <http://www.cdc.gov/nchs/datawh/nchsdefs/rates.htm#birth> last accessed 18 Apr 05

Numerator: California Center for Health Statistics, Vital Statistics, Births Statistical Master File. To order: <http://www.dhs.ca.gov/hisp/chs/OHIR/Catalog/DataCatalog.htm>. Last accessed 19 Apr 05.

Denominator: 1990-1999: State of California, Department of Finance, Race/Ethnic Population with Age and Sex Detail 1990-1999. 2000-2050 Projections: State of California, Department of Finance, Race/Ethnic Population with Age and Sex Detail, 2000-2050. Sacramento, CA, May 2004. Both files available at: http://www.dof.ca.gov/html/Demograp/DRU_datafiles/DRU_datafiles.htm. Last accessed 19 Apr 05.

Recommended Tables:

Can be analyzed using EpiBC:

- Births by mother's race/ethnicity -- Review Fertility Birth Rate Tables
- Births by mother's age
- Births by geographic area (ZIP code, if available)
- Map of distribution of births by geographic area (ZIP code)
- Births by parity

Notes: 2000-2050 Multirace/ethnic population projections must be allocated before use. See State of California, Department of Finance, 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. <http://www.dof.ca.gov/HTML/DEMOGRAP/MultiraceAllocns2000-2050.htm>. Last accessed 19 Apr 05.

C.L. = Confidence Limit = the boundary of the 95% confidence interval.

The Rate Table shows data for the total population in California and the local jurisdiction, over the most recent 12-year period for which data are available, in this case 1994 through 2005. For both California and the jurisdiction, columns indicate the number of events, the rate, and lower and upper confidence limits for that rate.

Below the rate table, we present information on the sources for the definitions of the indicator, the numerator, and the denominator. Recommended tables that can be made using EpiBC (if a birth-certificate indicator) or EpiHosp (if a discharge data indicator) are shown. Finally, we present any notes that we think are important to understand the data presented. On pages 2 and 3 of a given Rate Tab, we show the same rate-based data for the other race/ethnic groups, as described earlier.

Comparing the state and local rates and confidence intervals on Table 2, from start to end of period, we see that the local rate and its upper confidence intervals are always below the state rate. This allows us to conclude that the local rate was significantly lower than the state throughout the time period. In both cases, we also see that the rates appear to be dropping. This leads us to wonder if the drop in rates is statistically significant. Is the rate of change the same for the jurisdiction and state? That is, do we have a trend, and if so, is it linear or curvilinear? How does the jurisdiction compare with the state?

DISPLAYING AND TESTING TRENDS

If rates are on the Rates Tab, as for our example, we display local and state rates over the period on the Graph Tab and show results of statistical tests for trends and rate differences. If the data permit 12 years of rates, as they do here, we also test for non-linear trends.

THE TREND GRAPH

Figure 1 shows the graph made using the rates from Page 1 of the Graph Tab for our example indicator. Depending on the number of events the data allow us to calculate, a given graph might have 12, 6, 4, or no points. Note that the graphs do not include confidence intervals, which would be inappropriate for a trend test. Here, we focus on the rate of change over time rather than within a given period such as a year.

Figure 1: Fertility Rate Trend per 1,000 Female Population Age 15 to 17

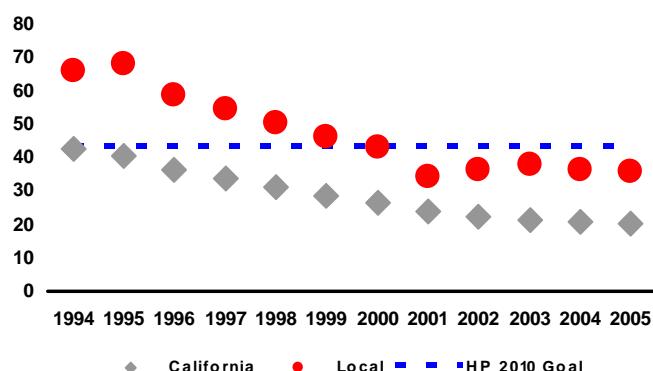


Figure 1 has 12 points, because the number of events was greater than 10 in all years. If the indicator had too few cases, the graph would be empty. Red dots (●) show the local jurisdiction rate and grey diamonds (◆) show the state rate. When an indicator has a Healthy People (HP) 2010 objective, it is shown with a dashed blue line. For this indicator, the HP 2010 objective is 43 births per 1,000 female population age 15 to 17.

STATISTICAL TESTS

We evaluated the indicator data on the rate tab two ways. First, we evaluated it to see if the trend we see is statistically significant. The emphasis here is on the rate of change over time, asking in essence if the local jurisdiction is changing at a rate similar to or different from the state and in what direction. Second, we tested if the local jurisdiction rate is significantly different from the state at the start and end of our time period. The emphasis here is whether local jurisdiction results are lower than, higher than, or the same as the state. Table 3 reports the results of these two types of statistical tests for our example: trend test and difference test.

Trend Test. Sometimes the "eyeball" test may make you think that significant linear trends are present when in fact the data are not linear or the trend is not significant. For example, the linear trend test may be non-significant because the data has some other shape like \cap or \cup . When the data has another shape, it would be inaccurate to describe a trend as linear.

We used JoinPoint software to test for trends. When JoinPoint finds that one or more segments of a time period have slopes that are significantly different from the slope of the previous period or finds significant shifts in the intercept, it breaks the trend into segments. This means that JoinPoint found a bend, or change in the angle of the trend line or intercept that is statistically significant from the previous segment at $P = 0.05$ or less. If bends were found, JoinPoint defines the time-based segment and describes the slope and intercept during each period. Each time the segment slope changes, its intercept also changes.

We allowed JoinPoint to search for segments if the series had 12 periods, as in our example. We report the multiple segment model when one or more segments have a slope significantly different from zero. Otherwise we report the simple linear model. If the numbers of local cases were too few to test 12 periods, we restricted JoinPoint to a simple linear trend test.

Table 3. Fertility Rate Trend Regression Results per 1,000 Female Population Age 15 to 17

Level	Date Range	Intercept		Slope			Sig?
		Est.	Std. Err	Est.	Std. Err	P-Value	
State	1994-2001	42.28	0.39	-2.70	0.10	0.000	Yes
	2001-2005	29.42	1.65	-0.86	0.17	0.001	Yes
Local	1994-2001	68.80	1.37	-4.62	0.34	0.000	Yes
	2001-2005	36.48	5.67	-0.01	0.59	0.992	No
State Avg	1994-1996	39.67	0.14				
Local Avg	vs State	64.15	1.23			0.000	Yes
State Avg	2003-2005	20.67	0.09				
Local Avg	vs State	36.65	0.80			0.000	Yes

State Trend. The first data row in the Level column contains the word "State." The "Date Range" column shows one or more time intervals. If a simple linear model was justified, one line of State results shows, and the Date Range column shows the 12-year range. In our example, JoinPoint found two segments, so two rows show with two ranges.

For each trend segment, the next column sets show the intercept and its standard error, the slope and its standard error, and the P-value associated with the slope. In the last column, we identify by "Yes" or "No" if the slope in the time segment is significantly different from zero at $P = 0.05$ or less.

JoinPoint rejected the hypothesis that the best fit for the state trend is linear and found that a 2-segment model offered the "best" description of the data. By best, we mean that this model explained more significantly variance than the simple linear model. From 1994 to 2001, the state declined at a rate of -2.70, as shown by the P-value of 0.000, then continued to decline significantly but less steeply (slope -0.86, P-value 0.001).

Local Trend. The first column contains the word "Local". In all other respects reporting of the local trend is like the state trend discussed above. For our example, JoinPoint found the local trend (-4.62, P = 0.000) was strongest between 1994 to 2001, then flattened thereafter (-0.01, P-value = 0.992). For more detailed information on the how the jurisdiction did, look at the slope estimate section of the table. The last column summarizes the "big picture": "Yes" means the trend is significant at P = 0.000 or less through 2001 and non-significant thereafter.

If the local jurisdiction has too few cases to calculate trends, we show the text "Too few cases", and the rest of the table is blank.

Difference of Slopes Test. If JoinPoint accepts a linear model for both the State and jurisdiction, we test to see if their slopes are significantly different from each other. That is, we test the null hypothesis that the local and state trends are equal, assuming the data are independent and had a normal distribution. In essence, we perform a T-test on the slopes.

Since JoinPoint found both state and models were nonlinear we do not show a difference of slopes test for it. When one area has a linear trend and another area has a curvilinear trend, or in this case when both are curvilinear, we cannot test if the slopes for these trends are different. But we can show the difference of slopes test for the fertility rate results for females age 40-44 in this same jurisdiction.

Table 4. Fertility Rate Trend Regression Results per 1,000 Women Age 40-44

Level	Date Range	Intercept		Slope			Sig?
		Est.	Std. Err	Est.	Std. Err	P-Value	
State	1994-2005	10.23	0.10	0.22	0.02	0.000	Yes
Local	1994-2005	8.34	0.39	0.14	0.06	0.043	Yes
	Different?					0.151	No
State Avg	1994-1996	10.58	0.05				
Local Avg	vs State	9.17	0.38			0.000	Yes
State Avg	2003-2005	12.59	0.05				
Local Avg	vs State	10.00	0.35			0.000	Yes

Table 4 shows that both the state and local jurisdictions had significant linear trends. When both trends are linear, line 3 of the "Level" column is blank and the "Date Range" column contains the word "Different?". The P-value is from the test that the local slope equals the state slope over the 12-year period.

Again, a "Yes" or "No" indicates whether the P-value is statistically significant. In this case, both trends are statistically significant, but the local rate of change is not statistically different from the state rate of change. In any circumstance where the difference of slopes test is not statistically significant, report the state rate, since the local rate is not different from that, and the state rate has more power. Only when the two trends are significantly different should the local trend be reported in favor of the state trend.

Difference of Rates Tests. The trend tests above describe the rate at which outcome indicators are changing and their direction (improving, getting worse, or no change), and, if the trends are simple linear, whether the local jurisdiction is changing at the same or a different rate than the state. But these do not answer a question of central importance to most jurisdictions: Did they do better, the same as, or worse than the state as a whole? To approach this question, we did difference of rates tests, comparing the local jurisdiction to the state for the first three years and the last three years of the trend period. These results also are shown on the graph tab, with Tables 3 and 4 containing the example results.

State Average in First 3 Years. The first column shows "State Avg", and the Date Range column identifies the first 3 years in the period. The three-year average state rate and its standard error

are displayed in the Intercept Estimate and Standard Error columns. In the first three years of the example period for Table 3, the average state rate was 39.67 births per 1,000 women age 15 to 17. In the first three years of the example period for Table 4, the average state rate was 10.58 per 1,000 women age 40-44. In both cases, notice that the jurisdiction standard error is wider than the state, since the jurisdiction is smaller.

Local Average in First 3 Years and Comparison to State. This is located on the line after the State Average. The first column contains "Local Avg" and the Years column contains the words "vs State". The three-year average local rate and its standard error are displayed in the Intercept Estimate and Standard Error columns. In the first three years of the trend period for Table 3, the local jurisdiction rate was 64.15 births per 1,000 women age 15 to 17. In the first three years of the trend period for Table 4, the local jurisdiction rate was 9.17 per 1,000 women age 40-44.

The "P-value" column shows the P-value for the test that the state and local 3-year rates are the same. "Yes" indicates that the P-value is less than 0.05 and the state and local rates are significantly different. "No" indicates that they are not significantly different. The example local jurisdiction had a statistically higher rate than the state during the 3-year period for teen births and lower than the state for older mothers.

State Average in Last 3 Years. Like first 3 Years described above, with the State as a whole having a rate of 20.67 births per 1,000 women age 15 to 17 and a rate of 12.59 births per 1,000 women age 40-44.

Local Average in Last 3 Years and Comparison to State. This is like the first three years described above, with the local jurisdiction having a rate of 36.65 births per 1,000 women age 15 to 17 and a rate of 10.00 per 1,000 women age 40-44. The "Yes" in the last column, and the P = 0.000 indicate that the local jurisdiction had statistically higher rates for teen births and a lower rate for older mothers than the state during the last three years in the 12-year period.

CONCLUSION

Reality is often complex, and complexity trumps simplicity when reporting statistical results. Thus when a statistically significant nonlinear trend exists, the interpretation must reflect this even in the presence of a statistically significant linear trend. In our example, the jurisdiction trend for teen births is nonlinear for both the state and local jurisdiction. In such a circumstance, the trends cannot be directly compared statistically and each must be summarized separately. Only when the best description for both trends is linear can their rates of change be compared.

The local rate of live births per 1,000 females age 15-17 declined linearly at the statistically significant rate of -4.62 per 1,000 live births during the period 1994-2001 with no significant change thereafter. Similarly, most of the state's decrease occurred between 1994-2001, at a rate of 2.70 per 1,000 live births, but it continued to drop at a lower but still significant rate thereafter. The local rate was significantly above the state rate at the start and end of the period.

On the other hand, the state and local jurisdiction each had a statistically significant trend for increased births to women age 40-44. The local trend was not significantly different from the state trend of 0.22 per 1,000 population.

In this jurisdiction, young women age 15-17 were much more likely to have a baby than the state average. On the other hand, older women in this jurisdiction were less likely to have a baby than the state average.

Over the period, both the state and jurisdiction rates fell below the Healthy People 2010 Objective for teen births (43 per 1,000). No objectives are set for births to older women.

Readers with a non-statistical background will find it most helpful to focus on the last column of the trend results table. If it says "Yes," the trend is significant at $P = 0.05$.

Those who would like a statistical refresher may wish to review the FHOP monograph, "Do We Have a Linear Trend?"¹⁰

10 Remy LL, Clay T, Oliva G. (Oct 2005) Do We Have a Linear Trend? A Beginner's Approach to Analysis of Trends in Community Health Indicators. Last accessed 27-Dec-2007 at:
http://www.ucsf.edu/fhop/_htm/publications/index.htm.

APPENDIX A: OTHER STATISTICAL ISSUES

The entire process of creating the Databooks is macro driven using SAS. The process starts by summarizing the indicator data and the appropriate denominator data to the geographic level of interest (state, county, super region (e.g., Bay Area) or sub-region (Berkeley, Long Beach, Pasadena, LA County Service Provider Areas). We calculate rates, feed the numerators and denominators into JoinPoint, bring the results back into SAS, and output the results directly into preformatted Excel template files.

We use JoinPoint to estimate linear trends for the jurisdiction and state and to test whether the resulting slope for each trend is significantly different from zero.^{11 12} The Statistical Research and Applications Branch (SRAB) of the National Cancer Institute developed JoinPoint as one among a set of new statistical methods and associated software tools for the analysis and reporting of cancer statistics. This group of powerful shareware statistical packages is appropriate for the analysis of any population-based data.¹³ In this set of software, JoinPoint was developed explicitly to estimate linear and curvilinear trends.

JoinPoint takes trend data and fits the simplest trend model that the data allow. The user supplies the minimum and maximum number of joinpoints. The program starts with the minimum number of joinpoints (e.g. 0 joinpoints, which is a straight line, or a standard trend test) and then tests whether more joinpoints are statistically significant and must be added to the model (up to that maximum number). This enables the user to test if an apparent change in trend from one period to another is statistically significant. The tests of significance use a Monte Carlo Permutation method.

Models may incorporate estimated variation for each point (e.g. when the responses are age adjusted rates) or use a Poisson model of variation. In addition, models may be linear on the log of the response (e.g. for calculating annual percentage rate change). The software allows viewing one graph for each joinpoint model, from the model with the minimum number of joinpoints to the model with maximum number of joinpoints.

Joinpoint uses a complex algorithm to decide whether to add a bend ("join point") to a simple linear model. For each line segment, or for the whole time period, Joinpoint tests whether the slope is equal to zero (i.e. the line is flat). When calculating the standard error of the slope of a segment it ignores points at the bend, which reduces the effective sample size when calculating the standard errors of the slopes. Consequently, Joinpoint tests comparing the slopes to zero are quite conservative.

Joinpoint program options we used to calculate trends are: 1) input numerators and denominators, 2) test for bends at whole years, 3) use a minimum of two years between bends and between a bend and either end of the data, 4) test for a maximum of 2 bends, 5) fit a linear (not log-linear) model with uncorrelated errors.

11 The Joinpoint software program was obtained at: <http://srab.cancer.gov/joinpoint/>. Last accessed 27 Dec 2007.

12 Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335-51 (correction: 2001;20:655).

13 See Methods & Software for Population-Based Cancer Statistics: <http://srab.cancer.gov/software/>. Last accessed 27 Dec 2007.

We decided to use the single line model unless one or more line segments had a slope significantly different from zero. If both the local and state data used a simple linear model, we tested the equality of the two slopes by using the estimates and standard errors reported by Joinpoint. The standard error of the difference was calculated as the square root of the sum of the squared standard errors.

In calculating state and local three-year rates, we used the total of the numerators and the average of the denominators over the 3 years. Then we divided the resulting rates and confidence limits by 3. In this calculation we used the same methodology as used to calculate 1-year rates and confidence intervals displayed on the rates table, namely the Wilson approximation method. The 95% confidence intervals were converted into an estimated standard error by dividing the width of the confidence intervals in half and then dividing by 1.96. We tested the equality of the two rates using the estimated rates and standard errors.