

California Department of Public Health

# Healthy Communities Data and Indicators Project: Website Toolkit

Adapted from:

Maizlish N, Tran D, Bustamante-Zamora DM. Healthy Communities Data and Indicators Project: How-To Manual and Illustrative Guide. Richmond, CA and San Francisco, CA: California Department of Public Health Office of Health Equity and University of California, San Francisco, Institute for Health & Aging; 2014. For original document please contact: [HClOHE@cdph.ca.gov](mailto:HClOHE@cdph.ca.gov).

12/01/2017

## Contents

The Healthy Communities Data and Indicators Project (HCI).....	2
How Were the Indicators Chosen? .....	2
How Were the Indicator Data Files and Templates Created? .....	3
What Is an Indicator? .....	3
Indicators and Comparisons .....	5
Data Files and Templates.....	5
Most ESRI Story maps include: .....	8
Interpreting Results .....	8
How Do Validity, Precision, and Bias Affect Indicators? .....	8
Causation vs. Correlation .....	10
Comparisons .....	11
Interpretation Framework .....	11
Use Cases .....	13
The Social Determinants of Health and their Link to HCI .....	13
Examples of HCI in Action .....	16
References .....	19
General Reading and Resources .....	19

## The Healthy Communities Data and Indicators Project (HCI)

The *Healthy Community Data and Indicators Project* (HCI) was started in 2012 with seed money from the Strategic Growth Council, a state agency supporting efforts to promote sustainability in California. The goal is to provide a standardized set of statistical measures (indicators), data files, and tools for planning healthy communities and evaluating the impact of plans, projects, policy, and environmental changes on community health. The indicators were based on the [Healthy Community Framework](#) developed by the [Health in All Policies \(HiAP\) Task Force of the Strategic Growth Council](#) with input from community stakeholders and public health organizations. The Healthy Community Framework identifies 20 key attributes of a healthy community through all stages of life, clustered in five broad categories.

### How Were the Indicators Chosen?

The staff of the California Department of Public Health conducted bibliographic reviews of numerous city, county, state, and national indicator projects and applied criteria (Table 1) to select indicators. More than 200 possible indicators were first matched to the content areas of aspirational goals in the *Healthy Community Framework*. Second, the indicators without detailed data at the level of census tract, town, or other small geography were filtered out. The remaining indicators were compared with indicators in state agency strategic plans for consistency. After taking into account feedback and comments of the HiAP Task Force, CDPH published a core list of 56 indicators in 2013. This list has been updated to 66 indicators as of November 2017 ([Appendix](#)).

**Table 1.** Criteria Used to Select Indicators

- 
- ▶ Validity of the indicator
    - Indicator measures what it purports to measure
    - Evidence links indicator to health outcomes
  - ▶ Technical feasibility and data properties of the indicator
    - Data source(s) are owned and collected by a recognized organization
    - Timeliness (time lag and frequency of updates)
    - Data quality (completeness, missing data, accuracy)
    - Geographic scale, census tract to region
    - Administrative accessibility (public vs. private, confidentiality, costs)
    - Current use and acceptability to stakeholders
    - Mechanics of data collection, aggregation, and reporting
  - ▶ Indicator is understandable and valuable to users
- 

Source: California Department of Public Health

## How Were the Indicator Data Files and Templates Created?

To produce statewide data, HCI used automated methods to download data from scores of public websites and process the data into Excel files with standard formats. These methods rely on statistical packages and programming techniques that are documented in our technical manual ([2015 version available online](#)). The goal was to provide detailed data by:

- time period,
- race/ethnicity, and
- geography

Depending on data availability, the data files have single years and/or three- to five-year aggregations between 2000 and present. The race/ethnicity categories follow those of the U.S. Census:

- Latino
- White (not Latino)
- African American (not Latino)
- Native Hawaiian/Other Pacific Islander (not Latino)
- Asian (not Latino)
- American Indian/Alaska Native (not Latino)
- Multiple Race (not Latino)
- Other (not Latino)

For geographic detail, we included the California statewide average, 1,523 cities, 8,057 census tracts, 58 counties, and 14 transportation planning regions. An Excel data file for one indicator could have 20,000 or more rows covering multiple geographies, time periods, and race/ethnicity groups. For some indicators, data were not available at the census tract or by race/ethnicity.

## What Is an Indicator?

An indicator is a number that describes the frequency of a characteristic in a population or geographic area. It is often a stand-in for a concept that is more complex or difficult to measure. For example, "the percentage of people who reside within a half mile of a park" is an indicator of geographical access to parks. It makes concrete a concept of access based on a reasonable walking distance of a half mile. Other measures of access to parks might include affordability (entrance fees), nearby transit stops or highway exits, personal safety, and hours and days of operation.

The actual values of indicators are specific to a place and time. An indicator can simply be thought of as what happened, to whom, when, and where. Many indicators of population frequency are a percentage or population rate (e.g., per 1,000 population).

$$\text{Indicator} = \frac{\text{What happened?}}{\text{To whom?}} \text{ and when and where?}$$

As a percentage, the denominator (“who”) is the number of people or households in the entire population of interest, and the numerator (“what”) are only those who experienced the event or situation. The indicator has a time period (“when”) and location (“where”) that the measurements were taken. For example, park access in Los Angeles County might be defined as:

$$\text{Percent of Los Angeles County residents living within a half mile of a park in 2010} = \frac{\text{Number of Los Angeles County residents living within a half mile of a park in 2010}}{\text{Number of Los Angeles County residents in 2010}} \times 100$$

In 2010, among the 6.4 million residents of Los Angeles, 3.4 million lived within a half mile of a park. Using this definition, 53 percent of county residents had access to a park within walking distance of their home.

$$53\% = \frac{3.4 \text{ million}}{6.4 \text{ million}} \times 100, \text{ Los Angeles, 2012}$$

Some indicators have numerators and denominators but are rates rather than percentages. For example, the number of serious bicycle injuries might be expressed in terms of the annual number of serious bicycle injuries per population.\* In the San Francisco Bay Area, in 2009, there were 237 serious bicycle injuries in a population of 6.97 million people.

$$\text{Rate of serious bicycle injuries per million population in the Bay Area in 2009} = \frac{\text{Number of serious injuries in Bay Area bicyclists in 2009}}{\text{Bay Area population in 2009}} \times 1,000,000$$

$$34 \text{ per million} = \frac{237 \text{ injuries}}{6,972,980 \text{ population}} \times 1,000,000 \text{ Bay Area, 2009}$$

If the rate is a decimal number with many leading zeroes, the result is multiplied by some factor of 10 (1,000 or 10,000 or 100,000 or 1,000,000) to make the rate easier to read. So, 0.000034 serious bicycle injuries per person above is equivalent to 34 injuries per million population\*. Where the population sizes may be different or changing, a common denominator like injuries per million population facilitates comparisons.

Finally, some indicators are expressed as an average of repeated measurements. For example, the average air levels of ozone are based on measurements from air monitors that sample air every hour of the day (24 measurements) at a specific location. The 24 measurements can be averaged as a single number for that day and location. The 365 daily averages might then be averaged for an annual average.

---

\* Because the same person can be injured in several different crashes, the measure of frequency is not a percentage.

## Indicators and Comparisons

An indicator value by itself may not tell us whether the results are good or bad. Comparing the results to a meaningful reference value provides context. Several natural reference values involve time, geography, and population subgroups. There may not be one correct reference value. You can explore several options by asking:

1. How do my community's (or county's, region's, or state's) current results compare with:
  - our community's indicator measured over the past several years?
  - other communities in our county, region, or state average in the most current year?
  - other communities that have the same size and demographics?
  - a goal set by ourselves or a state or federal agency?
  - the best outcome observed in the state?
  - an ideal goal?
2. Are there differences in indicator values between neighborhoods, race/ethnicity groups, or other socioeconomic characteristics in the community (or county, region, or state)?
3. If there are differences, do they occur in just a few indicators, or is there a pattern involving many indicators?

HCI provides several geographic options for comparisons as well as time series for the same community. Many issues should be considered in interpreting indicator results by time periods or geography. These are highlighted in **INTERPRETING RESULTS**.

## Data Files and Templates

The data source of an indicator's tables, graphs, and maps is located in an Excel data file. Templates or examples of tables, graphs, maps, and explanatory text are located in the PDF files. Both files can be downloaded from the [HCI website \(https://www.cdph.ca.gov/Programs/OHE/Pages/Healthy-Communities-Data-and-Indicators-Project-\(HCI\).aspx\)](https://www.cdph.ca.gov/Programs/OHE/Pages/Healthy-Communities-Data-and-Indicators-Project-(HCI).aspx).

Each Excel file follows a basic format, illustrated by the indicator for educational attainment (Figure 1). The first tab in the Excel worksheet contains the data for all available geographies, time periods, and races/ethnicities for the entire state of California. The second tab is a "data dictionary" that defines the column headers in the data file. Instructions to filter the data for time period, race/ethnicity group, and geography are in the third tab. The counties that compose the California regions in HCI are listed in the fourth tab.

The first two columns (A,B) of the data file identify the indicator. Column C gives the year(s) the data were analyzed or reported in the data source. The next columns (D, E) are a code and name of the different races/ethnicities. The next seven columns (F through L) provide geographic information. The "geotype" field describes the level of

geography for data in that row (RE=region, CT=census tract, PL=place/town/city, CO=county, etc.). Column M has the denominator of the indicator. Column N has the numerator, and column O is the indicator value. Columns P through S describe the statistical reliability of the indicator and have useful information for technically oriented users. Column T indicates the ranking of cities across the entire state from highest to lowest in 10 groups (deciles). Column U is a ratio of the specific geography and time period to the state average for that time period. A value of 1 means the census tract, place, or county is equal to the state average. For educational attainment, a ratio greater than 1 means the geography has higher (better) attainment than the state average; a ratio of less than 1 means the census tract or city has an educational attainment less than that of the state average. The last column (V) has the date the file was produced by the HCI and serves as a version reference. New versions will reflect corrections or additional years of data.



Of note, specific Excel files may have small variations in this basic layout, and some fields have missing data because data are not available or statistically reliable.

Figure 1. Layout of an Indicator Data File in Excel

What Happened? When? Who? Where? What Who? happened? Indicator Value Statistical Reliability Comparison

Indicator Identifier Year Reported Race/Ethnicity Geographic codes and names Denominator Numerator

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
ind_id	ind_definition	reportyear	race_eth_cod	race_eth_name	geotype	geotypevalue	geoname	county_name	county_fips	region_name	region_code	pop2010	pop2010_hs	pop2010_p_hs	SE	RSE	UL_95CI	LL_95CI	CA_deck	CA_FR	version	
2	369 High School or Greater	2000	9 Total	RE 01	01	Bay Area	Bay Area	01	3861886	459989	01	833645									1093288	4/2/2013 2:47
3	369 High School or Greater	2005-2007	9 Total	RE 01	01	Bay Area	Bay Area	01	4071921	478300	01	8620496	0.099336	0.7241391	86.09001	86.47962					1078562	4/2/2013 2:47
4	369 High School or Greater	2008-2010	9 Total	RE 01	01	Bay Area	Bay Area	01	4294465.6	4867925	01	86.67894	0.096829	0.7268859	86.48995	86.68872					1075421	4/2/2013 2:47
5	369 High School or Greater	2000	9 Total	RE 02	02	Butte	Butte	02	104303.73	132736	02	82.3									1073635	4/2/2013 2:47
6	369 High School or Greater	2005-2007	9 Total	RE 02	02	Butte	Butte	02	15536.3	132278	02	85	0.668693	4.4579534	83.68836	86.33064					10625	4/2/2013 2:47
7	369 High School or Greater	2008-2010	9 Total	RE 02	02	Butte	Butte	02	126823.27	140329	02	86.1	0.486322	3.4987208	85.16481	87.05319					1068238	4/2/2013 2:47
8	369 High School or Greater	2000	9 Total	RE 03	03	Central/Southeast Sierra	Central/Southeast Sierra	03	388292.84	127960	03	84.70838									132974	4/2/2013 2:47
9	369 High School or Greater	2005-2007	9 Total	RE 03	03	Central/Southeast Sierra	Central/Southeast Sierra	03	89477.386	30390	03	86.78867	0.55034	4.364066	85.73041	87.86634					1084858	4/2/2013 2:47
10	369 High School or Greater	2008-2010	9 Total	RE 03	03	Central/Southeast Sierra	Central/Southeast Sierra	03	32963.361	104337	03	89.67994	0.402489	3.9000636	88.89906	90.46882					1192654	4/2/2013 2:47
11	369 High School or Greater	2000	9 Total	RE 04	04	Monterey Bay	Monterey Bay	04	327782.07	448528	04	74.48684									9368836	4/2/2013 2:47
12	369 High School or Greater	2005-2007	9 Total	RE 04	04	Monterey Bay	Monterey Bay	04	338309.71	446574	04	75.7567	0.420921	1.7362349	74.33869	76.5817					9346959	4/2/2013 2:47
13	369 High School or Greater	2008-2010	9 Total	RE 04	04	Monterey Bay	Monterey Bay	04	347901.71	456731	04	76.57212	0.343996	144.36835	75.49789	76.84626					9345064	4/2/2013 2:47
14	369 High School or Greater	2000	9 Total	RE 05	05	North Coast	North Coast	05	67489.85	206396	05	80.31454									1052575	4/2/2013 2:47
15	369 High School or Greater	2005-2007	9 Total	RE 05	05	North Coast	North Coast	05	17786.84	208391	05	85.32846	0.460928	3.148468	84.42504	86.2388					1066806	4/2/2013 2:47
16	369 High School or Greater	2008-2010	9 Total	RE 05	05	North Coast	North Coast	05	188790.07	215449	05	87.42345	0.392864	3.1245798	86.65324	88.19366					1084858	4/2/2013 2:47
17	369 High School or Greater	2000	9 Total	RE 06	06	Northeast Sierra	Northeast Sierra	06	122978.21	142583	06	86.25026									112305	4/2/2013 2:47
18	369 High School or Greater	2005-2007	9 Total	RE 06	06	Northeast Sierra	Northeast Sierra	06	125413.35	145854	06	88.8486	0.46295	4.1443244	87.34278	89.75441					110867	4/2/2013 2:47
19	369 High School or Greater	2008-2010	9 Total	RE 06	06	Northeast Sierra	Northeast Sierra	06	130327.26	145032	06	90.55054	0.418991	4.4340218	89.72932	91.37076					1123456	4/2/2013 2:47
20	369 High School or Greater	2000	9 Total	RE 07	07	Northern Sacramento Valley	Northern Sacramento Valley	07	45461072	63272	07	7195022									939595	4/2/2013 2:47
21	369 High School or Greater	2005-2007	9 Total	RE 07	07	Northern Sacramento Valley	Northern Sacramento Valley	07	52812.363	69717	07	75.75249	0.920006	3.7942276	73.94928	77.5557					9346959	4/2/2013 2:47
22	369 High School or Greater	2008-2010	9 Total	RE 07	07	Northern Sacramento Valley	Northern Sacramento Valley	07	55451.364	72004	07	77.80081	0.802992	3.4919311	75.42695	78.57467					935345	4/2/2013 2:47
23	369 High School or Greater	2000	9 Total	RE 08	08	Sacramento Area	Sacramento Area	08	1024457.8	122328	08	83.7572									1090589	4/2/2013 2:47
24	369 High School or Greater	2005-2007	9 Total	RE 08	08	Sacramento Area	Sacramento Area	08	1228400.6	1427757	08	86.03709	0.381881	1.2975857	85.68898	86.39221					1075464	4/2/2013 2:47
25	369 High School or Greater	2008-2010	9 Total	RE 08	08	Sacramento Area	Sacramento Area	08	1272305.3	1479905	08	86.2098	0.182067	1.3301223	85.95813	86.66883					1070868	4/2/2013 2:47
26	369 High School or Greater	2000	9 Total	RE 09	09	San Diego	San Diego	09	1464788.1	1773327	09	82.6									1079521	4/2/2013 2:47
27	369 High School or Greater	2005-2007	9 Total	RE 09	09	San Diego	San Diego	09	1592063.5	1875222	09	84.9	0.128561	0.8058832	84.6637	85.1383					10625	4/2/2013 2:47
28	369 High School or Greater	2008-2010	9 Total	RE 09	09	San Diego	San Diego	09	1622935.9	1977598	09	85.1	0.182371	1.2239652	84.74255	85.45745					1085831	4/2/2013 2:47
29	369 High School or Greater	2000	9 Total	RE 10	10	San Joaquin Valley	San Joaquin Valley	10	1286934.2	198895	10	67.87061									883771	4/2/2013 2:47
30	369 High School or Greater	2005-2007	9 Total	RE 10	10	San Joaquin Valley	San Joaquin Valley	10	1533344.6	218442	10	71.82204	0.197299	0.7008893	71.42533	72.20874					8937776	4/2/2013 2:47
31	369 High School or Greater	2008-2010	9 Total	RE 10	10	San Joaquin Valley	San Joaquin Valley	10	1657077.4	2185884	10	71.98449	0.179286	0.6399523	71.63389	72.3359					883308	4/2/2013 2:47
32	369 High School or Greater	2000	9 Total	RE 11	11	San Luis Obispo	San Luis Obispo	11	136271.78	159196	11	85.6									1114563	4/2/2013 2:47
33	369 High School or Greater	2005-2007	9 Total	RE 11	11	San Luis Obispo	San Luis Obispo	11	148795.25	170006	11	87.5	0.486322	3.8905775	86.54681	88.45319					109375	4/2/2013 2:47
34	369 High School or Greater	2008-2010	9 Total	RE 11	11	San Luis Obispo	San Luis Obispo	11	156708.45	177272	11	88.4	0.425532	3.6883786	87.58596	89.23484					1096774	4/2/2013 2:47
35	369 High School or Greater	2000	9 Total	RE 12	12	Santa Barbara	Santa Barbara	12	195409.37	246729	12	79.2									10305	4/2/2013 2:47
36	369 High School or Greater	2005-2007	9 Total	RE 12	12	Santa Barbara	Santa Barbara	12	200436.08	248880	12	80.6	0.425532	2.1934635	79.78596	81.43484					10075	4/2/2013 2:47
37	369 High School or Greater	2008-2010	9 Total	RE 12	12	Santa Barbara	Santa Barbara	12	207114.52	259542	12	79.8	0.486322	2.4075356	78.84681	80.75319					9390074	4/2/2013 2:47
38	369 High School or Greater	2000	9 Total	RE 13	13	Shasta	Shasta	13	89357.576	107272	13	83.3									1084635	4/2/2013 2:47
39	369 High School or Greater	2005-2007	9 Total	RE 13	13	Shasta	Shasta	13	105259.29	120022	13	87.7	0.486322	3.9538389	86.74681	88.65319					109625	4/2/2013 2:47



The indicator data files can be downloaded from the HCI website and are also available on the [CHHS Open Data Portal](#) (search keyword: HCI). Subsets of data (state, county and city/town level data) for some of the indicators are also accessible via online data visualizations created using Tableau software, and interactive ESRI Story Maps. These online tools provide immediate access to the data without the need to download the Excel data files.

Most Tableau visualizations include:

- California dashboard: provides statewide data for the indicator. When available, it includes time series data, race/ethnicity and other demographic detail.
- County dashboard: provides county level data for the indicator. When available, it includes time series data, race/ethnicity and other demographic detail.
- Below county dashboard: provides a snapshot of city/town or school district level data for the indicator. When available, it includes time series data, race/ethnicity and other demographic detail.

Most ESRI Story maps include:

- Choropleth maps with the most recent data available for the indicator for counties, cities/towns (or school districts), and census tracts (when available).
- In the maps the colors correspond to a ratio that indicates how many times the local percent is higher or lower than the California average.

## Interpreting Results

This section provides an overview of the issues that analysts should consider before accepting the results of an indicator. Numerous technical issues that affect indicators are beyond the scope of this manual. Many of these technical issues have been highlighted in the limitations section of the .pdf files that accompany the HCI indicators posted at the CDPH HCI website. Users are encouraged to consult the *General Reading and Resources* at the end of this section and seek out experts in subject matter and analytic methods.

### How Do Validity, Precision, and Bias Affect Indicators?

Indicators should be a valid measurement of what they purport to measure, precise and free from biases. Bias is defined as the systematic difference between a measured value and its true value. The proverbial “thumb on the scale” or a poorly calibrated gas pump that always comes up short of the gallon mark are examples of systematic bias that overstate or understate the true value.

Bias is not the natural variation or random fluctuation that we might expect from taking a random sample of the population rather than measuring everyone or every household. For many indicators, it is impossible to measure everyone in the population, so we take random samples that are meant to be representative of the population. If we could take numerous samples from the same population, the results of an indicator would vary somewhat from sample to sample. The amount of variability depends on the sample

size. We have more confidence in results when the sample is large and based on thousands of observations, rather than when the sample has only a few observations, for example, less than 10. Samples with fewer observations, which commonly occur in small geographies such as a census tract, will be more variable and less reliable. This pattern does not mean that the results are biased. If we could take many small samples, their average would be the same as the true value, and the differences between the average and the samples would follow a random rather than a systematic pattern. HCI files have statistical information for advanced users (standard errors, 95 percent confidence intervals, relative standard error) to describe this variability.

Biases can occur in each component of “what happened, to whom, where, and when.”

### To Whom (Denominator)

Some indicators include individuals or households in the population who are not capable of experiencing the event in the numerator. These individuals should be excluded from the denominator. For example, prisoners, felons, and undocumented residents may not be eligible to vote and should be excluded from the denominator of the voter participation indicator. In some geographies, people living in group quarters or institutions (colleges, nursing homes, prisons, etc.) may have a disproportional impact on an indicator. For example, a high concentration of college students, who are still dependents or are not employed while studying, can skew poverty statistics and are generally excluded from the poverty indicator. Likewise, some people or households may be eligible for inclusion in the denominator but were undercounted in the sample. For example, undocumented immigrants may have been undercounted by the Census. Some transportation indicators have a denominator of residential population, but a numerator (road traffic injuries or miles traveled) that may include both residents and nonresidents. This mismatch may be particularly pronounced at small geographies but relatively unimportant at the regional or state level. Data suppression is a formal decision by some data-gathering organizations, like the U.S. Census, to protect confidentiality, especially at small geographies (census tracts or blocks). This factor sometimes creates a reporting bias by race/ethnicity that favors more information on racially or ethnically segregated areas.

### What Happened (Numerator)

Biases may occur if the events in the numerator are undercounted or over-counted. For example, it is known that fatal motor vehicle injuries reported by police fall short of the count based on death certificates. Some underreporting is due to delayed deaths that occur more than 30 days after the traffic collision was initially reported. Although undercounting is a concern, as long as the level of undercounting remains constant over time, it still may be possible to observe valid trends over time within the same jurisdiction.

### When (Time)

Many indicators are based on a sample of the population at one point in time. When an indicator is compared over two time periods for the same jurisdiction, few people or households are likely to have been in both samples. In fact, the population in the geographical area may have experienced significant in- or out-migration due to community succession, displacement, or gentrification. In these situations, changes in the indicator may not reflect a change in living conditions in the same residents but rather a change in the residential population.

Definitions of indicators or the sampling methods may change over time, which may produce changes in indicators that are not real. Data sources often publish tables on data comparability and make recommendations for the series of years that have comparable data.

It is ideal to compare indicators in non-overlapping time periods. However, some indicators may be based on moving averages of several years (e.g., 2005–2007, 2006–2008, 2007–2009). Independent samples from mutually exclusive time periods reduce a “dilution” bias.

Changes in an indicator may occur at the same time as other changes in the community environment. Some changes in indicators may be part of long-term trends, and programs implemented on a backdrop of progressively improving (or declining) performance may be taking credit (or blame) for inevitable changes.

### **Causation vs. Correlation**

Several aspects of the community environment, including other indicators, may appear to have an influence on each other. One explanation is that two indicators tracking together is a chance event. Someone might observe that two indicators—for example, access to public transit and drinking water quality—appear to be related or have a statistical correlation. Transit access tends to be lower in rural areas, and rural areas have a higher prevalence of drinking water problems. But this occurrence does not mean that low public transit access is a cause of poor drinking water quality. There is no evidence that bus and rail traffic influences the presence of coliform, nitrates, arsenic, or other major contaminants found in drinking water. Yet, it is possible for these two factors to be statistically related but not causally linked.

Variation in indicator results by geography or population subgroups may appear to be influenced by several competing factors. It is well documented that voter participation is strongly influenced by age, with older persons voting at higher rates than younger persons. We may observe that voter participation rates are higher in some census tracts than others. Could the census tracts with higher voter participation rates also have a higher proportion of older residents? This may be a plausible explanation for the variation of voter participation by census tract. This possibility could be confirmed or eliminated if we could compare voter participation rates by age group in each census tract. This type of indirect causal relationship between three variables—census tract, age, and voter participation—is called confounding and could explain some results.

## Comparisons

As presented earlier in this manual, the results of an indicator require context for understanding and interpretation. How to regard the results often depends on the answer to the question, "Compared with what?" Several options for reference values include comparisons to state, regional, or county average; your position or rank in the entire distribution of results over all census tracts, zip codes, counties, cities, or other geography; and self-set goals based on historical trends, best observed values, peer jurisdiction average, or an aspirational goal. The difference between your observed performance and your goals is what drives the next steps.

## Interpretation Framework

Observing differences between performance and comparison values allows us to evaluate how close we are to achieving the aspirational goals in the [Healthy Community Framework](#). The following sequence of questions can help guide follow-up activities:

- Are the differences between your performance measurement and goal(s) real?
- Does the size of the differences have any practical importance?
- Are there differences in population subgroups that are avoidable and unfair?
- What are the causes of the differences?
- What are potential solutions to close the performance gap?
- What can we do as a community, stakeholders, policy makers, and individuals to improve outcomes?

The first question raises technical issues that were discussed in the **Validity, Precision, and Bias** section. One additional technical issue is whether differences between geographies, population subgroups, or reference values are due to random variability in sampling. The statistical methods to answer this question are outside the scope of this manual<sup>\*\*</sup>, and readers should consult a subject matter expert and the references at the end of this section, including the [HCI Technical Manual](#).<sup>2</sup> If the potential biases and confounding do not appear to play an important role and one can rule out chance variation due to sampling error and random yearly fluctuation, then it is likely that the performance gap is real.

The size of the differences that matters can be informed by some indicator and comparison data. For example, the number of people affected is given by the numerator of the indicator, and this information is available in the HCI data files. Other factors such as the severity of the outcome, its economic impact, and availability, feasibility, and cost

---

<sup>\*\*</sup> As an approximation, when the difference between the local result and the state average divided by the standard error of the local result is greater than 1.96, the difference is beyond what we would expect by sampling or random error ( $\frac{abs(\%local - \%CA)}{se_{local}} > 1.96$  at  $p < 0.05$ ). Statisticians call this "statistical significance," but this term also applies to differences that may be small and not of practical importance.

of solutions are other considerations. Those affected, policy makers, and stakeholders in the community should engage each other on this question.

Some differences in indicator values may be due to factors that are unavoidable such as aging of the population. However, many differences may have complex roots in history and the social determinants of health in which population subgroups have experienced discrimination or unequal treatment. Again, it is up to the people who are affected, policy makers, and stakeholders in the community to engage on this topic.

Indicators describe “what happened, to whom, when, and where,” but not “why.” Other means are needed to explain “why.” Scientific inquiry that has a formal process to weigh evidence is one path to explain the causes of the differences. Personal and community experience are another path to discover the causes of differences in performance. Each of these approaches is complementary and is part of a community dialogue to identify strategies to improve outcomes. Experiences of communities that have been successful in implementing improvements or achieving high levels of performance may provide examples of best practices. A systematic evaluation of scientific research can also play a role in identifying strategies that may be beneficial. Stakeholders should take advantage of the information in the indicators to inform actions that will make their communities a healthier place to live, learn, work, and play.

## Use Cases

### The Social Determinants of Health and their Link to HCI

Chronic illness and injury account for more than 80 percent of deaths in California, and many Californians have multiple chronic conditions that lower their quality of life and disproportionately contribute to California's annual \$230 billion health-care spending.<sup>3</sup> Major risk factors for obesity, diabetes, and other chronic illness and injury include poor nutrition, sedentary lifestyles, smoking, and alcohol use. Each of these risk behaviors is profoundly influenced by our social, physical, and economic environments, which are collectively called the “social determinants of health.” Chronic illness and injury are also the key drivers of significant and persistent inequities in health outcomes, including substantial differences in life expectancy by income level and race/ethnicity (Box 1).

#### The Social Determinants of Health: How Social Factors Can Affect Health

“Where we live, learn, work, and play has a greater impact on how long and how well we live than medical care. And although health care is critical, we must broaden our view and find ways to enable more people to lead healthy lives and avoid getting sick in the first place.

**Communities.** Health and health-related behaviors have been linked with a range of neighborhood features, including the concentration of poverty; the density of convenience stores, liquor stores, and fast-food restaurants relative to grocery stores selling fresh foods; access to transportation; the condition of buildings; and the presence of sidewalks and places to play or exercise.

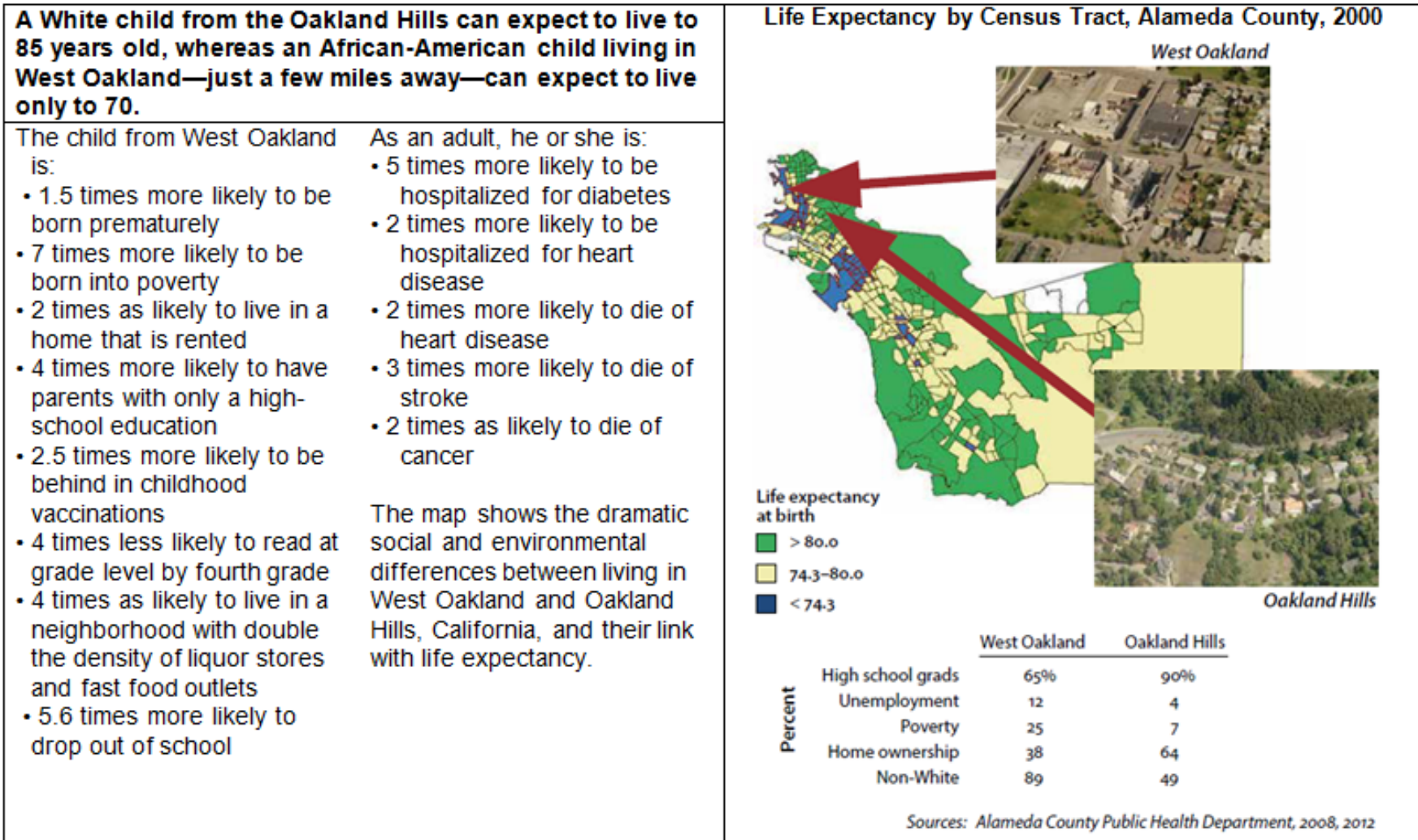
**Income.** A family’s income affects the health of both parents and children. More income increases access to nutritious food and other health-promoting goods and services, and can reduce stress by making it easier to cope with daily challenges. More income can buy the ability to live in a safe neighborhood with good public schools or to send children to private schools. These conditions can affect a child’s ultimate educational attainment, which in turn shapes job prospects and thus income levels in adulthood.

**Education.** Higher educational attainment can increase people’s knowledge, problem-solving, and coping skills, enabling them to make healthier choices. Education may also have powerful health effects by determining job prospects and thus earning potential. And education may also influence health through psychosocial pathways, by shaping people’s social networks and perceptions of their own social status.

**Racial or ethnic group.** Racial or ethnic differences in health can be explained in part by socioeconomic disadvantages that are the persistent legacy of discrimination. Chronic stress related to experiences of racial bias may also contribute to ill health—even without overt incidents of discrimination, and even among affluent and highly educated people of color.

**Childhood experiences.** Scientific advances have revealed that childhood experiences are particularly critical in shaping people's lifelong chances for good health. A range of biologic mechanisms, including responses to stress, are likely involved. Recent evidence indicates that physical and social conditions can influence whether particular genes are expressed or suppressed, making it clear that our genes alone do not determine our destiny" (Commission to Build a Healthier America<sup>4</sup>).

## Box 1: Place Matters: Community in Focus, Alameda County





## Examples of HCI in Action

Healthy Communities Data and Indicators can be used to:

- Describe the baseline community environment in which planning, projects, policies, and programs occur.
- Track changes over time to measure the impact of an intervention or program.
- Identify vulnerable populations, communities of concern, and “hot spots”.
- Describe how the social determinants of health vary by neighborhood, city, and county, and understand their health implications.
- Identify communities to prioritize for health interventions based on need.
- Provide feedback on whether policies, projects, and programs are proceeding as planned.
- Identify high-performance communities whose practices might help establish benchmarks and highlight “best practices”.
- Provide data for grants and to meet reporting requirements.

The examples on the following pages illustrate current and planned uses of the Healthy Communities Data and Indicators in different communities across California.

Examples of Uses of Healthy Communities Data and Indicators:

### Urban Planning

"Cities across California are using their **General Plan update** process to respond to the obesity epidemic. Some cities are including a separate health element in their general plan; others are adding health goals and policies in various general plan elements. Approximately 30 cities in California have added health goals and policies to their general plans. In Chula Vista, we are in the process of amending our General Plan and other planning documents to better emphasize access to healthy foods, walkability, pedestrian and bicycle safety, and a jobs-housing balance. The *Healthy Communities Data and Indicators Project* provides us with easy access to metrics on these type of topics and will help us help monitor our progress as we move forward." Brendan Reed, City of Chula Vista.

### Community Health Needs Assessment

"Napa County community members understand that improving the health of individuals, families, and communities requires a comprehensive understanding of health, one that considers all of the conditions in which people are born, grow, live, work, and age, including the health system. The Live Healthy Napa County collaborative is a public-private partnership bringing together representatives from health and health-care organizations, business, public safety, education, government, and the general public to develop a shared understanding and vision of a healthier Napa County. Our *Community Health Assessment* report presents data that reflect an understanding that health

extends beyond the medical setting. Thus, to improve health and well-being, the community strategies must consider the social, economic, behavioral, and structural factors that impact health. We were able to use the data files from the *Healthy Communities Data and Indicators Project* to build maps, graphs, and tables that appeared in our report." Dr. Jennifer Henn, Napa County Public Health.

### **Public Information Requests**

"Our public health department receives scores of requests from the public for data on the county. Some requests are from community groups who want to know about health, demographic, and socioeconomic issues in their communities. Other requests are from organizations that are applying for grants and need background information on community health. At times, these requests can be challenging for our staff, and we need a tool that the public can use to get this information on demand. We see the *Healthy Communities Data and Indicators Project* as a tool to help us meet the demand for public information about the health conditions in our communities. If there were such a tool, we would be able to refer many of our inquirers and know they would be getting the information they desire, allowing our staff to prioritize and follow up the most challenging requests." Matt Beyers, Epidemiologist, Alameda County Public Health Department.

### **Regional Collaboration on Improving Health and Reducing Health Disparities**

"The Public Health Alliance of Southern California is a collaboration of leaders from eight Southern California local health departments, who are responsible for the health of more than half of California's population. The Alliance addresses regional chronic disease prevention and reducing health disparities through policy, systems, and environmental change. Our current priority initiatives include multidisciplinary collaborations in transportation, food environments, and healthy community data indicators. In the transportation sector, we aim to advance health through improved transportation planning, using a coordinated regional approach that includes health metrics, performance-based measures, and cross sector data collection. Several indicators in the *Healthy Communities Data and Indicators Project* help us monitor active transport (walking and bicycling), transit access, and the targets in our regional transportation plan." Tracy Delaney, Executive Director, Public Health Alliance of Southern California.

### **Informing Decision-Making in Local Government**

"Our city council considers hundreds of issues of community interest throughout the year. Many times during our discussions and deliberations, we would like to bring up data and maps of neighborhoods on our mobile devices. This capability would add a dimension to our planning and decision-making process, and enable us to better consider how our decisions affect community health and contribute to a more efficient city government." Bay Area Elected Official.

**Accreditation**

"Like many jurisdictions, Orange County is actively pursuing accreditation for our local health department. A key part of accreditation is the creation of a Community Health Assessment, which analyzes population health status and community public health issues to drive programs and interventions that are of high quality and performance, advance health equity, and efficiently use limited resources. We have expanded our Community Health Assessment to include an assortment of emerging indicators on the built environment and the socioeconomic conditions affecting health to reflect both best practices and our values as an agency. The *Healthy Communities Data and Indicators Project* provides a substantial number of these cutting-edge indicators in a manner that is reliable, comparable, and valid, which supports our analysis of statistical trends and disparities among our diverse communities. The project is a crucial resource as the Orange County Health Care Agency and our allied community partners seek to make informed choices in the planning and implementation of those comprehensive, multi-component public health approaches that are key to accreditation and vital to addressing the most important causes of preventable disease." Travers Ichinose, Epidemiologist, Orange County Health Care Agency.

## References

1. Rudolph L, Sisson A, Caplan J, et al. *Health in All Policies Task Force Report to the Strategic Growth Council*. Sacramento, CA: California Department of Public Health; 2010. [www.cdph.ca.gov/programs/CCDPHP/Documents/HiAP\\_Final\\_Report\\_12%203%2010.pdf](http://www.cdph.ca.gov/programs/CCDPHP/Documents/HiAP_Final_Report_12%203%2010.pdf).
2. Bustamante-Zamora DM, Maizlish N, Tran D. *Healthy Communities Data and Indicators Project: Technical Manual*. Richmond, CA: Office of Health Equity, California Department of Public Health; 2014. [www.cdph.ca.gov/programs/pages/healthycommunityindicators.aspx](http://www.cdph.ca.gov/programs/pages/healthycommunityindicators.aspx).
3. Peck C, Logan J, Maizlish N, Van Court J. *The Burden of Chronic Disease and Injury in California, 2013*. Sacramento, CA: California Department of Public Health; 2013. [www.cdph.ca.gov/programs/Documents/BurdenReportOnline\\_04-04-13.pdf](http://www.cdph.ca.gov/programs/Documents/BurdenReportOnline_04-04-13.pdf).
4. Braveman P, Egerter S. *Overcoming Obstacles to Health in 2013 and Beyond*. Princeton, NJ: Robert Wood Foundation Commission to Build a Healthier America; 2013. [www.rwjf.org/en/research-publications/find-rwjf-research/2013/06/overcoming-obstacles-to-health-in-2013-and-beyond.html](http://www.rwjf.org/en/research-publications/find-rwjf-research/2013/06/overcoming-obstacles-to-health-in-2013-and-beyond.html).

## General Reading and Resources

### Policy, Planning, and Health

Stair P, Wooten H, Raimi M. *How to Create and Implement Healthy General Plans*. Oakland, CA and Berkeley, CA: Changelab Solutions and Raimi Associates; 2012. [changelabsolutions.org/sites/default/files/Healthy\\_General\\_Plans\\_Toolkit\\_Updated\\_20120517\\_0.pdf](http://changelabsolutions.org/sites/default/files/Healthy_General_Plans_Toolkit_Updated_20120517_0.pdf)

Rudolph L, Caplan J, Ben-Moshe K, Dillon L. *Health in All Policies: A Guide for State and Local Governments*. Washington, DC and Oakland, CA: American Public Health Association and Public Health Institute; 2013. <http://www.apha.org/~media/files/pdf/fact%20sheets/healthinallpoliciesguide169pages.ash>

University of Wisconsin Population Health Institute. *County Health Rankings and Road Maps. What Works for Health*. Princeton, NJ: Robert Wood Johnson Foundation; 2014. [www.countyhealthrankings.org/roadmaps/what-works-for-health](http://www.countyhealthrankings.org/roadmaps/what-works-for-health).

### Epidemiology and Biostatistics

Bonita R, Beaglehole R, Kjellström T. *Basic Epidemiology*. Second Edition. Geneva: World Health Organization, 2006. [whqlibdoc.who.int/publications/2006/9241547073\\_eng.pdf](http://whqlibdoc.who.int/publications/2006/9241547073_eng.pdf).

### Health and Related Data

- American Community Survey: socioeconomic, housing, and transportation data  
[www.census.gov/acs/www/](http://www.census.gov/acs/www/)  
[www.census.gov/acs/www/guidance\\_for\\_data\\_users/comparing\\_data/](http://www.census.gov/acs/www/guidance_for_data_users/comparing_data/)  
[www.census.gov/acs/www/Downloads/data\\_documentation/Statistical\\_Testing/2010StatisticalTesting3and5year.pdf](http://www.census.gov/acs/www/Downloads/data_documentation/Statistical_Testing/2010StatisticalTesting3and5year.pdf)
- California Health Interview Survey: health behaviors  
[healthpolicy.ucla.edu/chis/Pages/default.aspx](http://healthpolicy.ucla.edu/chis/Pages/default.aspx)
- Office of Statewide Health and Planning Development: hospital discharge/emergency room data  
[www.oshpd.ca.gov/](http://www.oshpd.ca.gov/)
- California Department of Public Health Vital Statistics: birth and death records  
[www.apps.cdph.ca.gov/vsq/](http://www.apps.cdph.ca.gov/vsq/)