HEALTHY PLACES INDEX GUIDANCE

FOR PUBLIC HEALTH SURVEILLANCE AND MONITORING HEALTH OUTCOMES, SERVICES, AND CLINICAL QUALITY IN HEALTHCARE SYSTEMS

BY NEIL MAIZLISH, PHD, MPH

JUNE 2023

CALIFORNIA Healthy Places Index™

Public Health Alliance™ OF SOUTHERN CALIFORNIA

Public Health Alliance™
OF SOUTHERN CALIFORNIA

WHO WE ARE

The Public Health Alliance powers collaborations between Local Health Departments and communities for transformational systems change.

Together, we represent 60% of California's population.

OUR VISION

Vibrant and activated communities achieving health, justice and opportunities for all.

OUR MISSION

Mobilize the transformative power of local public health for enduring health equity.

OUR CORE VALUES

- Courage
- High Impact
- Dynamic
- Ethical
- Unifying
- Inclusive

PUBLIC HEALTH ALLIANCE OF SOUTHERN CALIFORNIA
Public Health Institute, 555 12th Street, Suite #290, Oakland, CA 94607
https://www.thepublichealthalliance.org/

TABLE OF CONTENTS

PURPOSE

What is the Purpose of the Guidance?

This guidance supports health departments and healthcare organizations incorporate the Healthy Places Index (HPI) in health equity metrics for public health surveillance, population health management, and clinical quality improvement. Using indicators of socioeconomic status, education, housing, transportation, environmental pollution, built environment, and health care access, the HPI summarizes the healthiness of community conditions in patients' neighborhoods and addresses the social determinants of health as a driver of health outcomes. This is a complement to individual patient social needs and service referrals.

Who is the Audience?

The target users of this guidance are a team composed of data analysts, information technologists, and clinicians. Program managers should be generally familiar with the guidance so they can assure that their staff has the appropriate skill sets and tools to collect, analyze, interpret, and communicate the findings.

What Topics Are Covered?

The guidance reviews standard definitions of health outcomes, measures of frequency of health events in the general and patient populations, assignment of HPI scores to individual patients, construction of health equity metrics, analysis strategies using stratification of HPI scores in combination with demographic characteristics (age, gender, race/ethnicity, language, etc.) and time series, interpretation of results, including a checklist of potential explanations for both positive and negative results. We suggest ways in which results can be linked to policy action. This guidance is not a primer in biostatistics, epidemiology, or health services research, but resources are provided to build capacity in these areas.

What is Our Goal?

We aim to elevate a community of practice of users of the Healthy Places Index. Our goal is to provide a solid conceptual basis for analyzing health data so users can focus on the results and the implications for community health improvement. For technical assistance, please contact us at AskHPI@thepublichealthalliance.org.

BACKGROUND

Since 2018, the Healthy Places index (HPI) has been used by hundreds of California organizations across multiple sectors to help direct more than $2 billion in funding to improve the health of California communities, particularly those with the least resources. When community members, health professionals, advocates, and policy makers discuss health disparities or health inequities, they are referring to the observation that some segments of our society have a greater frequency of ill health or poor health outcomes than others.

Research demonstrates that most of these differences are not merely biological but are socially determined, and are avoidable, unfair, unjust, and related to social position based on age, race/ethnicity, gender identity, sexual orientation, class, tribal affiliation, and rurality. Differences in health status without an explicit social context are called health disparities.[1] Health inequities reflect a historical and ongoing maldistribution of community resources that are the product of policies, systems, and environments (PSE) that are the fabric of our modern lives. This means that these differences are largely preventable and amenable to change by health-promoting PSE. Many if not most of these resources and PSE originate outside the health sector. Based on scientific evidence, the "social determinants of health" names and categorizes these resources. The HPI 3.0 organizes 23 social determinants of health present in nearly 8000 California census tracts (approximating neighborhoods) into 8 domains or policy action areas: economic, education, social, housing, transportation, neighborhood built-environment, air and drinking water quality, and healthcare access. A composite score reflecting cumulative community conditions is calculated for California census tracts, whose scores are ranked from the least- to the most-healthy community conditions.[2]

In the California policy environment, people living in the 25% most disadvantaged census tracts are often earmarked for increased investments. This supports our practice to group census tracts into quartiles of HPI scores with a discrete blue-green color scale to visualize the ranking on our HPI mapping platform. Dark blue and dark green represents the least and most healthy community conditions, respectively. This sets the stage for examining the differences in the frequency of health outcomes in the four different population groupings by HPI quartile.

We will share real-world examples of graduated differences or dose-response across HPI quartiles showing people living in the least healthy community conditions have the least favorable health outcomes. We will discuss other ways the HPI score can be categorized for analyses of health outcomes.

The purpose of this guidance is to support health analysts with best practices to describe, analyze, and interpret differences in the frequency of health outcomes in communities with different HPI scores. The practices blend standard metrics and methods long-used in public health, epidemiology, health services research, biostatistics, and clinical quality improvement. Data scientists working in public health agencies and those in healthcare systems are our primary audience. Because of the confidential nature of health information, these are the two groups that have access to individual patient information, and can construct metrics of frequency of health outcomes and link these to the Healthy Places Index.

DEFINING HEALTH OUTCOMES, POPULATIONS, AND FREQUENCY

HEALTH OUTCOMES

Health outcomes is a broad term for a spectrum health events from self- and community-affirmed healthiness through progressive stages of symptoms, illness, disability, injuries, and death. This includes health services to diagnose and treat illness in patients and to prevent and promote health in community members by health professionals, clinics, laboratories, hospitals, and other healthcare facilities. For the purpose of measuring health equity, we present two main contexts: the community burden of disease and clinical quality of care.

Assessing the community burden of disease is an essential function of local and state health departments. The health outcomes that form the core of the community burden of disease includes approximately 40 communicable diseases that by state law must be reported to local health jurisdictions by health professionals and facilities. Local health departments are also responsible for registering birth and deaths, whose certificates include diagnostic and clinical information such causes of death and gestational age and birth weight, respectively. Local health departments access data on hospital discharges and emergency department visits of their residents as reported by California's short stay hospitals and emergency departments. Local health departments combine these sources with many others to develop periodic community health status reports and community health needs assessments in conjunction with non-profit hospitals. These reports are used to prioritize public health action and collaborate with community partners to develop plans to improve the community's health. These health improvement plans often include improving access to healthcare and other services as well and identifying populations with high rates of disease and illness living in geographic areas with low levels of resources.

Health outcomes are classified in a standardized manner using diagnostic and other criteria. They often match categories established by the World Health Organization's International Classification of Diseases, 10th Revision.[3] The National Notifiable Disease Surveillance System of the Centers for Disease Control and Prevention provide case definitions for use by state and local health departments.[4] In the arena of clinical quality improvement, federal agencies such as the Centers for Medicare & Medicaid Services (CMS)[5] and the Agency for Healthcare Research and Quality (AHRQ),[6] and non-profit healthcare organizations such as the National Committee for Quality Assurance (NCQA/HEDIS)[7] have established guidance for health plans and hospitals on defining health outcomes. These definitions are based on information collected during patient encounters from electronic and paper medical records and ordering and billing data that use standardized code sets for diagnosis (ICD),[3] treatment (CPT),[8] laboratory (LOINC),[9] imaging, and pharmacy (NDC).[10]

Before carrying out any analyses and to inform a written plan for data analysis, the analysis team should conduct a literature review that includes the definitions of health outcomes and health services published by official public health (local, state, and federal), and healthcare agencies,

professional organizations, and academic health researchers.

The advantage of using established definitions of cases includes vetting by governmental and non-governmental agencies for validity and practicality, and making the strengths and weaknesses of definitions transparent. Standard definitions may also help control bias (covered in more detail later). Using standard definitions will potentially make normative data available to allow you to contextualize, benchmark, and compare your findings with others. Standard definitions will also enhance communication among health professionals who use standard definitions as part of a shared professional vocabulary. Standard definitions also have detailed written documentation (see references above). The resource page in California County Health Status Profiles, 2022 provides additional details.[11]

FREQUENCY OF HEALTH OUTCOMES

Just as health outcomes have standardized definitions, so do the measures of their frequency. There are two dimensions to frequency: 1) how often they occur in the population from which they arise, and 2) how often do they occur with respect to time. Depending on the context, "population" may be the general population or a subset in community burden of disease or a specific group of patients in a health plan with a specific diagnosis.

The basic format of metrics of frequency used in public health and healthcare quality is:

$$\textit{Rate or Percent} = \frac{\textit{Health Outcome (Numerator)}}{\textit{Population (Denominator)}} \textit{ at a specified time or time interval}$$

Definitions generally incorporate time in one of two measurement periods. If outcomes and the population are enumerated in one point in time, the frequency measure is called prevalence. If new health outcomes or cases are accumulated in a specific time interval (week, month, year, etc.) the frequency measure is called incidence. For practical reasons, the general population is often represented as an average or the midpoint date over the time period. In managed care

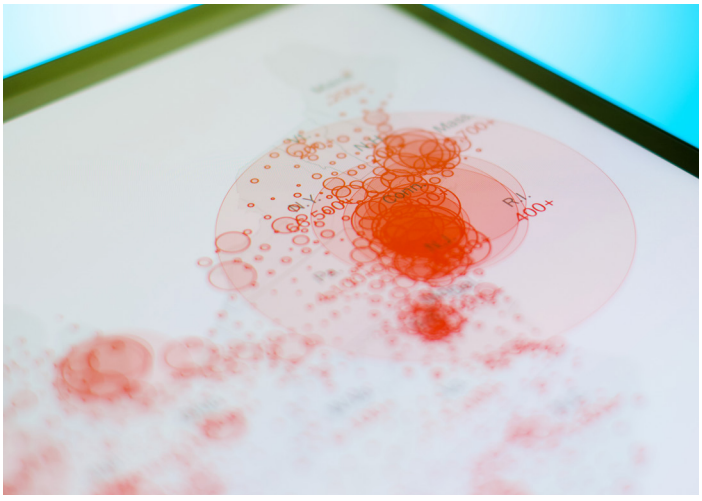
organizations the eligible population is based on uninterrupted months of enrollment usually within calendar years of group health contracts. Incidence is the most common frequency metric for assessing community burden of disease, healthcare utilization and healthcare quality. Deaths or mortality are a special kind of incidence. Population health surveys such as the California Health Interview Survey (CHIS)[12] and patient satisfaction surveys used to assess quality of healthcare (Consumer Assessment of Healthcare Providers and Systems, CAHPS)[13] generally measure prevalence.

In public health settings, the denominator reflects the community from which the burden of disease occurs and references the entire population, which includes both the individuals with the health outcome and individuals who are healthy. Depending on the health outcome, the population may not be the entire population but one that is restricted to a subgroup based on age, gender, or clinical criteria. For example, only women who are female at birth can experience ovarian or cervical cancers, so cancer incidence for these sites of cancers may include only females at birth, or more precisely females without hysterectomies. Some clinical care guidelines such as well baby visits, only apply to children below a certain age. Premature births are another example in which the population is restricted to all births or those with specific characteristics, such a singleton (non-twin, triplets, etc.). The references above delineate both the demographic and clinical criteria for inclusion or exclusion for the numerator and denominator of health outcome metrics.

Public health surveillance uses general population counts and subsets based on data from the U.S. Census Bureau (decennial census or American Community Survey, or Current Population Survey)[14] with refinements from state agencies (Department of Finance).[15] Managed care organizations can base populations on member-months of enrollment. Other health organizations such as preferred provider organizations (PPOs) networks and indemnity insurers may only be aware of persons covered by their policy that seek medical care and generate a medical bill (numerator), and may not have a business reason to compile data on all policy holders and their dependents (denominator).

In the literature review previously mentioned, review the definitions of the frequency metric for the health outcome of interest. Consider the use of standardized metrics of frequency published by official public health and healthcare agencies, professional organizations, and academic health researchers.

Like standardized health outcome definitions, using established definitions of frequency includes vetting for validity and practicality, making the strengths and weaknesses of definitions transparent, and helping control potential bias. Using standard definitions will potentially make normative data available to allow you to contextualize, benchmark, and compare your findings with others. Standard definitions will also enhance communication among health professionals who use standard definitions as part of a shared professional vocabulary. Standard definitions also have detailed written documentation (See references above).



Incidence and mortality metrics are a rate expressed as the number of events per population unit. When the number of health events is small, it is customary to multiply the numerator/denominator fraction by 1,000 or 10,000, or 100,000, or even per million, so that the metric is not overwhelmed by many 0s preceding the first non-zero digit. So, 20 events in a population of 200,000 people could be expressed as 0.0001 or 1 per 10,000 (or 10 per 100,000). For public health reporting, there are often specific reporting conventions regarding the population basis to communicate the frequency of health outcomes.

Analysts should refer to official organizations reporting conventions to express the rate or percentage of health events. Using rates and percentages has the obvious advantage when comparing populations of different sizes. Thus, for the analysis of health equity among subgroups, rates and percents are almost always used as health outcome metrics. While the sheer number of cases is important, it not by itself an indicator of health disparities between groups.

Percents are often a metric to describe self-reported health status or patient satisfaction when either a sample of the general population or a sample of patients is used in the denominator of the percentage. This type of percent should be distinguished from the percent distribution of a specific diagnosis as a proportion of all diagnoses or patients with a diagnosis (percent of all hospitalized patients with a cancer diagnosis). A percent specific to particular diagnosis is called **proportional morbidity or proportional mortality**. This metric has no measure of the underlying population, just "numerator" or patient data. These percents are subject to the limitation that the percents for each diagnosis must ultimately add up to 100% of diagnoses. This creates a mathematical see-saw: a large increase in cases of one diagnosis depresses the percentage for other diagnoses. This is likely to have happened during the COVID-19 pandemic. Underlying proportional morbidity analyses is the assumption that the overall rate of disease (all diagnoses combined) is the same between two different population subgroups. In health equity analysis, this assumption is almost never true.

HEALTH EQUITY METRICS AND THE HEALTHY PLACES INDEX

Health outcome rates and percents are a component of health equity metrics, which are based on subdividing populations along one or more dimensions of health equity. Both outcomes and populations—numerator and denominator of the rate or percent, respectively—are subdivided by criteria such as age, gender, race/ethnicity and HPI quartile. Subgrouping the overall rate by population subgroups is called stratification. It is an important technique to examine potential health inequities, but it is not a health equity metric. Health equity metrics are either the difference or the ratio of the rates of outcomes in two or more groups.[1] One of the groups must be designated as a reference.

$$\textit{Difference in Rates} = \textit{Rate of Health Outcome in Group}_1 - \textit{Rate of Health Outcome in Group}_2$$

$$\textit{Ratio of Rates, RR} = \frac{\textit{Rate of Health Outcome in Group}_1}{\textit{Rate of Health Outcome in Group}_2}$$

For example, in 2020, the fatal injury rate for California men was 96.5 per 100,000 males (18,896 deaths in a population of 19,577,489) and, for females the rate was 335 per 100,000 females (6,638 deaths in a population of 19,790,589). The rate ratio is 2.9 and the rate difference is 63 per 100,000.

There are two general approaches to designating a reference group (group 2 in the equations above). The first approach is done *a priori* (i.e., decided before data analysis) and is based on historical and current scientific evidence from a review of the medical and public health literature. The subgroup that was least impacted or that benefited from policies of social exclusion of other groups is designated as the referent (Group 2). When such a referent is not identified, it is customary to compare all subgroups to the group with the lowest rate (of an adverse outcome). Sometimes the overage average rate or percent is used as a benchmark. Rather than health inequity, differences without an explicit social context are better described as health disparities. That a

historically discriminated subgroup does not have the poorest health outcome may have several competing explanations which will be explored in the Interpretation section.

A common variant of the health equity metric, based on the difference in rates, is the attributable number or percentage of cases if the most disadvantaged group had the rate of the most advantaged. This is tantamount to saying if we could eliminate health inequities in the subgroup experiencing the poorest rates, how many cases would be avoided in this subgroup. Sometimes these cases are called excess cases, or attributable cases.

$$\text{Avoidable Cases in Group}_1 = \text{Rate Difference}_{\text{Group 1-2}} \times \text{Population (Denominator)}_{\text{Group 1}}$$

In the example above, avoidable cases in men is 63/100,000 × 19,577,489 or 12,321.

$$\text{Population Attributable Fraction,\%} = 100 \times (RR - 1)$$

Rate ratios are dimensionless and help identify subgroups that should be prioritized for follow-up, even when the group is numerically small compared to other population subgroups. Rate ratios are generally more easily communicated than rate differences, which must retain their units to be interpretable. A rate ratio of 2 can be communicated as Group 1 having a rate that is twice as high or 100% higher than the reference group. Using large numbers like 100% has a different emotional appeal than a smaller number (like 2), although the expressions are mathematically equivalent. It is perfectly plausible that the attributable cases in a numerically large subgroup with a low RR is greater than a small subgroup with a very large RR. Attributable cases depend on both the magnitude of the differences in rates AND the size of the population experiencing poorer rates.

HEALTH EQUITY METRICS INCORPORATING THE HEALTHY PLACES INDEX

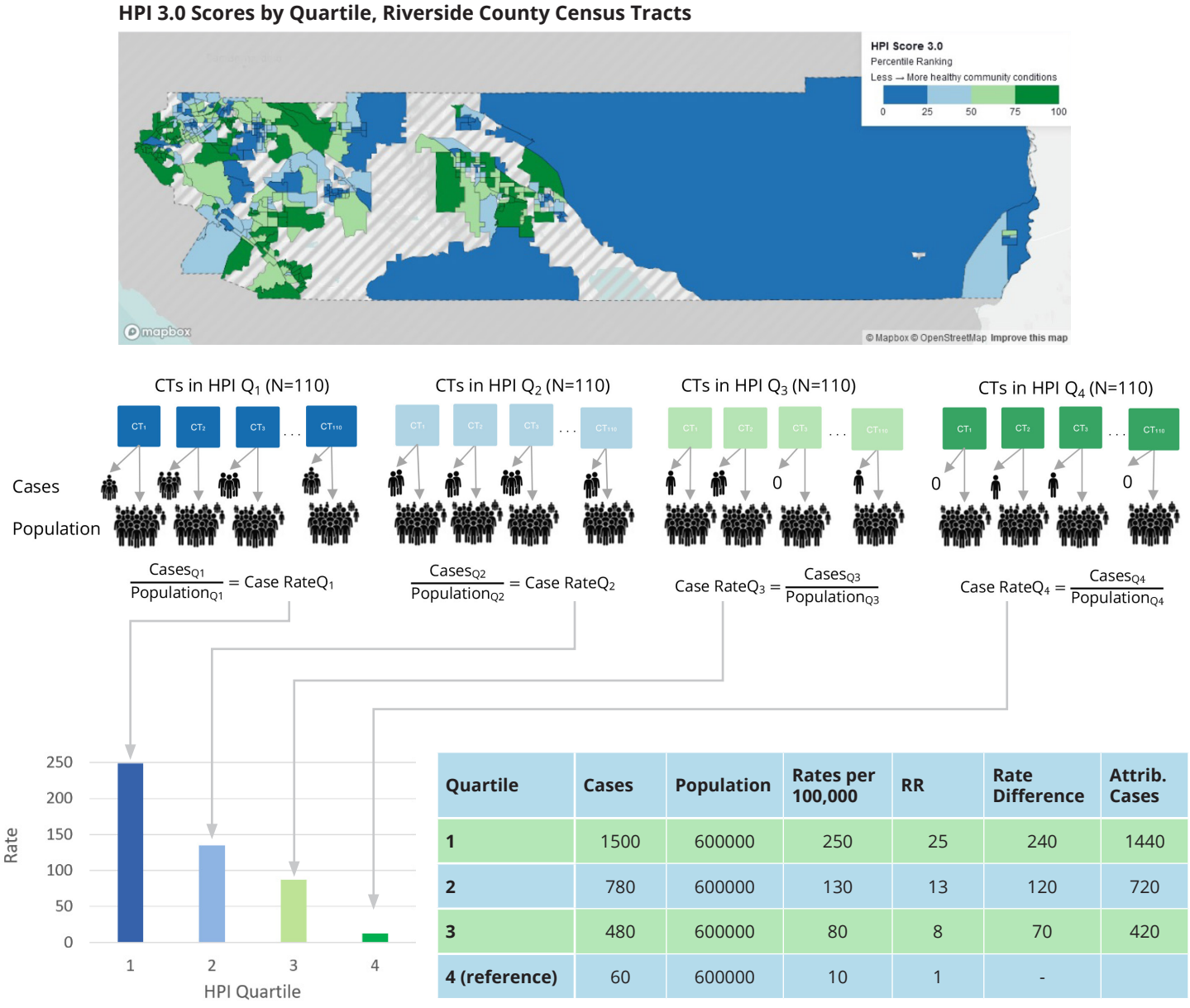
Constructing health equity metrics with the Healthy Places Index is a special instance of stratification that subdivides cases (numerator) and populations (denominator) by the HPI quartile of residential census tracts, or less ideally, ZIP Code Tabulation Areas (ZCTA). People living in the same neighborhood are often more like each other than people living in neighborhoods farther away. This is, in part, the legacy of residential racial discrimination that segregated neighborhoods and local zoning ordinances that disfavored mixed income and mixed race neighborhoods. Nonetheless, neighborhood average characteristics (such as median income, years of education, percent of homeowners, level of housing overcrowding, etc.) can be ascribed to both cases (numerator) and populations (denominator).[16]

Assigning a Patient and Population a HPI Score

The census tract of a patient can be ascertained by automated means through look-up lists, and geocoding, which uses mapping software to assign x and y coordinates of residences, and places them in census tract boundaries. Each tract has a HPI score which can be linked to the patient, and the entire census tract population can be assigned a HPI score. Cases from groups of census tracts with the same range of HPI scores (quartiles) can be aggregated together. And, likewise, population counts in groups of census tracts with the same range of HPI scores can be aggregated. This is illustrated in Figure 1 using hypothetical COVID-19 cases in "California" County.

Figure 1. Stratification of Census tract cases and population and aggregation to create population rates by HPI Quartile

- ~440 census tracts (CTs) in Riverside County with ~4000 people in each
- Census tracts from least to most healthy community conditions in 4 equal groups (~110 tracts and 600,00 people)



Quartile	Cases	Population	Rates per 100,000	RR	Rate Difference	Attrib. Cases
1	1500	600000	250	25	240	1440
2	780	600000	130	13	120	720
3	480	600000	80	8	70	420
4 (reference)	60	600000	10	1	-	

Compared to the people living in the HPI quartile of census tracts with the healthiest community conditions (Q4), other HPI population quartiles experienced progressively higher COVID-19 case rates and health inequity. The least resourced communities had rates 25 times greater than the most resourced communities. The rate ratios, RR, are measures of health inequities rather than health disparities, given a long line of evidence that social policies (e.g., redlining) have narrowed the resource base in the spheres of socio-economic development, education, housing, transportation, and environmental exposures.[17] There is a dose-response relationship from HPI quartile Q1 to Q4, which is an example of a "social gradient" in disease.

Zero Case Census Tracts

One practical consideration is ensuring all census tracts in the health jurisdiction or healthcare service area are included in the aggregation of cases and populations. Some enterprise data systems create line lists based only on census tracts reporting cases. Census tracts without cases (zero cases) that have an eligible population should be included in the aggregation the population denominator.

Census Tract vs. ZCTAS

We recommend using census tracts rather than ZCTAs as the geographic unit to calculate the HPI scores and rankings. Census tracts are smaller in area and population than ZCTAs, which can show substantial variation in disease rates and social determinants of health compared to census tracts. This loss in variability may attenuate the dose-response gradient. Census tracts rather than ZCTAs are the preferred geographic unit to score and rank the Healthy Places Index.

Local vs. Statewide HPI Rankings of Census tracts

The geographic extent of analyses will influence the relevance of using the statewide ranking of HPI scores versus a local or county ranking of HPI scores. Most local health jurisdictions prefer to re-rank the raw HPI score and calculate percentiles for only census tracts in their county. The rank geographies function of the HPI mapping application (https://map.healthyplacesindex.org/) creates re-ranked HPI percentiles, which can be downloaded from the map.

Matching of Time Period of Cases, Populations, and HPI Version

Rates composed of numerators (cases) and denominators (population) that are temporally matched are the most accurate. The HPI is updated at 3- to 5-year intervals and combine data that are largely 5-year annual averages. HPI 2.0 covers 2011-2015 and HPI 3.0 covers 2015-2019. Users should pick the HPI version that best matches the time period of the cases and population.

Longitudinal Comparisons Across HPI Versions

HPI scores and rankings are specific to their respective version, and should be compared with great caution. The differences across versions reflect changes in individual indicator values and life expectancy at birth, census tract eligibility, and domain weights. Change in individual indicator values appears to make the greatest contribution to census tract rankings.[18] Change in indicator values for the nearly 8000 California census tracts includes Z-scores for each indicator which has a different mean in each HPI version. True compatibility across time periods would use the same baseline mean for the calculation of Z-cores that make up the overall HPI index.

Geographic Coverage

Some use cases of the HPI require complete geographic coverage for California. However, the HPI is calculated for census tracts and ZCTAs with 2015-2019 populations of at least 1500 people and for census tracts with less than 50% group quarters. The rationale behind the exclusion criteria was to increase statistical reliability and validity. The 270 excluded tracts comprise 3% of all California census tracts and roughly 500,000 people (1.4% of California population). Rural areas with low resources are overrepresented in these excluded tracts.

If complete geographic coverage is required in a HPI use case, cases and populations of HPI-excluded census tracts cases and populations can

be assigned to the least healthy HPI quartile (Q1).[19] This is less labor intensive and less confusing than altering the HPI methodology or creating non-standard values for HPI scores and percentiles. Any derivative index based on a modified HPI methodology cannot be called the Healthy Places Index, which is the copyrighted intellectual property of the Public Health Alliance of Southern California.

Populations Experiencing Homelessness

Representing populations experiencing homelessness in the Healthy Places Index is challenging. Data on these populations are not reported at the census tract or ZIP code. Some people experiencing homelessness temporarily live in shelters with addresses that can be geocoded to census tract and others live unsheltered on streets or in tent encampments without official addresses. Public health and health care systems that provide services for people experiencing homelessness may record their address as "homeless", or give a shelter address, or the last known address. This may provide information for a numerator of a health equity metric, but the lack of an estimate of the population experiencing homelessness (denominator) makes it difficult to construct an equity metric with the Healthy Places Index, stratified by homelessness.

Including population experiencing homelessness in the least healthy HPI quartile (Q1) may create the least amount of misclassification. The resource level in populations experiencing homelessness is critically low and often are segregated to neighborhoods that are similarly low resourced. Excluding people experiencing homelessness from analyses may underestimate disease rates in the least healthy HPI quartile, Q1 (i.e., excluded from numerator but included in the general population, which includes the population experiencing homelessness).

Populations with Limited HPI Quartiles Representation

Medi-Cal and community clinic populations live in geographic areas that have a large overlap with HPI Q1 and little overlap if any with Q4. Equity analyses may not show a dose-response because there is little or no representation beyond Q1. It is inappropriate to declare that health inequities have been eliminated if the analysis is largely confined to a single HPI quartile.

Additional Populations Stratifiable by the Healthy Places Index

HPI population quartiles can be additionally stratified by other dimensions of health equity from data sources such as the decennial census and the American Community Survey. These include age, sex, race/ethnicity, language, region, and occupation. A significant caveat is that the stratification of the numerator and denominator must uses the same definitions and categories. For example, cases aged 0 to 17 years must be paired with the population aged 0 to 17 years. Race/ethnicity in the numerator and denominator must be defined using the same criteria, including whether data were based on self-identification or a judgment by staff in health systems without asking the patient. This may be challenging because public health and healthcare systems might use race/ethnicity categories that do not match those of the Census, or combine categories in different ways. For example, Asian and Pacific Islanders may be combined in medical records, but are distinct categories in the Census or visa versa. Some patients, particularly Latino, self-classify into "Other", creating a large data artifact, when compared to Census categories.

Best Practices for Race/Ethnicity

Mutually exclusive race/ethnicity classifications are but one way of classifying race/ethnicity of cases and populations. The classification used by the federal Office of Management and Budget (OMB) has had an outsized influence on the categories that have become *de facto* in public health surveillance, although this is beginning to change. The OMB classification is problematic because it is not accompanied by explicit instructions to disaggregate the major categories, which for Hispanic, Asian, and Native Hawaiian/Pacific Islander can obscure important heterogeneity in rates of health outcomes between subgroups within categories. Mutually exclusive categorization also tends to numerically minimize Native American/Alaskan Natives and Native Hawaiian/Pacific Islanders who may also identify with other race/ethnicity groups. For these groups, census designations "alone and in combination with other races" vs. "alone" may be more representative.[20] A larger population will also enhance the statistical reliability of the rate.

To maintain validity, the definition of the race/ethnicity category in the cases (numerator) should match the denominator (population). When possible, disaggregate the stratification to the maximum extent. In addition to mutually exclusive categorization, consider non-mutually exclusive classifications, especially for groups that are numerically small. Details of these categories and their locations throughout California are available at the HPI mapping applications (https://map.healthyplacesindex.org/).

Occurrence/Residence Mismatch

In public health surveillance, the populations used in rates and percentages are residential. To avoid geographical mismatch, health outcomes must also reflect the residence of the patient. Some data sources include both location of occurrence and residence of the patient. For example, for fatal traffic collisions death certificates record the location of the event and the residence of the decedent. However, in SWITRS,[21] a traffic

ANALYSIS STRATEGY

collision database maintained by the California Highway Patrol, only the location of the collision is recorded. The crash location may be in a census tract with a different HPI score than that of the decedent. Occupational injuries may also be reported by employer worksite, whose census tract HPI score may not correspond with the workers residence. The HPI may be not be suitable for assessing the health equity dimensions of health outcomes where health events are attributed to a service location rather than a patients' residence.

For health equity, the analytic strategy should focus on a series of questions:

- Do the data have sufficient integrity to even do an analysis?
- What is the size of the health equity gaps (difference or ratio of rates) between HPI quartiles?
- Are the gaps increasing or decreasing over time?
- Is there a "dose-response" relationship of the health outcome over HPI quartiles? Dose-response refers to increasing or decreasing rates (or rate ratio) of the outcome with increasing or decreasing HPI quartile.
- Does the dose-response relationship change in subgroups of a third variable such as age, gender, race/ethnicity?
- Is the rate of the outcome in the highest and lowest HPI quartile so far above or so far below a benchmark that it is not meaningful to differentiate among HPI quartiles when all quartiles are uniformly deficient or excellent?

DATA INTEGRITY

Data integrity refers to the validity, accuracy, and completeness of the data, particularly case data that are collected in enterprise data systems for public health and healthcare. Validity refers to what is purported to be measured in concept is actually measured. For example, if a population was defined as people with diabetes based on a sole diagnostic test with a qualifying ICD-10 code, many patients whose test was ordered to *rule-out* diabetes will be included in the population. This is a data artifact because commercial diagnostic laboratories reject (for billing purposes) a lab test request without a tentative diagnosis or

a diagnosis that does not correspond to the purpose of the test. The validity of a study that uses this (unrealistic criteria) would be highly suspect. At least a second confirmatory diagnosis with diabetes would make the definition of the population more plausible. Accuracy is related to precision of measurement. For example, rather than the age at last birthday, age might be rounded to the nearest 0 or 5 year, or age computed from dates may have incorrect or illogical dates (admitted to hospital before the birthdate). There are many kinds of data entry errors: digits that are deleted, added, incorrect, or reversed). As long as these types of errors are infrequent (<1% of a variable), the data are adequate for surveillance and monitoring purposes. Completeness of data refers to both complete ascertainment of all eligible cases as well as non-missing values of covariabes: census tract or geocodeable residential address to assign an HPI score, age, gender, race/ethnicity, etc.

While there are no hard-and-fast rules for the amount of missing data that seriously weakens the integrity of a data analysis, one can perform a sensitivity analysis by assigning a specific, but hypothetical value to missing data (e.g., all missing race/ethnicity assigned to Black race group). If the conclusions of the study would be the same with and without omitting the missing data, one could state this to support the findings. If the conclusions were different, then a statement saying that missing data cannot be excluded as a potential explanation of results would be appropriate.

Restricting the data analysis only to cases with complete data risks making the analysis

population unrepresentative of the actual population, introducing a potential bias, especially when missing data do not occur randomly across all population subgroups. We do not recommend imputing missing HPI scores. This violates our user agreement by creating a derivative product that cannot be named the "Healthy Places Index."

For routine public health surveillance and population health management in healthcare systems, we recommend simple, accessible statistical analyses based on categorizing HPI scores into quartiles (either statewide, county, or service areas specific). HPI scores can be used to construct a continuous variable in advanced regression techniques (e.g., logistic, Poisson, negative binomial). This format of the HPI requires staff with advanced training and is more suitable to research settings. Communicating the results of regression coefficients (e.g., increase in disease rates per unit change in HPI score or percentile) is challenging for a non-technical audience.

UNIVARIATE ANALYSES: THE RELATIONSHIP BETWEEN A HEALTH OUTCOME AND HPI

Assuming the data quality assessment signals data integrity, the starting place for analysis of health equity with the HPI is a univariate analysis. (Describing the demographics of the population and inter-relationship between demographic variables is also useful). The univariate analysis gets its name from the analysis of one health outcome with one ("uni") equity variable, starting with the HPI. A univariate analysis of a health outcome and categories of HPI quartiles at a single point in time or time period generates 4 data points, which can be tabulated, and graphed as bar or line chart. The following examples illustrate health outcomes in 1) public health (Table 1/Figure 2: Congenital syphilis surveillance), 2) healthcare (Table 2/Figure 3: discharge diagnoses of mental and behavioral conditions from hospitals and emergency departments), and 3) preventive clinical services (Figure 4: COVID-19 vaccination coverage). The analyses are statewide and use statewide ranking of HPI scores into quartiles and use the American Community Survey to define population denominators based on census tracts (syphilis) or ZCTAs (mental/behavioral health, vaccine coverage). These types of analyses help

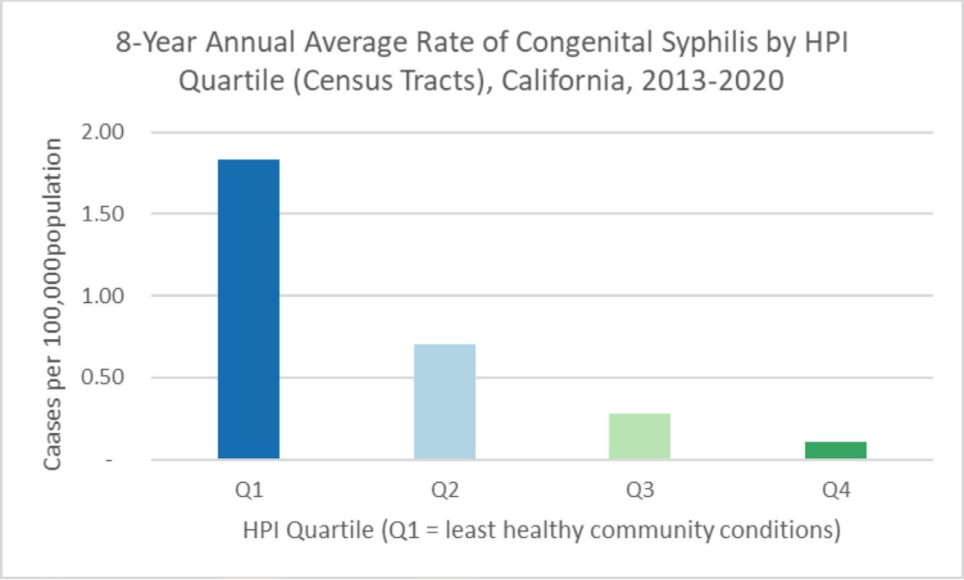
Table 1. Eight-Year Annual Average Rate of Congenital Syphilis by HPI 3.0 Quartile (Census Tracts), California, 2013-2020

HPI Quartile*	Annual Average Cases	Population	Rate per 100,000	Rate Ratio#
Q1	106	5,800,000	1.83	17
Q2	38	5,400,000	0.70	7
Q3	19	7,000,000	0.28	3
Q4	7	6,700,000	0.11	1 (reference)

* Least healthy community conditions
# Rate Ratio = Rate/Reference Rate
Note: Data in table have been rounded to simplify calculations

SOURCES: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California (excludes Los Angeles and San Francisco counties)

Figure 2. Eight-Year Annual Average Rate of Congenital Syphilis by HPI 2.0 Quartile (Census Tracts), California, 2013-2020



SOURCES: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California (excludes Los Angeles and San Francisco counties)

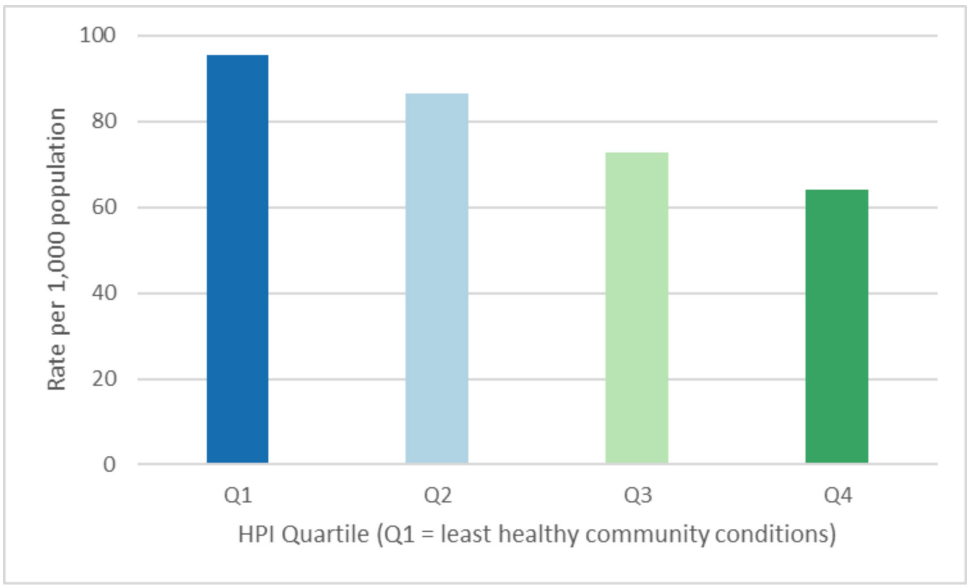
Table 2. Rate and Number of Behavioral/Mental Health Emergency Department and Hospital Discharges by HPI 3.0 Quartile (ZCTAs), California, 2020

HPI Quartile*	Discharges	Population	Rate per 1,000	Rate Ratio#
Q1	924,173	9,679,402	95	1.5
Q2	874,681	10,109,746	87	1.4
Q3	718,596	9,856,787	73	1.1
Q4	598,209	9,328,849	64	1.0 (reference)

* Least healthy community conditions
# Rate Ratio = Rate/Reference Rate

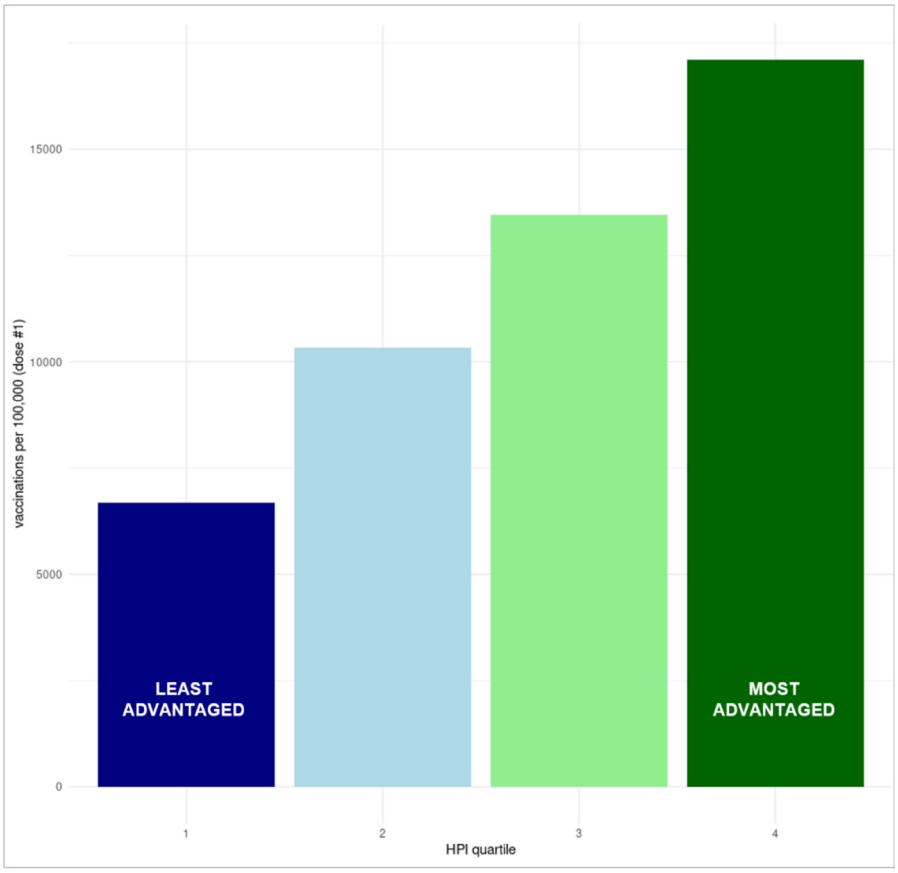
SOURCES: Department of Health Care Access and Information, American Community Survey, Public Health Alliance of Southern California

Figure 3. Rate of Behavioral/Mental Health Emergency Department and Hospital Discharges by HPI 3.0 Quartile (ZCTAs), California, 2020



Sources: Department of Health Care Access and Information, American Community Survey, Public Health Alliance of Southern California

Figure 4. Rate* of Vaccination Coverage by HPI 3.0 Quartile (ZCTAs), California, Feb. 2021



* Rate of first dose per 100,000 population (aged 16 and older)

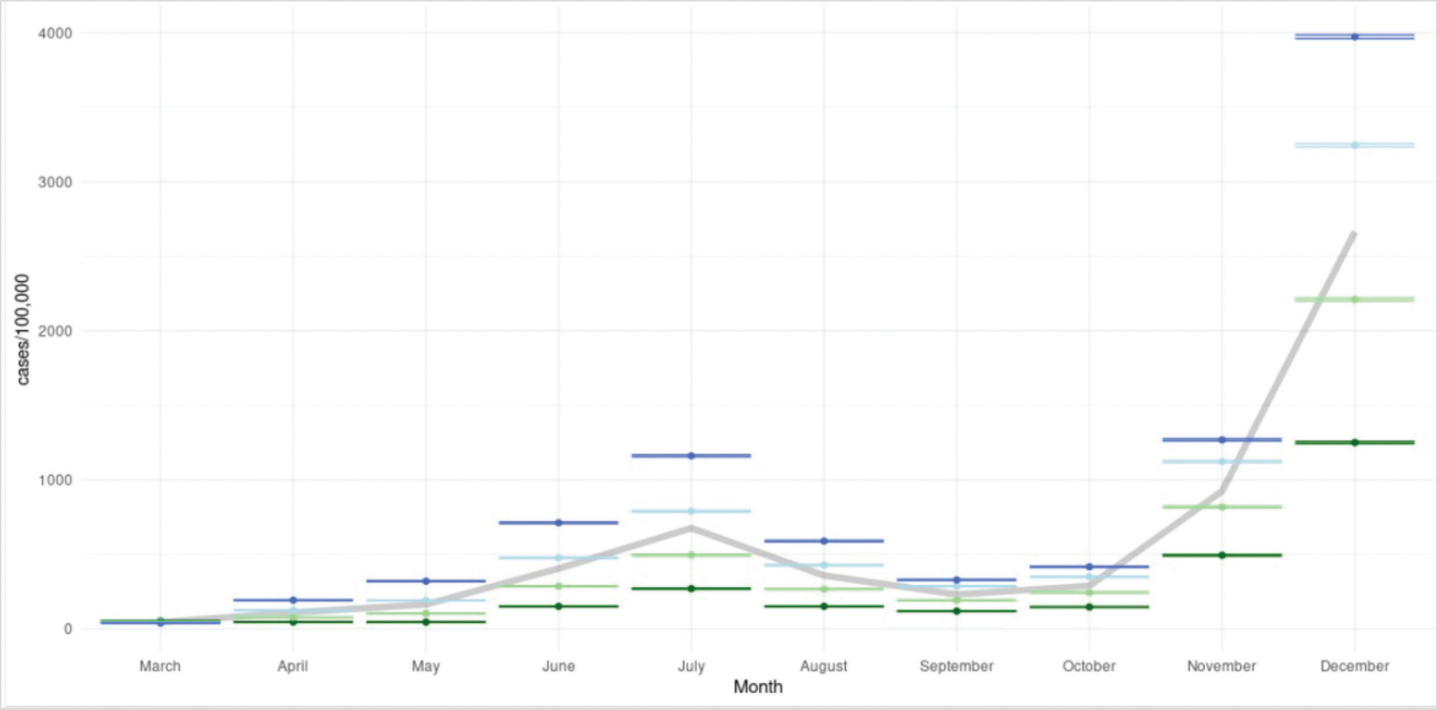
Sources: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California

address the second bullet point above, regarding the size of the differences between Q1 and the other HPI quartiles as measured by the rate ratio of Q1 to Q4. In the examples, we observe a dose-response. We might describe the congenital syphilis rate as being 17 times greater in Q1 than Q4; the rate of mental health discharges being 50% higher (RR=1.5) in Q1 compared to Q4; and the rate of vaccination coverage being 62% lower in Q1 compared to Q4, or that Q4 has 2.6 times the vaccination rate as Q1.

UNIVARIATE ANALYSES IN A TIME SERIES

Routine public health surveillance and clinical quality improvement are ongoing efforts in time to detect sudden rises in illness rates, monitor time trends in chronic disease and mortality, or as an evaluation or feedback tool to see if clinical or other actions in a plan-do-study-act (PSDA) cycle is changing the frequency of the health outcome or changing the size of the gaps in health equity variables. Time series require the calculation of rates by HPI quartile at time intervals. The time intervals depend on the incidence rates themselves: rare diseases require a longer time period to accumulate a meaningful number of numerator events than common diseases

Figure 5. COVID-19 Case Rates by HPI 3.0 Quartile and Month, California, March - December, 2020



Sources: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California
Note: Grey solid line is California average rate

or illnesses. For public health surveillance, the time periods are formalized by state health departments and the Centers for Disease Control. The key question in examining times series stratified by HPI quartile is "do the differences increase, decrease, or stay the same over time?" A corollary is "Are there trigger events (like a new strain of COVID-19) that appears to magnify health inequities (differences between the HPI quartiles) during pandemic surges? This is illustrated in Figure 5.

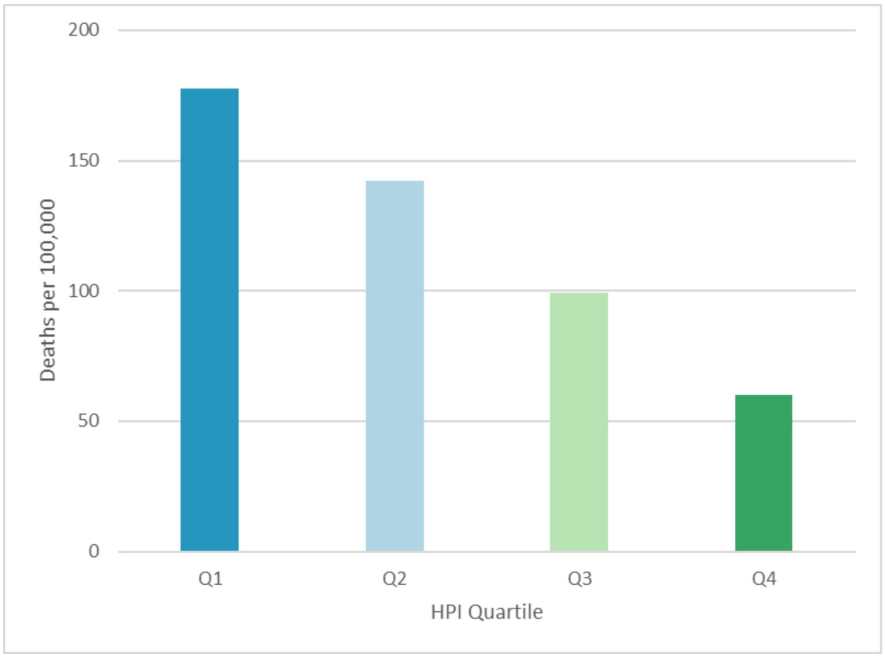
BIVARIATE ANALYSES

Bivariate analyses get their name from the analysis of one health outcome with two ("bi") stratification variables, one of which is the HPI quartile. The second stratification variable may be categories of age, gender, race/ethnicity, region, language, etc.

Many health data systems routinely include patient race and ethnicity, which is an important focus of health equity. With the inclusion of the HPI, it is possible to examine the independent and combined impacts of "race" and "place". This is an essential bivariate analysis.

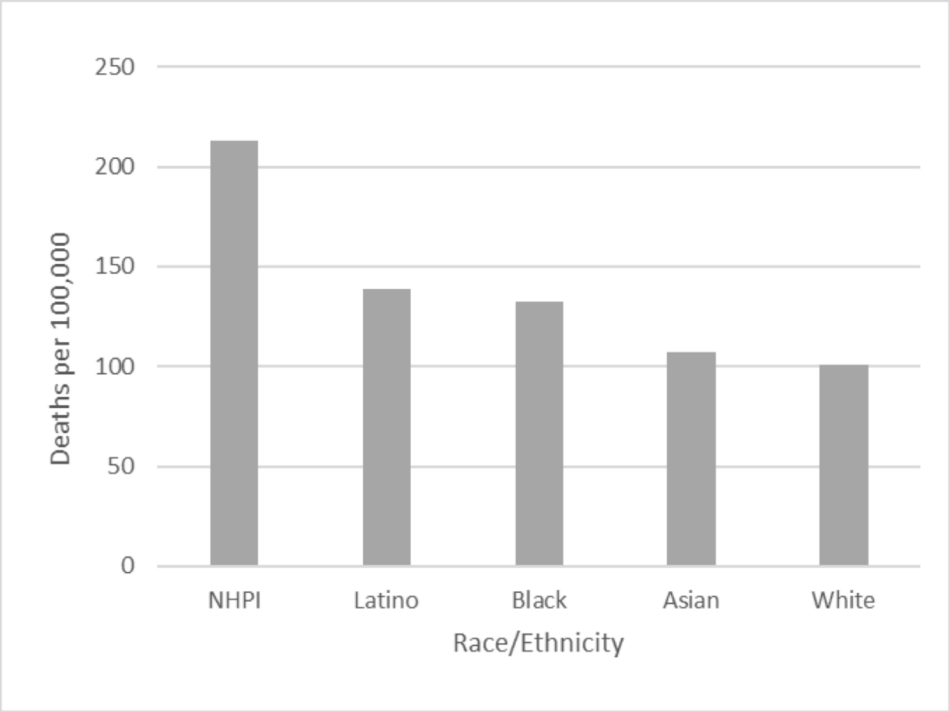
We illustrate this with California COVID-19 mortality during the first year of the COVID-19 pandemic using HPI and race/ethnicity. We first examine—separately—univariate COVID-19 mortality by HPI quartile (Figure 6) and univariate COVID-19 mortality by race/ethnicity (Figure 7). Table 3 and Figures 8a and 8b stratify cases, populations, and COVID-19 mortality rates by HPI quartile and race/ethnicity simultaneously. The bivariate analysis reveals that within each race/ethnicity group there is a social gradient with HPI that is strongly linear for all groups except Native Hawaiian/Pacific Islander. Not only are NHPI rates greater at each HPI quartile than other groups, their rates remain persistently high across the first three HPI quartiles before a large decline in Q4. Figures 8a and 8b plot the same data, but it is perhaps easier to see in Figure 8b the progressive linear decline in mortality with increasing HPI

Figure 6. Cumulative COVID-19 Mortality Rates per 100,000 by HPI 3.0 Quartile (Q1 = Least Healthy), California, February 2021



Sources: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California

Figure 7. Cumulative COVID-19 Mortality Rates (per 100,000) by Race/Ethnicity, California, February 2021



Sources: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California

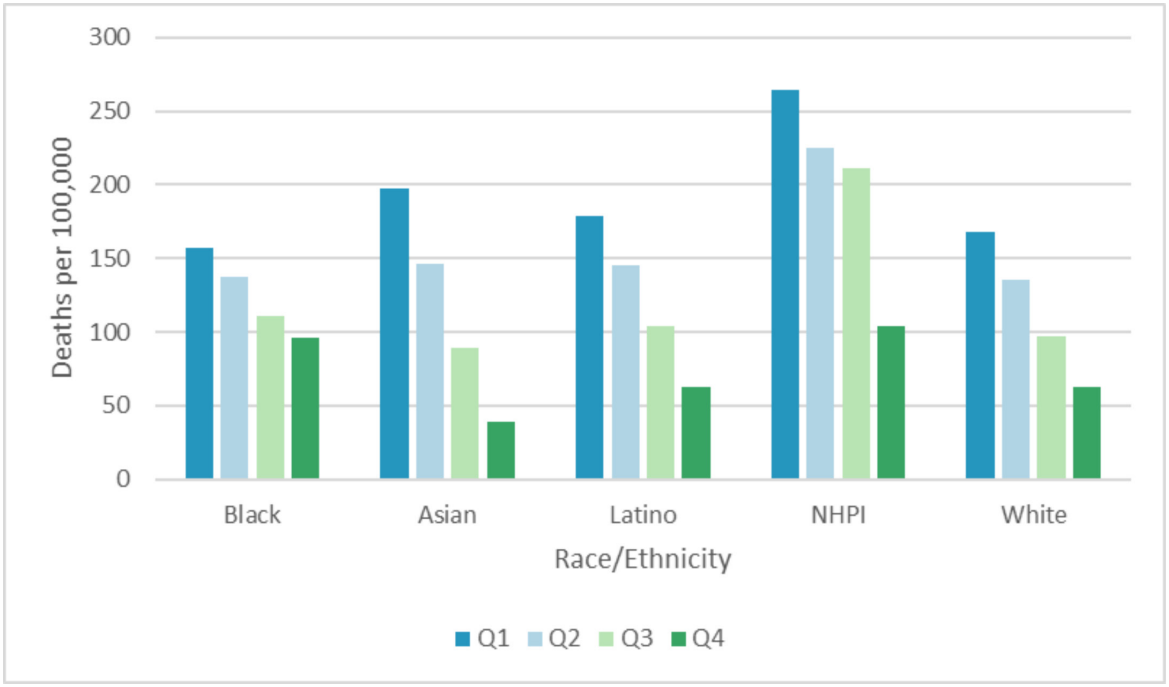
Table 3. COVID-19 Mortality Rates per 100,000 by Race/Ethnicity and HPI 3.0 Quartile (Q1 = Least Healthy), California, Reported by CDPH 2/24/2021

Race/ Ethnicity	Q1			Q2			Q3			Q4		
	Deaths	Population	Rate	Deaths	Population	Rate	Deaths	Population	Rate	Deaths	Population	Rate
Asian	1,953	986,616	198	1,963	1,344,777	146	1,362	1,530,290	89	644	1,652,391	39
Black	1,181	752,440	157	785	571,144	138	496	446,599	111	313	325,560	96
Latino	10,094	5,639,070	179	6,285	4,334,580	145	3,250	3,124,742	104	1,172	1,859,531	63
NHPI*	110	41,726	264	87	38,578	225	71	33,727	211	24	22,833	104
White	2,865	1,705,108	168	3,890	2,860,488	136	3,891	4,011,400	97	3,306	5,247,262	63

* NHPI, Native Hawaiian/Pacific Islander

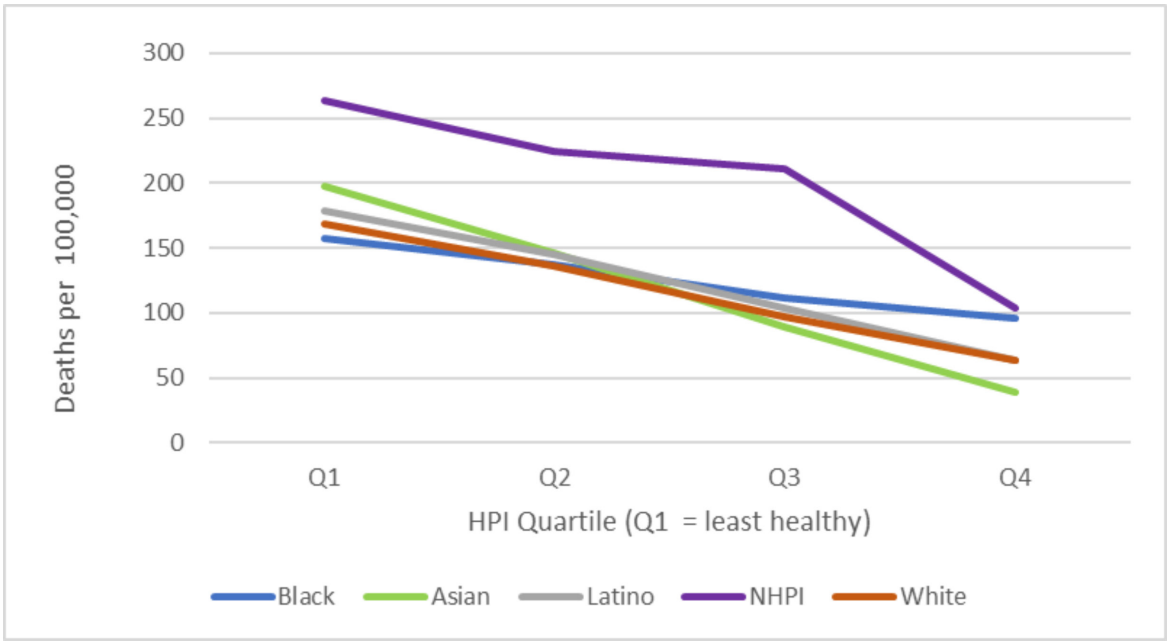
SOURCES: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California

Figure 8a. Bar Chart Series: COVID-19 Mortality Rates per 100,000 by Race/Ethnicity and HPI Quartile (Q1 = Least Healthy)



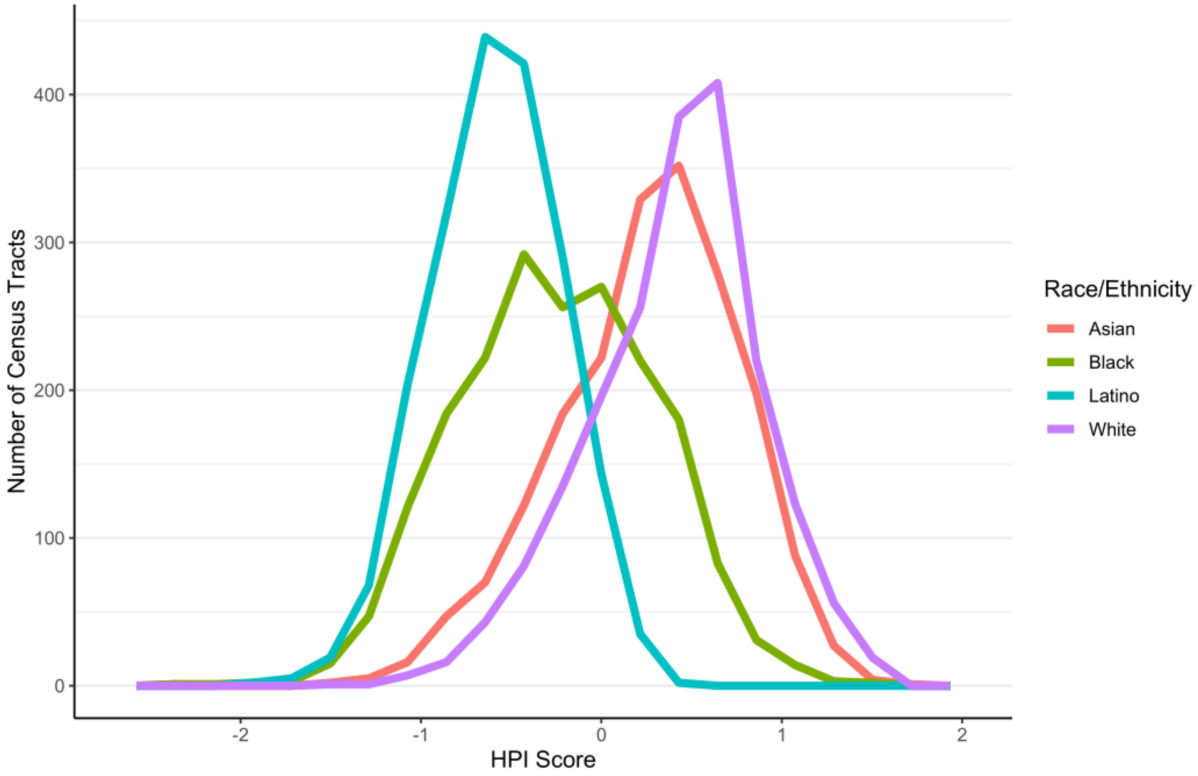
Sources: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California, Reported by CDPH 2/24/2021

Figure 8b. Line Graph Series: COVID-19 Mortality Rates per 100,000 by Race/Ethnicity and HPI Quartile (Q1 = Least Healthy)



Sources: California Department of Public Health, American Community Survey, Public Health Alliance of Southern California, Reported by CDPH 2/24/2021

Figure 9. Distribution of HPI 3.0 Scores in Quartile of Census Tracts With the Highest Proportion of Asian, Black, Latino, and White Residents, California, 2015-2019 (Higher HPI Score = Healthier Community Conditions)



SOURCES: Public Health Alliance of Southern California, American Community Survey

quartile for all but the NHPI group. One or more groups exhibiting a different HPI dose-response pattern across an independent variable like race/ethnicity is called statistical interaction, or, in epidemiologic terms, effect modification. This can be important to prioritize subgroups for additional resources. Based on Tables 3 and Figures 8a-b one can make the argument that Q1 populations irrespective of race/ethnicity and NHPI irrespective of their HPI quartile should be prioritized. Race AND place complement each other in the lens of health inequity.

Bivariate time series are an extension of the univariate time series concept. Visualizations may take more effort. For example, for HPI quartiles (4 groups) and race/ethnicity (8 or more groups), it is easier to see patterns between race/ethnicity groups when the time series are presented as panels in a grid in which each panel is a different race/ethnicity and the series are HPI quartiles (Figure9). Scaling the x (time) and y (outcome rates) in a consistent fashion across all panel facilitates comparisons.

MULTIVARIATE ANALYSES

Stratification can be extended to 3 or more equity variables (e.g., "trivariate") in an analysis of a health outcome. Unless the population is large, such as the entire State of California, higher levels of stratification often rarify the population into many subgroups with small populations and small or zero numbers of health events. While it is possible to ameliorate this by combining categories (e.g., White versus all other race/ethnicity groups combined), considerable detail may be lost and create an undetected pattern of low risks in numerically prevalent groups that obscures high risks in numerically small groups within the larger population.

Statistical models using regression techniques are a tool to analyze multiple health equity variables. Depending on the nature of the outcome and its distribution, logistic, Poisson, or negative binomial regression may be appropriate. A more complete

INTERPRETATION OF DIFFERENCES IN GROUPS

discussion is beyond the scope of this guidance. It is best to engage a biostatistician, epidemiologist, or other health sciences researcher to help with these types of advanced techniques. Although the models differ in detail, they are examples of linear additive models, and, to be analytically thorough, they should include the HPI as an equity variable and a means to detect statistical interactions between the HPI and other equity variables. These types of models can also use HPI score as a continuous variable, which may add statistical precision to the analysis.

Physicians, when presented with a patient exhibiting a specific constellation of signs, symptoms, and objective findings, apply a decision tree or logic to include or exclude possible diagnoses that match the information at hand. This is called a differential diagnosis. Epidemiologists employ a population version of this logic tree to potentially explain the results of statistical analyses. Elements of this epidemiologic decision tree are presented below.

RANDOM SAMPLE VARIATION

Cases and populations at specified time period can be thought of a sample of cases drawn from a much longer time line. In a theoretical world in which characteristics of cases and populations are unchanging, the average characteristics (age, gender, disease status, etc.) of the individuals will vary somewhat from sample to sample. Most samples will have an average that is close the average of the entire population. However, some samples, due to the luck of the draw, may have characteristics far different from the average, and may appear to drawn from a completely different population altogether. Statistical techniques are used to assess whether the difference in averages between two or more samples is a statistical fluke, acknowledging that random but extreme samples do occur every once in while. In the context of equity analyses, different subgroups within a category, such as HPI quartiles, can be thought as samples in which the average characteristic (e.g., disease rate) is subject to sample variation.

It is beyond this this guidance to be a primer on parametric statistics or probability. The reader should consult basic textbooks in epidemiology and biostatistics to learn about the variability in rates and proportions are characterized using standard deviations and confidence intervals) and various types of statistical tests (e.g., chi square) for count data that follow different statistical distributions (binomial, Poisson, etc.).[22] The terms "statistically different" or "significant difference" are specific terms of art in statistics, and do not convey a generic meaning that the differences are important or clinically meaningful. Free software (EpiInfo), developed by the Centers for Disease Control and Prevention, are available to carry out this type of statistical evaluation.[23]

Suffice it to say that as part of an "epidemiologic differential diagnosis", one important consideration is assessing whether differences between groups are likely or very unlikely explained by random sampling variation. A truism is that for a given difference, the smaller the sample size, the more difficult it will be to rule out random sample variation as an explanation of differences between subgroups. Sample size is also related to the concept of statistical power or the ability to discern whether differences are from sampling variation within one group or between two distinct populations.

It is worth noting that in the univariate analysis samples size per HPI quartiles is approximately equal. With additional stratification, sample sizes may be severely reduced, especially among race/ethnicity groups that are less numerous in the population (i.e., Native Hawaiian/Pacific Islander, American Indian/Alaskan Native).

CONFOUNDING AND THIRD VARIABLE RELATIONSHIPS

Another possible explanation of differences in rates between HPI quartiles arises when another equity variable that is associated with the same health outcome is also associated with the HPI. What if the Q1 only had a population of people aged 65 and older, and other HPI quartiles had progressively younger populations? Because age is associated with many health outcomes, would a finding that lower HPI quartiles have higher rates be really HPI differences masquerading as age differences? If the oddball distribution of age and HPI were reversed—Q4 were the oldest and Q1 the youngest, we may even see a counter-intuitive "social gradient." That age and HPI are independent risk factors for disease is based on evidence published in the scientific literature that should have been part of your literature review. What makes a third variable a confounding variable is the degree to which in your study population they have an overlapping distribution (e.g., older people of higher or lower HPI scores). Bi-variate stratification is a technique to unravel the role of confounding and other ways two independent variables may interact.

The distribution of HPI scores and race/ethnicity is somewhat overlapping (Figure 9). However, the race/place bivariate example of COVID-19 mortality rates shows that within each race/ethnicity group there is a consistent and strong social gradient. While it is true that Latinos and Blacks are overrepresented in Q1 compared to Asians and Whites, this unequal distribution can be ruled out as the reason for Q1 having the highest COVID-19 mortality rates.

EFFECT MODIFICATION

The example of COVID-19 mortality (Table 3, Figures 8a-b) also illustrates how bivariate stratification can identify population subgroups whose HPI dose-response may be different.

Figure 8b illustrates that the HPI dose-response was similar for Asian, Black, White, and Latino populations whose linear dose-response curves are almost superimposed on each other. However, Native Hawaiian/Pacific Islanders appear elevated across the first three HPI quartiles and rapidly approach the other groups in Q4. In this context we can say that race/ethnicity appears to modify the impact of HPI in NHPIs. It is beyond the scope of this guidance to discuss the statistical tools to assess heterogeneity of associations across a third variable. Readers should consult reference materials.[22]

DISEASE SEVERITY AND RISK ADJUSTMENT

Multiple, overlapping health conditions in the same person, or co-morbidities, is a measure of disease severity and an important component of clinical risk adjustment.[24] Patients and populations with co-morbidities are likely to be overrepresented in HPI Q1. However, it may be unwise to "correct" or statistically adjust for co-morbidities when they too are the consequence of low HPI scores. It is a basic methodological tenet in epidemiology and health services research that that a confounding factor not be caused by the exposure (HPI score). For assessment of health equity, this condition may not be true. That is, the outcome of interest and co-morbidities have the same causal origins in the social determinants of health. Incentives (via pay-for-performance) within healthcare systems may misplace responsibilities on individual clinicians, and risk adjustment using measures of the social determinants of health may minimize differences between equity subgroups. Risk adjustment in this context may obscure the community origins of disparities and poor health outcomes, and may distract healthcare systems from acting more forcefully to work with partners to improve community conditions. Some researchers make counter arguments that without risk adjustment, some healthcare systems and providers will simply avoid patients with challenging health and social conditions, exacerbating access to quality care.[25]

SYSTEMATIC BIASES

Bias is process and result of non-random difference between the true value of health outcome, HPI score and any other analytic variable and that which was measured. Random error was briefly discussed earlier in the context of variation in population samples. However, there are potential biases that are related to non-random error.

Population and Case Selection Bias

The healthcare system and other social systems can exert strong influences that discourage whole populations and those who are ill *not* to seek healthcare or delay healthcare until they are severely ill and in late stages of the disease process (severity bias). In order to be included in public health and clinical data, one must come in contact with the healthcare system. Access to healthcare, or the lack thereof, may be differentially expressed across HPI quartiles. It should be noted that the HPI includes health insurance coverage in adults as an individual indicator and domain, as well as other indicators that are linked to access such as income and mobility (automobile access).

Well-established factors that limit access include, lack of health insurance, underinsurance (high co-pays and high deductibles), linguistic barriers, geographic and transportation barriers (lack of private vehicle, transit access to healthcare facilities), cultural barriers, and immigration status, homelessness, and racist behaviors of health workers that range from unconscious to overt.[26] Many of these barriers to access are overrepresented in HPI quartile 1. Healthcare systems, through contracts with employers and healthcare providers, may not accept certain types of insurance (Medi-Cal), limit their service area, or otherwise seek out healthier populations in their underwriting and marketing strategies. Racial bias and cultural insensitivity in healthcare delivery has been well-documented[26] and may lead patients to minimize contact with healthcare providers. Thus, there is a tension been the underlying social drivers of disease causation and the social drivers of access to the healthcare system. In the case of death and severe illness, barriers to access may be overcome, even if only to issue a death certificate.

Because of this dynamic tension, the true HPI social gradient may be stronger than observed because of differential barriers to access in Q1 (i.e., equally ill patients are less likely to be counted in Q1). Under some circumstances the selection

pressures may be so great and illness mild enough that a social gradient may occur in a counter intuitive direction.

It is difficult to establish that selection bias has occurred in a specific example. Some diseases have formal clinical staging of disease progression. As a hypothetical example, the distribution of high-grade incident breast cancer may skew towards HPI Q1, which might suggest a delay in care or low utilization of preventive screening in HPI Q1. This is analogous to the actual observation that Black women have a skewed distribution of high-grade (i.e., more likely to spread) incident breast cancers, and have mortality rates that are higher than White women despite having overall lower incidence rate.[27]

Lack of dose response may be due to circumstances where the population is primarily restricted to HPI Quartile 1. This may occur in samples exclusively drawn from community clinics or Medi-Cal which significantly overlap with HPI Q1.

Non-Response Bias

This type of bias is particularly relevant for population surveys or patient satisfaction questionnaires that are based on voluntary participation. A large percentage of people asked to participate in these types of surveys often do not respond for a variety of reasons. They simply may not be at home when an interviewer calls or visits their home and interviewers give up after repeated attempts. There may be language barriers or poor translations of English-origin questionnaires. Respondents with low literacy in any language may not be able to understand questions or response options. There may be a lack of trust to share personal details. To counter these barriers, validated questionnaires in multiple languages delivered by bilingual and bicultural interviewers using multimodal outreach and administration (internet, mail, landline, cell phone) are commonly employed.

In some circumstances it might be possible to assess the direction and magnitude of non-response bias, particularly in healthcare systems. The health status of non-respondents as well as equity variables may be abstracted from clinical records. It might be possible to show systematic differences in the frequency of health events by age, gender, and HPI score between responders and non-responders. This will shed insight on both the direction and magnitude of potential bias based on rate ratios of responders and non-responders and magnitude the percentage of non-responders in the eligible population.

Another approach for healthcare facilities (hospital, community clinic, etc.) is compare HPI census tract distributions or average scores and sociodemographic characteristics of patients in the facility's catchment area with the characteristics of the underlying population.[28] There are many methods for defining catchment area based on the frequency of patient visits by geographic units such as ZCTAs or census tracts and travel times and distances. Once a catchment area is defined, the distribution of HPI scores (or an average) of patients a facility can be plotted with that of the underlying population, similar to Figure 9. Observing that the facility patients HPI scores are more left- or right-skewed than the general population in the same catchment area may suggest that population selection may be occurring. Why geographic areas with low HPI scores might appear to be under-represented or avoiding the facility might spark a discussion regarding community perception of the hospital, a closer examination of barriers to access, marketing strategies, and community outreach.

Missing Data

Missing data, including exclusion of eligible population and cases, and missing information on equity variables may lead to several types of potential bias. Reduced number of cases with complete information diminishes the effective sample size and the statistical reliability of the results. This not only affects univariate analyses, but reduces the usable number of observations for bivariate and multivariate analyses. When the amount of missing data is very large, it is reasonable to question the validity of the results altogether. However, even in public health surveillance systems that depend on individual healthcare providers to recognize and report cases, significant underreporting is not a barrier to observe trends that trigger preventive action, especially when the health outcome is serious (e.g., COVID-19).

Missing data may exhibit patterns that are not random for the health outcome and health equity variables. A troublesome type of bias occurs when missing data are most pronounced in those that have the health outcome and low HPI scores or other equity variables. The HPI score is predicated on a geocodeable residential address. For example, people experiencing homelessness may not have or report their address or are recorded in data systems with an unknown address. Homeless people may be homeless, in part, because they experienced health problems that led to their homelessness. Moreover, the unsheltered homeless experience dire living conditions that cause and exacerbate health problems. Homelessness is a type of severe resource deprivation most likely to occur in HPI Q1 populations. Missing data on homelessness will most likely lead to an underestimate of true disease rates in HPI Q1, and make the dose-response appear to be less steep had there not been missing data.

This type of differential missing data can play out in more subtle fashion. In clinical quality improvement, eligible population of patients may fail to return to get lab tests or treatments in part because they are too sick to schedule an appointment. If low HPI scores are more common in this group, missing data on the health outcome will bias the disease rates downward in HPI Q1.

One approach to assessing the magnitude of bias is to assign cases with missing data the worst outcome. For example, measures of glycemic control in people with diabetes requires routine laboratory tests for hemoglobin A1c within a specific measurement period. Rather than restrict the analysis only to those who had laboratory tests, analysts might score the cases with missing information as not-in- glycemic control. This would set the upper limit of the true value for the rate or percent not meeting the criteria for adequate glycemic control.

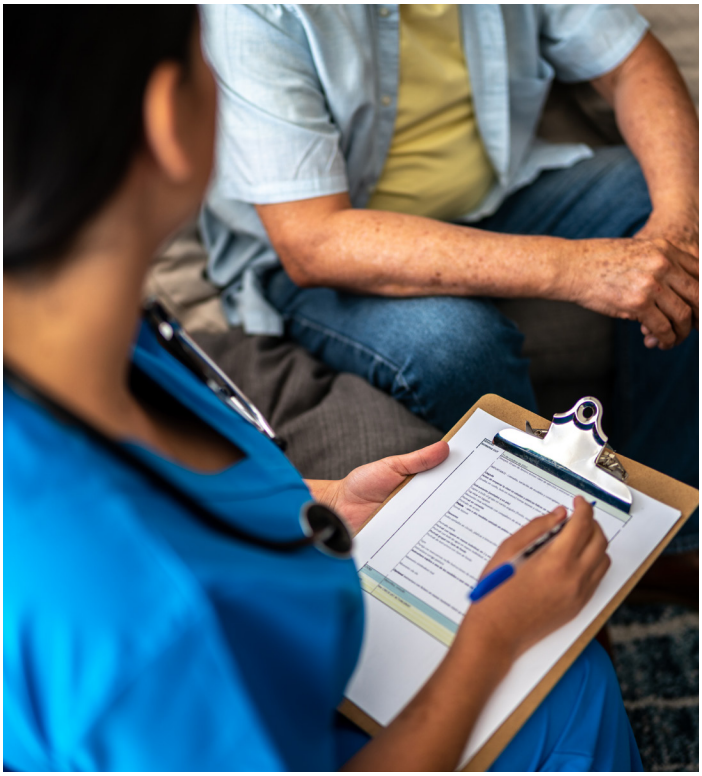
Inaccurate and Misclassified Data

We previously touched on inaccurate collection of race/ethnicity information. Self-identification is the preferred method to collect, but self-identified race/ethnicity may be reclassified by healthcare systems that strictly follow federal guidelines for classification of race/ethnicity. Misclassification of race/ethnicity in healthcare is not random among race/ethnicity categories. Populations and patients who are American Indian/Alaskan Native, Multi-race, and Latino are more likely to be misclassified, particularly when persons classifying these groups are not members of the group they are classifying.[29] There has been a long-standing cultural preference among Latinos to respond as "Other" to questions of their race. Extra effort to collect race/ethnicity data at disaggregated levels requires changes to data collection practices of healthcare systems and public health, some of which are underway by government agencies,[30] standards organizations, and vendors of electronic health records.

Experimenter Bias

Experimenter bias may range from subtle to flagrant and may occur at any stage of data collection, analysis, and interpretation, and reporting/communication of results. By experimenter we mean all the participants in public health surveillance or clinical quality improvement programs—from executive leadership to front desk staff collecting information. Sometimes individuals may not even be aware that routine decisions they make might be introducing bias into a study. Even the lack of interest to follow-up on the assessment of bias constitutes a bias.

The bulwark against experimenter bias includes the standardization of definitions for outcomes and health equity variables, standardization of data collection processes, explicit data quality control and assurance, rigorous training of participants, supervision, and protocolization of the public health surveillance activities and clinical quality improvement programs. Standardization of protocols is provided by governmental and non-governmental standard setting organizations in healthcare and public health. Analysts should develop a written data analysis plan before any analyses are conducted. The plan should identify the pre-planned analyses and even create empty, but formatted "key" tables that cover univariate and bivariate analyses. A written plan helps guard against improvisational data analysis that seeks to find "statistically significant" differences. This approach to data analysis is called data dredging and degrades the value of statistical testing, and, taken to extreme, may be a sign of incompetence or fraudulent conduct.



OVERALL ASSESSMENT OF BIAS

The preceding sections outline some, but not the potential biases that might constitute alternative explanations for results. As such, they form part of a checklist that should generate discussion among participants in the surveillance or clinical quality process. Even if the magnitude of the biases cannot be assessed, the likely net direction of the biases should be systematically considered and discussed. In formulating recommendations or a plan of action to follow up the results, additional data gathering and data audits should be considered to help resolve the direction and magnitude of potential biases.

ACTION PLANNING

In public health surveillance and clinical quality improvement, results, even after considering statistical analyses and assessment of biases, may not be definitive. Definitive results may be unachievable in local settings and may depend on more resource intensive research undertaken by academic centers covering multiple locations and studies.[31]

The threshold for action is considerably lowered when a dose-response relationship between HPI scores and health outcomes have been demonstrated in the data and that the relationship persists over levels of other health equity variables such as age, gender, and race/ethnicity. The larger the ratio between Q1 and Q4 rates, the more likely that social determinants of health are playing a contributory role. In the congenital syphilis surveillance example, a Q1/Q4 rate ratio of 17 is extraordinary, and few would doubt that social determinants play a role in the disease process. More modest Q1/Q4 rate ratios, including null (RR=1) or counterintuitive results may be due to biases or random variation in small samples. The observation of a modest Q1/Q4 rate ratio and gradual dose-response across HPI quartiles in consecutive years also suggest that the results are consistent and durable, which also lowers the threshold for action. Modest RRs in a large population portend many cases that are potentially preventable. Not taking action also risks standing idlily by while your public (public health) or members/patients (healthcare systems) continue to experience potentially avoidable adverse health outcomes.

Fortunately, there is a roadmap to taking "upstream" actions based on health equity analyses using the Healthy Places Index and the Health Outcomes Equity Tool (highlighted in the next section). First, the identity of Q1 census tracts (or ZCTAs) and their community conditions are inputs to the analysis. Using the HPI mapping tool (https://map.healthyplacesindex.org/), it is possible to visualize Q1 tracts. The mapping tool allows users to drill down to specific census tracts, or a pooled area of adjacent or clustered census tracts in order to examine which of the eight domains (policy action areas) and individual indicators are driving the low HPI score. Depending on the pattern of community conditions, it will be possible to identify the sectors (economic, education, social, housing, transportation, environmental pollution, neighborhood built-environment, and healthcare access) for which policy and services may be strengthened. Each of the 23 individual HPI indicators is linked to a policy guide (https://policies.healthyplacesindex.org/) which provides a menu of policy options to improve community conditions. The map and policy guides may also suggest strategies and partners for community engagement.

Many organizations may initially approach follow-up of results without a sense of where to start. The HPI mapping platform and policy guides provide multiple entry points for community engagement and policy action. The results of the analyses, maps of Q1 census tracts, and policy options can be shared and discussed with clinicians and community partners to identify and prioritize promising community-based actions, investments and policy options.

REFERENCES

1. Braveman P. Health Disparities and health equity: Concepts and measurement. *Ann Rev Public Health*. 2006;27:167-194. doi:10.1146/annurev.publhealth.27.021405.102103

2. Maizlish N, Delaney T, Dowling H, Chapman D, Sabo R, Woolf S, et al. N, Delaney T, Dowling H, et al. California Healthy Places Index: Frames Matter. *Public Health Rep*. 2019;134:354-362.

3. World Health Organization. *International Classification of Diseases, 10th Revision*. Geneva, Switzerland: World health Organization; 2019. Accessed April 8, 2023. https://www.who.int/classifications/classification-of-diseases

4. National Notifiable Diseases Surveillance System (NNDSS). *Surveillance Case Definitions for Current and Historical Conditions*. Atlanta, GA: Centers for Disease Control and Prevention; 2023. Accessed April 8, 2023. https://ndc.services.cdc.gov/

5. Centers for Medicare & Medicaid Services. *Specifications Manual for National Hospital Inpatient Quality Measures. Version 5.12*. Washington, DC: Centers for Medicare & Medicaid Services; 2022.

6. Agency for Healthcare Research and Quality. *AHRQ Quality Indicators*. Rockville, MD: Agency for Healthcare Research and Quality; 2022. Accessed September 5, 2022. https://qualityindicators.ahrq.gov/

7. National Committee for Quality Assurance. *Healthcare Effectiveness Data and Information Set (HEDIS).* Washington, DC: National Committee for Quality Assurance; 2023. Accessed April 8, 2023. https://www.ncqa.org/hedis/measures/

8. American Medical Association. *Current Procedural Terminology (CPT).* Chicago, IL: American Medical Association; 2023. Accessed April 8, 2023. https://www.ama-assn.org/amaone/cpt-current-procedural-terminology

9. Regenstrief Institute. *Logical Observation Identifiers Names and Codes (LOINC).* Indianapolis, IN: Regenstrief Institute; 2023. Accessed April 8, 2023. https://loinc.org/

10. U.S. Food & Drug Administration. *National Drug Code (NDC)*. Silver Spring, MD; 2023. Accessed April 8, 2023. https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory

11. California Department of Public Health, California Conference of Local Health Officers. *County Health Status Profiles, 2022*. Sacramento, CA: California Department of Public Health Accessed April 8, 2023. https://www.cdph.ca.gov/programs/chsi/pages/county-health-status-profiles.aspx

12. UCLA Center for Health Policy Research. *California Health Interview Survey (CHIS)*. Los Angeles, CA University of California; 2022. http://healthpolicy.ucla.edu/chis/Pages/default.aspx

13. Agency for Healthcare Research and Quality. *Consumer Assessment of Healthcare Providers and Systems (CAHPS)*. Rockville, MD: Agency for Healthcare Research and Quality; 2023. Accessed April 8, 2023. https://www.ahrq.gov/cahps/index.html

14. U.S. Census Bureau. *Surveys (American Community Survey, Decennial Census)*. Washington, DC: U.S. Census Bureau; 2023. Accessed April 8, 2023. https://data.census.gov/advanced

15. California Department of Finance. *Demographic Reports*. Sacramento, CA: California Department of Finance; 2023. Accessed April 8, 2023. https://dof.ca.gov/reports/demographic-reports/

16. Krieger N, Chen J, Waterman P, Rehkopf D, Subramanian S. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: The Public Health Disparities Geocoding Project. *American Journal of Public Health*. 2005;95:312-323.

17. Wilkerson R, Marmot M eds. *Social Determinants of Health: The Solid Facts. Second Edition*. Geneva, Switzerland: World Health Organization; 2003. Accessed April 10, 2023. https://www.euro.who.int/__data/assets/pdf_file/0005/98438/e81384.pdf

18. Bodenreider C, Damicis A, Delaney T, et al. *Healthy Places Index (3.0): Technical Report*. Los Angeles, CA: Public Health Alliance of Southern California; 2020. Accessed October 25, 2022. https://assets.website-files.com/613a633a3add5db901277f96/63320a9e98493bbdcc03d509_HPI3TechnicalReport2022-09-20.pdf

19. Public Health Alliance of Southern California. *Why Isn't There an HPI Score for My Community?* Oakland, CA: Public Health Alliance of Southern California, Public Health Institute

20. Urban Indian Health Institute. *Best Practices for American Indian and Alaska Native Data Collection*. Seattle, WA: Seattle Indian Health Board; 2020. Accessed April 18, 2022. https://www.uihi.org/resources/best-practices-for-american-indian-and-alaska-native-data-collection

21. California Highway Patrol. *Statewide Integrated Traffic Records System (SWITRS)*. Sacramento, CA: California Highway Patrol; 2023. Accessed April 10, 2023. https://www.chp.ca.gov/programs-services/services-information/switrs-internet-statewide-integrated-traffic-records-system

22. Centers for Disease Control and Prevention. *An Introduction to Applied Epidemiology and Biostatistics. Third Edition*. Atlanta, GA: Centers for Disease Control and Prevention; 2012. https://www.cdc.gov/csels/dsepd/ss1978/index.html

23. Centers for Disease Control and Prevention. *EpiInfo*. Atlanta, GA: Centers for Disease Control and Prevention; 2023. Accessed April 9, 2023. https://www.cdc.gov/epiinfo/index.html

24. Agency for Healthcare Research and Quality. *Making Adjustments to Health Care Quality Scores*. Rockville, MD: Agency for Healthcare Research and Quality; 2023. Accessed April 8, 2023. https://www.ahrq.gov/talkingquality/translate/scores/adjustment-scoring.html

25. Rogstad T, Gupta S, Connolly J, Shrank W, Roberts E. Social Risk Adjustment In The Hospital Readmissions Reduction Program: A Systematic Review And Implications For Policy. *Health Affairs*. 2022;41(9):1307-1315. doi:https://doi.org/10.1377/hlthaff.2022.00614

26. Agency for Healthcare Research and Quality. *National Healthcare Quality and Disparities Report*. Rockville, MD: Agency for Healthcare Research and Quality; 2023. https://www.ahrq.gov/research/findings/nhqrdr/nhqdr22/index.html

27. Giaquinto A, Sung H, Miller K, et al. Breast Cancer Statistics, 2022. *CA: A Cancer Journal for Clinicians*. 2022;72:524-541. doi:10.3322/caac.21754

28. Saini V, Brownlee S, Gopinath V, Smith P, Chalmers K, Garber J. *Methodology of the Lown Institute Hospitals Index for Social Responsibility*. Needham, MA: The Lown Institute; 2022. Accessed April 8, 2023. https://lownhospitalsindex.org/wp-content/uploads/2022/06/2022_Lown_Index_Methodology.pdf

29. Wikipedia contributors. *Racial and Ethnic Misclassification in the United States*. Wikipedia, The Free Encyclopedia; 2023. Accessed April 10, 2023. https://en.wikipedia.org/w/index.php?title=Racial_and_ethnic_misclassification_in_the_United_States&oldid=1136441241

30. Office of Management and Budget. Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards. *Federal Register*. 2023;88:5375-5384. doi:https://www.federalregister.gov/documents/2023/01/27/2023-01635/initial-proposals-for-updating-ombs-race-and-ethnicity-statistical-standards

31. Braveman P, Egerter S, Woolf S, Marks J. When do we know enough to recommend action on the social determinants of health? *Am J Prev Med*. 2011;40(11):S58-S66.