

# Medical Claims Data Characterizes Heat Health Risk for Low-Income and Agricultural Communities in California

I. Avery Bick<sup>\*1,2,3</sup>, Gina Hervey<sup>1,2</sup>, and Jenny Suckale<sup>4</sup>

<sup>1</sup>Climate and Energy Policy Program, Stanford University,  
Stanford, CA

<sup>2</sup>Environmental and Natural Resources Law and Policy Program,  
Stanford University, Stanford, CA

<sup>3</sup>Norwegian Institute for Nature Research, Oslo, Norway

<sup>4</sup>Doerr School of Sustainability, Stanford University, Stanford, CA

\*Email: [averybick@stanford.edu](mailto:averybick@stanford.edu)

## Abstract

Heat exposure is a significant cause of morbidity and mortality around the world that is expected to worsen with climate change. In California, efforts to map the health risks of heat exposure have focused mainly on compiling indices that assess heat vulnerability theoretically, rather than empirically. Here, we use complete Medicaid claims data from 2011-2019 to map heat impacts for low-income Californians at a ZIP code scale. We find that the top 10% of ZIP codes by heat-related claim rates tend to have lower median income, a higher percentage of farm workers, and a higher rate of mobile homes, compared to the lower 90% of ZIP codes. We find that heat-related claim rates increase 24.4% for every 1 C ° in majority cropland areas, compared to 20.6% for every 1 C ° in majority built-up areas. Comparing heat-related claim rates against three common heat and socioeconomic vulnerability indices, we find that two have a weak to moderate positive correlation, while the third is weakly negatively correlated. These results highlight the importance of considering medical claims data when mapping heat impacts and designing interventions to reduce heat-related health risks and public costs.

## Keywords

Heat, Public Health, Climate, Medical Data, Agriculture

## Synopsis

Health impacts of heat exposure are increasing with climate change. Existing tools for mapping heat risk typically quantify hypothetical vulnerability to heat, but do not integrate actual historical data on heat-related illnesses. This study uses Medicaid claims data to map historical heat risk in California, identify specific communities and demographics at risk, and evaluate the effectiveness of vulnerability indices and policy interventions.

## Introduction

Heat exposure now causes more deaths in the United States than any other weather event [1]. From 2004 to 2018, an average of 702 people died annually from heat-related causes [2], increasing to 1,930 deaths annually between 2019-2023 [3]. This trend in heat-related deaths is projected to continue in the coming years, both in the United States and worldwide [4, 5, 6]. In the United States, California is one of the states where heat waves regularly affect millions of people, especially in the agricultural sector, where many low-income workers are exposed to high temperatures [7, 8].

Increased mortality from heat is only one of many heat-related health impacts [9]. Heat exposure can also cause dehydration, heat exhaustion, heat stroke, and acute kidney failure, all of which are preventable in principle [10] and with improved methods of identifying high heat risk, could be minimized. Potential methods for identifying where and to whom such health impacts occur include the use of vulnerability indices that composite many risk factors and, increasingly, medical claims and emergency room data [6, 11, 12].

Attention to heat risk is increasing in California, as evidenced by the creation of AB 2238 in 2022, which requires the state to develop a statewide extreme heat ranking system, and the Extreme Heat Action Plan that outlines comprehensive state actions on extreme heat [13]. However, a recent assessment of Heat Action Plans in California shows that many plans are incomplete and would benefit from a more deliberate focus on health outcomes and the consideration of health data [14]. Nonetheless, California does not currently have a systematic statewide assessment of heat-related health impacts.

The goal of this paper is to better understand the spatial distribution of heat risk in California and the socioeconomic and environmental factors associated with heat-related illnesses in low-income households. We use complete patient-level claims data from 2011-2019 for Medi-Cal [15, 16], California’s Medicaid system, considering both inpatient and outpatient claim counts and rates for heat-related illnesses within the International Classification of Diseases (ICD) category “Effects of Heat and Light” on a ZIP-code scale. This ICD category includes heatstroke, heat exhaustion, heat edema (heat-induced swelling), and heat syncope (heat-induced fainting and dizziness). The data set captures approximately 13 million certified eligibles in 2019 [17], about 32% of the State’s population that year [18].



By calculating the total heat-related claims and claim rate (per eligible) on the ZIP-code scale, we are able to identify specific communities facing disproportionate heat risk in California and compare their characteristics. More specifically, we consider median household income and agricultural workforce size derived from U.S. Census data [19], mobile home space counts from the California Department of Housing and Community Development [20], and land cover from the European Space Agency WorldCover dataset [21]. The main drawback of this high spatial resolution is that small ZIP codes have few Medicaid eligibles, introducing significant noise into the estimation of claim rates. We mitigate this issue through a Bayesian spatial smoothing technique [22] based on information from surrounding ZIP codes.

We find that heat health risks among low-income households vary dramatically across communities in California driven by both temperature and socioeconomic characteristics. This result has important implications for California's Extreme Heat and Community Resilience Program (EHCRP) where local and regional governments can apply to support projects aimed at reducing heat risk [23]. However, the metrics for determining awardees do not include heat risk metrics, let alone heat-related hospitalization data [24]. The program prioritizes support for disadvantaged and heat-vulnerable communities through the designation of "disadvantaged communities" with the CalEnviroScreen index [25]. An important open question is whether CalEnviroScreen and other indices guide funding towards community where it would have the biggest impact on heat-risk reduction.

California uses a suite of vulnerability indices from various state and federal agencies to measure chronic heat vulnerability and real-time heat hazard. These vulnerability indices include the California Heat Assessment Tool (CHAT) Heat Health Action Index [26], the Centers for Disease Control and Prevention (CDC) Heat and Health Index [27], and CalEnviroScreen [25]. These composite indices include numerous heat-vulnerability covariates such as temperature, age, and occupation but do not empirically measure heat-related medical outcomes. As a consequence, these indices are not always consistent with real-world data documenting impacts [28] and have limited ability to predict how heat risk varies by region [29].

A complement to heat-vulnerability indices are monitoring tools such as CHAT's "heat health events" (HHEs) metric, which estimates the number of annual heat events by region which may threaten public health, based on historical data [26]. The focus on historical data enables a quantification of relative exposure to extreme heat events, but also directs the metric towards areas where heat has posed large risks to public health in the past. However, past heat exposure is not necessarily predictive of future impacts, partly because mortality risk increases rapidly beyond a certain threshold temperature [30, 31] and partly because vulnerability might be a more important driver of risk than exposure [32]. Another set of heat risk indicators, such as CalHeatScore [33] and the National Weather Service's HeatRisk tool [34], are geared towards short-term warnings and forecasts of extreme heat. One data product that does measure quantities of heat-related illness is the CDC Heat and Health Tracker [35], which releases

daily rates of emergency department visits, but these rates are aggregated regionally, limiting their applicability to community-level interventions.

## Methods

### Medicaid Claims

We source heat-related illness claims from complete Medicaid Inpatient [15] and Outpatient [16] data from 2011-2019 within California, supplemented by Medicaid Demographic [36] and Eligibility [17] data. The combined dataset includes claims for patients at outpatient facilities, such as urgent care, emergency rooms, and physicians' offices, as well as patients who underwent inpatient stays at hospitals. It also includes the age and residential ZIP codes of patients.

The Medicaid database utilizes the International Classification of Diseases (ICD) coding system, a standardized system for classifying diseases that is often applied for reporting medical claims. ICD diagnosis codes are primarily used for billing purposes, and thus do not always represent instances of bona fide diagnoses, but are commonly applied in epidemiological studies [37]. In October 2015, as part of a national effort, California shifted from reporting medical claims in ICD version 9 to version 10 [38]. During the same year, the use of the Medicaid Analytic eXtract (MAX) system for the compilation and distribution of Medicaid data was discontinued in favor of the Analytic File (TAF) system of the transformed Medicaid Statistical Information System, which is generally more comprehensive with higher data quality [38].

We aggregate both TAF and MAX data and count all codes falling under the ICD class "Effects of Heat and Light", denoted by the prefixes "992" in ICD-9 and "T67" in ICD-10. This class includes heatstroke, heat cramps, heat exhaustion, heat fatigue, heat edema (swelling in extremities), and heat syncope (dizziness or fainting). Each individual outpatient claim can have up to two diagnosis codes, while inpatient claims can have up to 12. Medical claims data is subject to restrictions to protect patient privacy that limit the display of any cell with a size of 15 or less for outputs that include individuals under 18. Thus, any count of Medicaid claims or unique patients between 1-15 in a ZIP code is displayed as "1-15", though counts of 0 are shown. We also exclude rates calculated using cell sizes of 1-15.

We perform all processing of Medicaid data on the secure Redivis cloud platform [39]. First, we filter all MAX and TAF inpatient and outpatient claims data for residents of California. Next, we drop visits that do not have a documented patient residential ZIP code, a total of 547 claims, less than 1% of the total of 100,928 relevant heat-related claims. We then aggregate heat-related health codes based on the residential ZIP codes of patients for that year, as reported in the Medicaid Personal Summary dataset [36]. Finally, we prepare count data by aggregating heat-related illness claims by patient residential ZIP code.

To develop a crude average annual claim rate by ZIP, we divide the claim

counts by the mean monthly Medicaid eligibles in each ZIP across our study period (see Fig. 1B), based on data provided by the California Department of Health Care Services [40]. We then divide by the nine years in our study period, 2011-2019, to calculate the crude rate. Since ZIP codes differ significantly in the mean of monthly Medicaid eligibles, we employ spatial smoothing to reduce noise from small-number bias by borrowing information from neighboring regions [41]. More specifically, we use the BYM (Besag, York and Mollié) hierarchical Bayesian model [42] designed for the spatial smoothing of epidemiological data.

The BYM model accounts for spatial patterns using an Intrinsic Conditional Autoregressive Model (ICAR), which assumes that neighboring regions have similar disease rates, while also accounting for unstructured, random variation in each region. We apply a reparameterized version of the model, BYM2 [22], which introduces a mixing parameter that quantifies the proportion of variance associated with spatial patterns versus unstructured, random variation.

We define the neighbors of each ZIP as the 10 nearest ZIPs based on geographic centroids. The neighborhood structure in BYM2 is captured by a symmetrized adjacency matrix. The symmetrization process can create more neighbors for some regions, as regions selected by others can gain additional relationships. This larger neighbor sample can decrease variance for certain regions. Within the BYM2 model, this is corrected by scaling the ICAR spatial component based on the average neighborhood size. This process ensures fair comparison between spatial patterns and random noise [22]. In our BYM2 implementation, we specify a negative binomial distribution for likelihood to account for overdispersion in disease counts, common with real-world disease data [41]:

$$\mu_i = E_i \cdot \exp \left( \beta_0 + \sigma \left[ \sqrt{1 - \rho} \cdot \theta_i + \sqrt{\frac{\rho}{s}} \cdot \phi_i \right] \right) \quad (1)$$

where:

$\mu_i$  = expected claim count in ZIP  $i$

$E_i$  = baseline expected count for ZIP  $i$  (ZIP eligible population  $\times$  global rate)

$\beta_0$  = intercept

$\sigma$  = scaling parameter

$\rho$  = mixing parameter

$\theta_i$  = unstructured random effect

$\phi_i$  = spatial pattern random effect

$s$  = scaling factor

After calculating the expected claim count,  $\mu_i$ , by ZIP  $i$ , we calculate smoothed rates by dividing the expected claim count by the mean Medicaid eligible population in each ZIP.

## Geospatial Datasets

We source socioeconomic data and ZIP code geometries from the 2015-2019 U.S. American Community Survey data provided by the Integrated Public Use Microdata Series (IPUMS) National Historical GIS (geographic information system) tool [19]. This dataset represents ZIP code geographies as ZIP code Tabulation areas (ZCTA), an approximation of ZIP postal codes used by the U.S. Census. These datasets include median household income (IPUMS field ALW1E001) and the number of workers in farming, forestry, and fishing by sex (IPUMS variables ALY6E031 and ALY6E067). We calculate the overall fraction of the workforce in the farming, forestry, and fishing industries by ZIP by summing fields ALY6E031 and ALY6E067 and dividing by the total workforce size (IPUMS field ALY6E001).

We source daily maximum air temperature from the 1 km<sup>2</sup> resolution Daymet [43] dataset, then calculate the mean maximum daily air temperature for each ZCTA across our study period of 2011-2019 (see Fig. 1C). We derive the land cover of each ZCTA from the 10 m resolution ESA WorldCover 2020 Global land cover product, which is based on imagery from the Sentinel-1 and 2 satellites [21] (see Fig. 1A). Using zonal statistics, we calculate the relative fraction of each ZCTA covered by each land use type. We focus on three land cover types that have been shown in literature to modulate heat-related illness risk [44, 45, 8, 46]: cropland, defined as land sowed and harvested at least once per year, built-up areas, covered by man-made structures and roads, and tree cover, which characterizes areas with at least 10% tree cover. All zonal statistics calculations were performed using the Python library rasterstats [47].

We also investigate associations of heat-related claim rates with relative rates of mobile homes by ZIP. We derive the number of mobile home spaces per ZIP code from records of active mobile home parks in 2024 from the California Department of Housing and Community Development [20]. We then assume that the 2024 data is indicative of the years 2011-2019 as well, since these specific years are not available in the database. We calculate relative rates of mobile homes per ZIP code by dividing the total number of active mobile home spaces by the mean Medicaid eligible count in each ZIP code across the study period.

## Vulnerability Indices

We compare our ZIP heat-related claim rates against the California Heat Assessment Tool (CHAT) Heat Health Action Index [26], the CDC Heat and Health Index [27], and the CalEnviroScreen Population Characteristics Score [25]. The CDC Heat and Health Index and the CalEnviroScreen data are available at ZIP-code resolution, and thus directly comparable against heat-related claim rates. However, the CHAT Heat Health Action Index data is at a Census Tract scale. We use a 2019 ZIP to Tract crosswalk file [48] from the U.S. Department of Housing and Urban Development to convert the Census Tract-scale index to ZIP-scale. This crosswalk provides the fractions of residential addresses in each

ZIP that come from Census Tracts within that ZIP [49]. Thus, we calculate a ZIP-scale CHAT Heat Health Action Index by performing a weighted average of the tract-scale Heat Health Action Index values based on the fraction of ZIP code residential addresses contained within each Census Tract.

## Analyses

### Evaluating the Relationship of Socioeconomic and Environmental Characteristics to Heat Risk

To determine whether the distributions of these demographic indicators were significantly different in high- and low-heat risk ZIP codes, we first split all ZIP codes into groups of relative high risk (those within the top 10% of BYM2-smoothed heat-related claim rates) and low risk (the remaining ZIP codes). We apply a one-sided permutation tests with 1000 resamples and mean statistics (`scipy.stats.permutation_test`) [50] to evaluate whether the distributions of median household income, percentage of farm workers in the workforce, ZIP code land use coverage fraction (tree cover, cropland, and built-up area), and mobile home space rates are significantly different between low- and high-heat risk ZIP codes. We also use Spearman rank correlation tests [50] to evaluate the direction, strength, and significance of monotonic relationships between BYM2-smoothed heat-related claim rates and mean daily maximum temperature and the three aforementioned vulnerability indices.

### Heat Wave Analysis

To determine whether Medicaid claim data can capture signals of heat health events such as heat waves, we calculate the daily Medicaid claim count before, during, and after the 2017 heat wave in the Central Valley of California. We first subset California ZIP codes to those that fall at least 50% within the bounds of the Central Valley. In the absence of official Central Valley borders, we use the borders as described by the 2003 Central Valley Historic Mapping Project [51].

We identify the start date (2017-06-17) and the end date (2017-06-27) of the heat wave based on a 2024 report from the California Department of Insurance [52]. For the heat wave analysis, daily heat-related claim counts were calculated in a similar fashion as previously described in the main analysis. We first filter inpatient and outpatient claims to remove claims without residential ZIP codes, as well as claims with residential ZIP codes outside the designated Central Valley region. We then sum and compare the counts of heat-related claims during the 11 days of the heat wave to the 11 days before and after the heat wave.

### Heat Claims by Age and Gender

To compare heat-related claim counts of men and women by age, we process inpatient and outpatient claim counts. We filter claims to include only California and to remove ZIP codes without patient residential ZIP code data. We then derive the patient sex and calculate the age of the patient at the date of service

using the birth date provided in the Medicaid Personal Summary dataset [36]. We drop any claims without information on date of birth, date of service, or patient sex. Finally, we aggregate total heat-related claim counts by integer age for male and female patients separately across the study period into a histogram.

### Regression Model

Following the hypothesis that farm workers face high heat risk, we fit a multiple linear regression model using ordinary least squares (OLS) to model log-transformed heat claim rates as a function of temperature and land use. We calculate the log-transformed heat claim rates by first aggregating heat-related claim counts by ZIP code and day, dropping any claims without data on patient residential ZIP codes. We then aggregate the daily ZIP counts by daily maximum temperature (in 0.5°C buckets) and dominant land cover type. We set the temperature range of interest between 20°C and 48 °C, as few heat-related illness claims occurred below this range in “cropland” areas or above this range in both “cropland” and “built-up” areas. We then filter the analysis to ZIP codes that are either majority “built-up” or majority “cropland”. We also calculate the number of person-days of heat exposure to each 0.5°C temperature bucket for each land cover type by summing the mean Medicaid eligible count for each ZIP and day.

After log-transforming the claims per person-day data ( $r$ ), we apply the OLS model with a categorical variable ( $C$ ) based on whether dominant land cover (LULC) was cropland or built-up:

$$\log(r) \sim T + C(LULC) + T : C(LULC)$$

where the  $C(LULC)$  term creates separate y-intercepts for baseline heat illness claims per person-day categorized by land cover type. We also include an interaction term  $T : C(LULC)$ , which allows different slopes for each land cover type, allowing temperature ( $T$ ) to affect claims per person-day differently in built-up and agricultural areas. To account for heteroscedasticity in model residuals, we report heteroscedasticity-robust standard errors using HC3 covariance estimation [53]. Model parameters are reported in the Supporting Information. We construct the model using the Python library statsmodels [54].

## Results

### Heterogeneous Spatial Distribution of Heat-Related Claims and Rates

Between 2011 and 2019, we identify 98,456 outpatient claims stemming from 27,079 beneficiaries for heat-related illnesses billed to California Medicaid and 1,925 inpatient heat-related claims from 1,706 patients. We map claim counts and smoothed claim rates on the ZIP code scale in Fig. 2a and Fig. 2b, respectively. Heat-related claim counts cluster in the southeast, parts of the Central

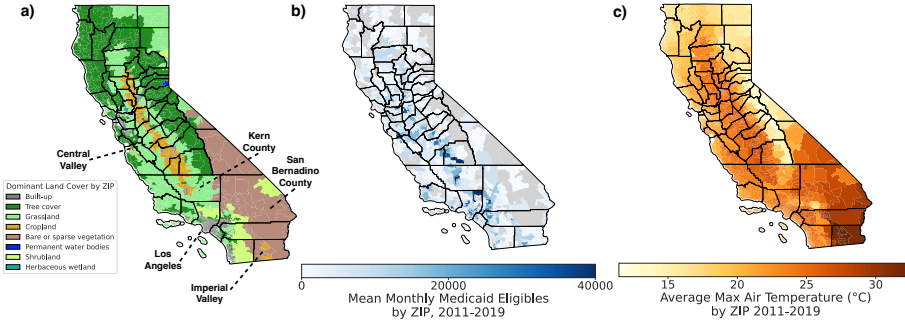


Figure 1: A) Majority land cover [21], B) Medicaid enrollment [?] at the ZIP code scale in California, and C) average daily maximum temperature (C°) by ZIP, 2011-2019 [43]. County borders shown in black.

Valley, and generally in rural ZIP codes (Fig. 2a). The ZIP codes with the highest claim counts (92243, 93307, 92227, 92201, and 93706 respectively) contain a mix of built-up and cropland areas, with relatively high mean air temperatures (Fig. 1c) and low tree cover (Fig. 1a).

The BYM2 [22] spatial smoothing model converged well, with all parameters achieving a Gelman-Rubin statistic ( $\hat{R}$ ) of 1.00 (SI Table 1). The spatial distribution of rates (Fig. 2b) is even more heterogeneous than the count (Fig. 2a) across ZIP codes ( $\sigma = 4.56$ , 95% HDI: [3.94, 5.19]). We find that 98.8% of variation in heat-related claim rates explained by spatial or neighborhood effects and only 1.2% of variation due to random effects in individual ZIP codes (SI Table 1). The primary clusters of high heat-related claim rates are located in the agricultural Imperial Valley and Kern County areas, San Bernardino County, and select rural ZIP codes in Northern California (Fig. 2b).

Among more populated ZIP codes with at least 1,000 mean monthly Medicaid eligibles, smoothed heat-related claim rates were generally highest in southeastern regions that contain a mix of agricultural, urban, and desert landscapes (Fig. 1a). These areas include ZIP codes within Imperial County (92257, 92283, 92227, 92243, 92281), San Bernardino County (92363), and Riverside County (92225). We also note high rates in some more urban areas, such as Santa Clara County (95113) and in sparsely populated rural ZIP codes in Northern California in Trinity (95595), Humboldt (95511), and Colusa (95955) counties.

## Socioeconomic and Environmental Characteristics of High-Risk ZIP codes

To identify the socioeconomic and environmental characteristics of high-heat risk ZIP codes, we compare the top 10% of ZIP codes by heat-related claim rates (high-risk ZIPs) between 2011-2019 to the remaining ZIP codes (low-risk ZIPs). In this comparison, we only consider ZIP codes with at least 1,000 mean monthly eligibles (Fig. 3). Via permutation test (see Methods), we find that high-risk

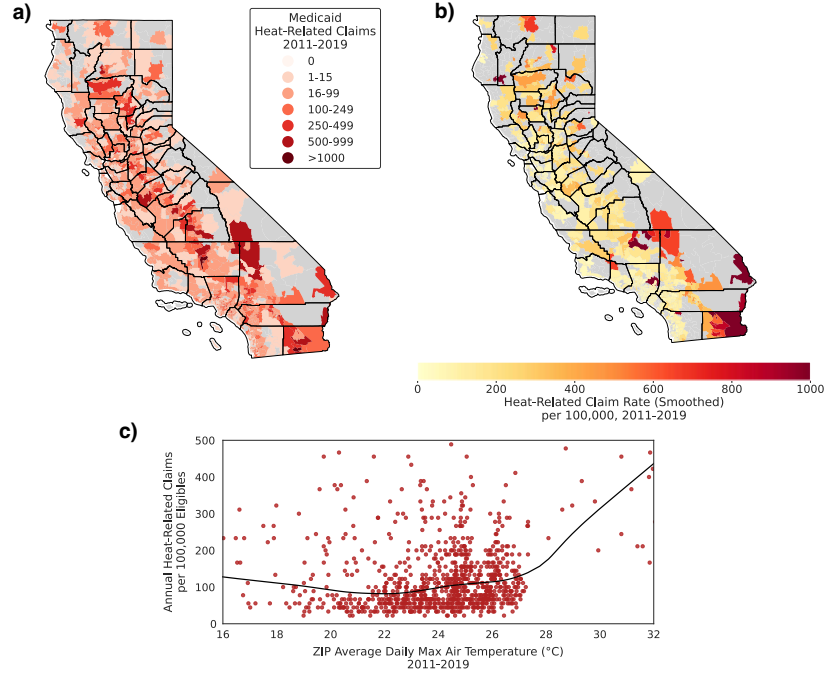


Figure 2: Heat exposure and heat-related medical claims across California. A) The total inpatient and outpatient heat-related Medicaid claims by ZIP from 2011-2019 [15, 16]. B) Population-adjusted annual heat-related claims per Medicaid eligible by ZIP from 2011-2019. C) Scatter plot of average annual rates of heat-related claims across 2011-2019 versus the average maximum daily air temperature for all California ZIP codes. The black line is a locally-weighted smoothing curve (LOESS) indicating trends in the scatter.



ZIP codes tend to have higher percentages of farm workers in the workforce than low-risk ZIP codes ( $N_{highrisk}=142$ ,  $N_{lowrisk}=1,214$ , test stat.: 0.0187,  $p=0.008$ ) (Fig. 3b), lower median household income ( $N_{highrisk}=140$ ,  $N_{lowrisk}=1204$ , test stat.: -\$24,553,  $p < 0.001$ ) (Fig. 3d), and a significantly higher rate of mobile home spaces per Medicaid eligible ( $N_{highrisk}=112$ ,  $N_{lowrisk}=781$ , test stat.: 0.0673,  $p < 0.001$ ) (Fig. 3b).

Interestingly, we find that high-risk ZIPs tend to have slightly higher tree cover fraction ( $N_{highrisk}=143$ ,  $N_{lowrisk}=1221$ , test stat.: 0.0660,  $p=0.002$ ) (Fig. 3e) and significantly lower fraction of built-up area ( $N_{highrisk}=143$ ,  $N_{lowrisk}=1221$ , test stat.: -0.311,  $p < 0.001$ ) than low-risk ZIPs (Fig. 3f). One possible explanation is that rural areas face higher heat risk, potentially due to a greater fraction of outdoor workers or other lifestyle or demographic differences, despite also having more tree cover. The urban heat island effect associated with built-up areas may be counteracted by other factors, such as fewer outdoor workers or greater air conditioning access. We also note that heat-related claim counts are similar for young men and women, but are higher for middle-aged men compared to middle-aged women (Fig. 4). Indeed, the median age of patients with heat-related claims is 35 years old, and approximately 58% are male. A potential reason is that middle-aged men are generally more likely to work outdoors [55, 56, 57].

We use an OLS model to compare how normalized heat-related claim rates change with temperature in majority cropland ZIP codes versus majority built-up ZIP codes. The OLS model shows a strong fit ( $R^2=0.978$ ,  $F(3, 110)=1403$ ) and heteroscedasticity-robust standard errors (HC3) as listed in Table S2. Temperature has a strong positive association with heat-related claim rates in both land cover types. However, in majority built-up areas, heat-related claim rates increase by about 20.6% for every 1°C ( $\beta = 0.187$ ,  $SE = 0.003$ ,  $p < 0.001$ ), while claim rates increase more strongly by 24.4% per 1°C in cropland-dominant areas ( $\beta = 0.0307$ ,  $SE = 0.008$ ,  $p < 0.001$ ). Residual diagnostics suggest non-normality in the residuals (Jarque-Bera  $p < 0.001$ ), with negative skewness (1.04) and high kurtosis (8.13), indicating a left-skewed and heavy-tailed residual distribution and the probable presence of outliers. The full OLS model summary is available in the supplemental information (Table S2).

## Heat-Related Claims Increase with High Temperatures and Heat Wave Events

We identified a significant, though weak-to-moderate, positive monotonic relationship between ZIP code average maximum daily air temperature and the corresponding monthly rates of heat-related illnesses in Fig. 2c (Spearman Rank Corr.: 0.183,  $p < 0.001$ ). This weak correlation suggests that claim rates are likely confounded by variables other than maximum daily temperature. For additional verification of the suitability of claims data to monitor heat-related illness rates, we analyzed claims during the 2017 Central Valley Heat Wave, which occurred from June 17-27 [52]. We observed a tenfold increase in heat-related inpatient and outpatient claims compared to the previous 11 days, an

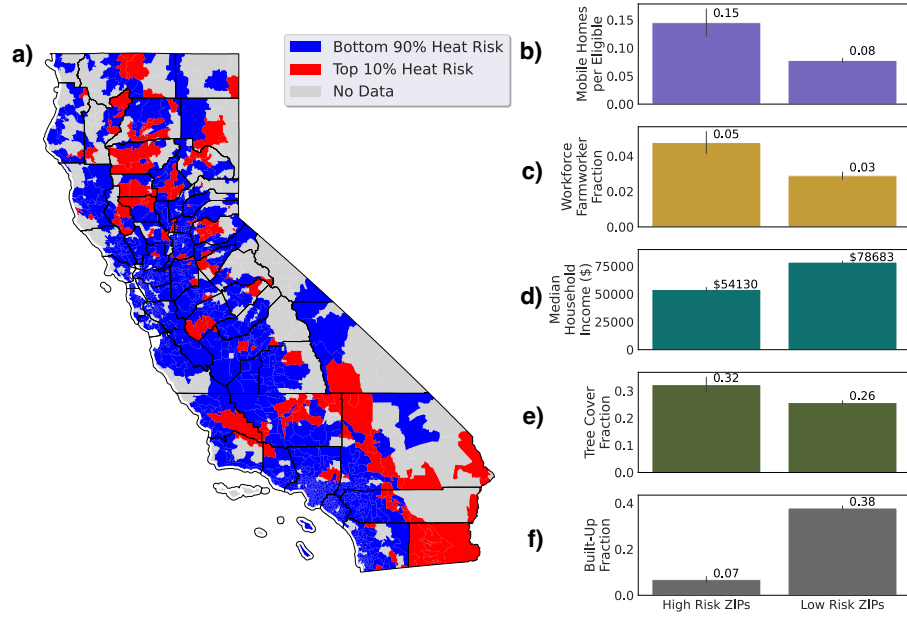


Figure 3: A) The top 10% highest heat-risk ZIP codes versus the remaining 90% of ZIP codes, based on BYM2-smoothed heat claim rates. Grey areas indicate ZIPs with no calculated rates, due to insufficient data on the Medicaid eligible population [17]. Bar plots illustrate the difference in selected demographics in high- and low-risk areas: B) the farmworker fraction of the workforce [19], C) the rate of mobile home spaces per Medicaid eligible [20], D) median household income [19], E) ZIP code tree cover fraction [21], and F) ZIP code built-up fraction [21]. Error bars on plots represent the standard error associated with the distributions of the five analyzed covariates.

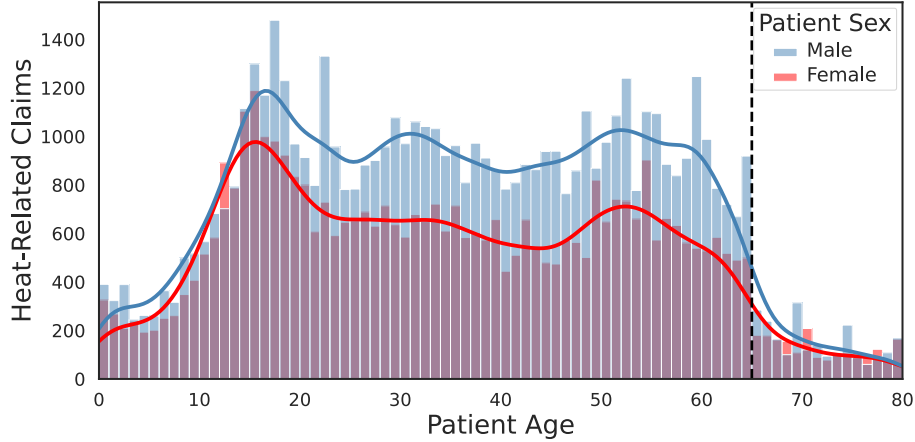


Figure 4: Histogram of heat-related inpatient and outpatient Medicaid claims in California between 2011-2019 by patient sex and age. The dashed black line indicates age 65, when many patients switch to Medicare as their primary provider.

increase from 111 to 1,143 claims (see Fig. S2). The highest claim counts occurred in ZIP codes in Bakersfield (93307, 93304) and Fresno (93728, 93726, 93706).

### Vulnerability Indices May Not Identify High-heat Risk Communities

The high spatial variability of heat-related claim rates in Fig. 2b poses a policy challenge. Measurably reducing heat risk with limited public resources requires decision-makers to identify and prioritize high-heat risk communities in programs like EHCRP. One common option for identifying high-risk communities is through composite vulnerability indices such as CalEnviroScreen, raising the question of whether high-risk communities are correctly identified through these kinds of algorithms.

In Fig. 6, we compare the CDC Heat Health Index (a), the CalEnviroScreen Population Characteristics Score (b) and the Heat Health Action Index from the California Heat Assessment Tool (c) against the annual heat-related claims at the ZIP code level. Only the CDC Heat Health Index increases with claims count ( $\rho=0.393$ ,  $p < 0.001$ ). This index differs from the others, because it includes historical heat health events, health sensitivities to heat, demographic sensitivities such as age and income, and natural and built environment conditions such as mobile home occupancy and air quality [27].

In contrast, we find a weak but significant negative rank correlation between ZIP code claim rates and the CHAT Heat Health Action Index ( $\rho=-0.099$ ,  $p <$

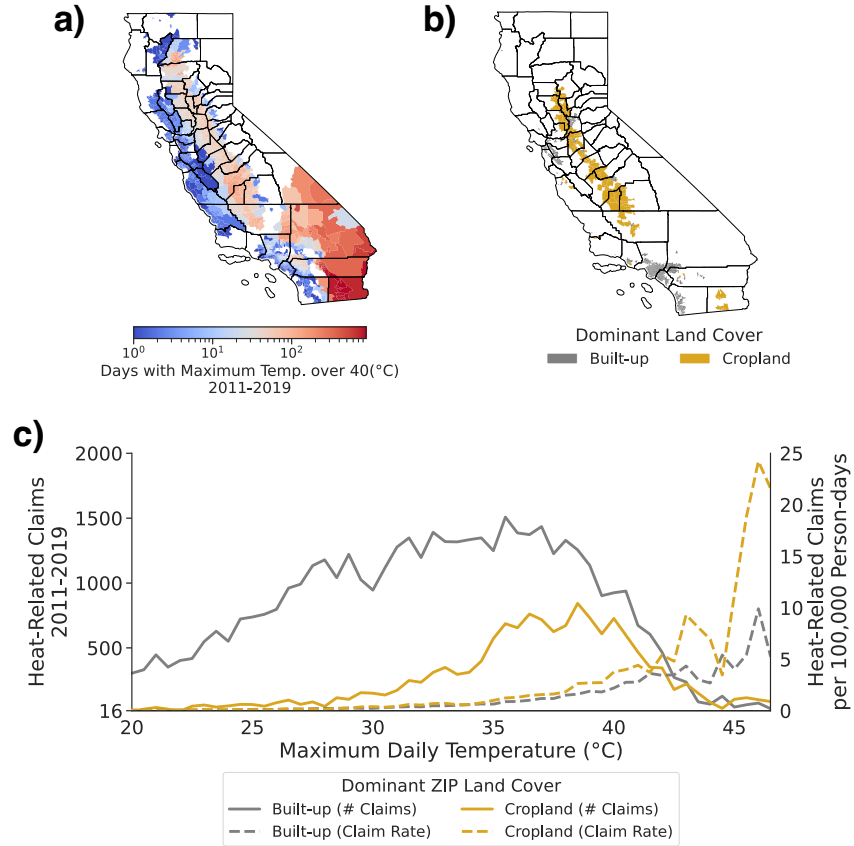


Figure 5: A comparison of heat-related Medicaid claims in ZIP codes with majority built-up and majority cropland land cover based on normalized heat exposure. A) The number of days with a maximum temperature over 40°C. B) ZIP codes included in the analysis had either built-up or cropland as their majority land cover. C) Daily temperature in each ZIP was rounded to the nearest 0.5°C, then the total number of days and claims per dominant land cover type were aggregated into these 0.5°C buckets. The total number of claims per dominant land cover were then normalized by the total person-days of exposure for Medicaid eligibles to calculate heat-related claims per 100,000 person-days of exposure by maximum daily temperature and dominant land cover class.

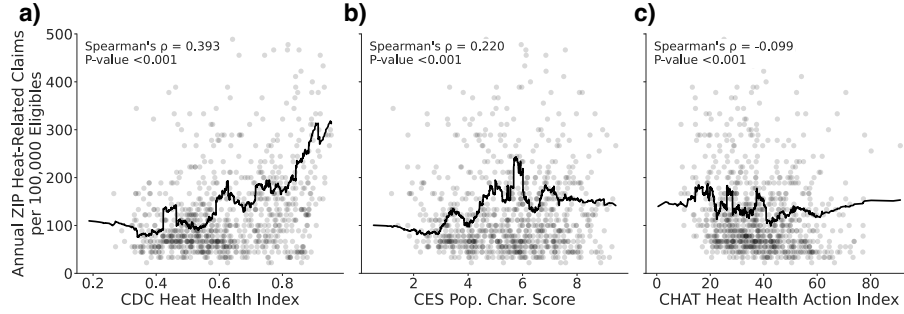


Figure 6: Comparison of smoothed heat-related Medicaid claim rates with three indices: A) the Center for Disease Control Heat Health Index, B) The CalEnviroScreen Population Characteristics Score, and C) the California Heat Assessment Tool Heat Health Action Index. Spearman rank correlation test statistics are displayed on each plot. Lines represent moving averages with a window of 100 points.

0.001). The CHAT Heat Health Action index also uses information on demographic sensitivities, health sensitivities, and the natural and built environment, but does not incorporate any data on prior heat-related illnesses [26]. For comparison, the CDC Heat Health Action Index incorporates emergency room visit data. While not a heat health index, we also compare smoothed ZIP code heat-related claim rates against the Population Characteristics Score of CalEnviroScreen, which is intended to represent the vulnerability of communities to pollution, finding a very weak positive monotonic agreement ( $\rho=0.220$ ,  $p < 0.001$ ).

## Discussion

California, like many other regions in the United States and worldwide, faces rapidly increasing heat risk [58, 59]. Tackling this challenge effectively starts with an understanding of who is most at risk. Much of the existing literature on heat-related illness emphasizes population-wide outcomes—such as total mortality [60] or emergency department visits [6] and focuses predominately on metropolitan areas [60, 61]. In contrast, our work shows that rural ZIP codes face disproportionately high heat risk, even after correction for noise stemming from low population (see Fig. 2b and 3a).

Imperial County alone contains 6 of the 10 top ZIP codes based on smoothed heat-related claim rates for ZIP codes with at least 1,000 mean Medicaid eligibles. Several of the other rural ZIP codes with high heat-related claim rates are in Northern California within Trinity (95595), Humboldt (95511), Siskiyou (96064), and Colusa (95955) counties, despite these areas facing much lower average maximum daily temperatures than Southern California. The heightened risk may be due to high rates of outdoor workers in the regional forestry

industry of Northern California [62, 19], or due to comorbidities associated with exposure to both extreme heat and smoke from wildfires [63].

In Fig. 3, we show that the top 10% of ZIP codes by smoothed heat-related claim rates had, on average, a \$24,553 lower median household income and a 86.5% higher rate of mobile home spaces per Medicaid eligible than other ZIP codes (0.145 versus 0.0778 spaces per eligible). Low income [29, 45, 64], poor housing quality or lack of housing [65, 66, 67] and living in mobile homes [65] have previously been associated with high exposure to heat and high rates of heat-related illness, partly because of poor insulation and high cooling costs for residents who are already lower-income [68]. Low income may also restrict air conditioning usage due to energy costs [46, 69].

ZIP codes with high heat risk have, on average, a 66% larger percentage of agricultural, forestry, and fishing workers in their workforce relative to ZIP codes with low heat risk (see Fig. 3b), suggesting that elevated risk could partly be driven by an outdoor workforce. California’s outdoor agricultural labor force is mostly comprised of young to middle-aged males, as detailed in the MICASA Survey [55] and reports from the California Research Bureau [56, 57]. High occupational heat exposure for outdoor workers and farm workers [70, 71, 72, 73, 74] and middle-aged men in general [75, 70] has been noted in prior studies. A 2015 study found that farm workers are approximately 35 times more likely to have a heat illness compared to other workers [70, 76].

California law requires “shade” when ambient outdoor temperatures exceed 80 degrees Fahrenheit, and workers must be “encouraged” to take a break if they express heat-illness symptoms [77]. However, the law does not specify how a farm operator should determine the outdoor temperature and does not reference heat-risk indexes, maps, or other heat-illness resources. Of the 1,200 workers surveyed in a 2021 UC Merced farmworker study, 20% reported lacking temperature monitoring on hot days, 15% reported insufficiently shaded breaks when it was over 80 degrees Fahrenheit, and 22% reported no heat illness monitoring when the temperature exceeded 95 degrees [76].

Similarly, a recent California OSHA audit in 2024 reported a 40% vacancy in enforcement staff, with only 22% of inspections being proactive, compared to a 42% national average [78]. California OSHA’s 2024 Annual Report acknowledges these problems, particularly compared to other outdoor workplaces. The agency did not reach their fiscal year 2024 goal regarding issuing at least one percent more “serious” agricultural outdoor workplace citations compared to the prior fiscal year. They also did not succeed in achieving a one percent increase in the number of agricultural employees “removed from serious hazards as a result of these inspections.” Both of these metrics were achieved for outdoor construction workers [79].

In addition to better protecting workers in agricultural workplaces, our work suggests that state funding aimed at reducing heat risk is not optimally distributed. In the EHCRP program and other state initiatives, the designation as a disadvantaged community directs funding, as exemplified by AB 1550, which mandates that 25% of California’s Greenhouse Gas Reduction Fund is allocated to these communities [80]. SB 535 “Land Use: General Plan: Safety

Element" identifies disadvantaged communities as the top 25% Census Tracts in the CalEnviroScreen index [25], which accounts for exposure to pollution, but not to heat risk. It is therefore not necessarily surprising that the CalEnviroScreen Population Characteristics Score does not correlate with high heat-related claims (Fig. 7b).

The generic view of vulnerability as unrelated to a particular stressor embedded in the construction of CalEnviroScreen index has been criticized extensively in the scientific literature [81, 82, 83]. These concerns have not yet translated to a change in policy, partly because the definition of disadvantaged communities is codified in laws such as SB 535 "Land Use: General Plan: Safety Element", which identifies disadvantaged communities as the top 25% Census Tracts in the CalEnviroScreen index [25]. The designation as a disadvantaged community then directs funding, such as with AB 1550, which mandates that 25% of California's Greenhouse Gas Reduction Fund is allocated to these communities [80]. Additionally, SB 1000 requires local governments to construct environmental justice plans, which may leverage CalEnviroScreen or other indices such as the CHAT Heat Health Action Index in local efforts to map risk [84].

Our results show that through this process, many high-risk communities are not currently prioritized in EHCRP funding. For example, in 2024 and 2025, EHCRP announced 46 recipients for grants to plan and implement efforts to reduce extreme heat risk across the State [85], funding scientifically-backed and worthwhile projects, such as solar shade structures, green infrastructure, and upgrades to cooling centers. Of the approximately \$32 million distributed by EHCRP, only one small planning grant of \$531,764 (1.7% of the total funds distributed) was awarded within Imperial County. This share corresponds roughly to what each of California's 58 counties would receive if EHCRP funds were distributed equally among them, despite the out-sized heat illness claim counts and rates within Imperial County.

We are not the first to point to Imperial County as a high heat-risk location. In 2025, the California Department of Public Health classified Imperial County as having the highest rate of occupational heat-related illness emergency department visits of all California counties between 2016-2023 [86]. This assessment in conjunction with our result highlights that considering medical claims data relevant to environmental exposures, such as heat illness claims, alongside with composite indices could elevate the relevancy of EHCRP and related programs for climate resiliency and risk reduction.

While medical claims data sheds light on community needs, it also has drawbacks. ICD codes are used for billing purposes, suggesting a patient's condition or disease requires further tests or investigation rather than indicating a final diagnosis [37]. They can also include errors resulting from medical coder inexperience or miscommunication between doctors and patients [37]. There may also be differences in coding practice across specific doctors, hospitals, or regions. Thus, these claims data cannot quantify the exact number of bona fide heat-related illnesses, but instead indicate relative risk as well as the public cost of heat-related illness claims across the state.

Another caveat specific to our study is that Medicaid only covers a subset of

Californians: low-income communities and those under age 65. At age 65, patients can enroll in Medicare, which then becomes the primary payer for medical services before Medicaid, so many claims from elderly patients are not captured here. It is also probable that many heat-related illnesses from farmworkers are not captured. A 2022 report from the Public Policy Institute of California estimated that 58.3% of farmworkers in California between 2010-2018 were undocumented; these workers use Medicaid at lower rates than documented workers [87]. Lastly, in order to focus our analysis on specific heat-related claims, we only include claims here with a diagnosis code from the ICD-10 "Effects of Heat and Light" category. We do not capture other heat-induced claims, such as those related to mental health conditions or kidney issues [6, 88].

The first step towards a more rigorous approach to heat health planning and intervention is an improved understanding of the spatially heterogeneous impacts of heat on health. Heat-illness data can help create a more accurate representation of heat risk. Which in turn, can support high-impact, efficient, and equitable planning and resource distribution for heat-risk minimization interventions. Claims data may also help to quantify the public and private monetary costs associated with heat-related illnesses and evaluate the financial impacts of interventions.

## Conclusions

Improving knowledge of the spatial and demographic distribution of heat-related health impacts is critical for informing efficient and equitable distribution of public resources, especially as risk is projected to increase with climate change. In this study, we map heat-related illness claims for low-income Californians on Medicaid between 2011-2019, highlighting specific communities at-risk, and demonstrating that common existing metrics for mapping heat risk, such as composite heat vulnerability indices, have limited capacity to identify high-risk communities. We also show disproportionate heat risk in rural areas, primarily those with lower median income, higher rates of mobile homes, and higher rates of farm workers in the workforce. We note that heat-related claim rates in cropland-dominant communities increase faster with temperature than in majority built-up communities, in agreement with many previous studies that highlight elevated risk for outdoor and farm workers.

## Data Availability

The raw data underlying this study are not publicly available due to the presence of personal health information. However, we share a limited version of the dataset that includes ZIP-code scale aggregations of heat-related illness claim counts across California between 2011-2019. We also share a table of heat-related illness counts aggregated by age and sex across California between 2011-2019. All counts and rates based on small cells of 1-15 have been redacted.



These data are available at Zenodo.org. We make python notebooks and data processing workflows available at <https://doi.org/10.71778/ywt4-2r72>.

## Acknowledgements

Data for this project were accessed using the Stanford Center for Population Health Sciences Data Core. The PHS Data Core is supported by a National Institutes of Health (NIH) National Center for Advancing Translational Science Clinical and Translational Science Award (UL1TR003142) and from internal Stanford funding. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results.

Avery Bick was funded by the Stanford Accelerator grant on “Addressing Blind Spots in Environmental Justice: The Central Valley as a Proof of Concept” and by a doctoral research grant 323294 from the Research Council of Norway.

## Supporting information

- SI.pdf: Supporting Information: Parameters of BYM2 claim count smoothing algorithm, OLS regression model results for temperature effects on claims by land use type, and effects of the 2017 Central Valley Heat Wave on temporal claim counts

## References

- [1] National Weather Service. Weather related fatality and injury statistics. Technical report, NOAA, 2024.
- [2] Ambarish Vaidyanathan. Heat-related deaths—united states, 2004–2018. *MMWR. Morbidity and mortality weekly report*, 69, 2020.
- [3] Jeffrey T Howard, Nicole Androne, Karl C Alcover, and Alexis R Santos-Lozada. Trends of heat-related deaths in the us, 1999-2023. *JAMA*, 332(14):1203–1204, 2024.
- [4] Jeremy S Pal and Elfatih AB Eltahir. Future temperature in southwest asia projected to exceed a threshold for human adaptability. *Nature climate change*, 6(2):197–200, 2016.
- [5] Samuel Lüthi, Christopher Fairless, Erich M Fischer, Noah Scovronick, Ben Armstrong, Micheline De Sousa Zanotti Stagliorio Coelho, Yue Leon Guo, Yuming Guo, Yasushi Honda, Veronika Huber, et al. Rapid increase in the risk of heat-related mortality. *Nature communications*, 14(1):4894, 2023.
- [6] Carlos F Gould, Sam Heft-Neal, Alexandra K Heaney, Eran Bendavid, Christopher W Callahan, Mathew V Kiang, Josh Graff Zivin, and Marshall Burke. Temperature extremes impact mortality and morbidity differently. *Science Advances*, 2025.
- [7] Kristen Guirguis, Alexander Gershunov, Alexander Tardy, and Rupa Basu. The impact of recent heat waves on human health in california. *Journal of Applied Meteorology and Climatology*, 53(1):3–19, January 2014.
- [8] Maria Stoecklin-Marois, Tamara Hennessy-Burt, Diane Mitchell, and Marc Schenker. Heat-related illness knowledge and practices among california hired farm workers in the micasa study. *Industrial health*, 51(1):47–55, 2013.
- [9] Rupa Basu and Jonathan M Samet. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiologic reviews*, 24(2):190–202, 2002.
- [10] Cecilia Sorensen and Jeremy Hess. Treatment and prevention of heat-related illness. *New England Journal of Medicine*, 387(15):1404–1413, 2022.
- [11] Andrea W. Harris and Jennifer Albrecht. Association between homelessness and temperature-related health emergencies in the united states. *Weather, Climate, and Society*, 16(4):633–638, October 2024.
- [12] Amruta Nori-Sarma, Chad Milando, Kate R. Weinberger, Jeremy J. Hess, Nicole A. Errett, and Gregory A. Wellenius. Association between the 2021 heat wave in portland, oregon, and seattle, washington, and emergency department visits. *JAMA*, 328(23):2360, December 2022.

- [13] Luz Rivas. Ca ab2238 (2021-2022).
- [14] Michael T. Schmeltz, Jason A. Smith, Isabella Olmos, and Erin Quintero. Extreme heat governance: A critical analysis of heat action plans in california. *American Journal of Public Health*, 113(1):15–19, 2023. PMID: 36516387.
- [15] Stanford Center for Population Health Sciences. Medicaid 100% [2011-2019] inpatient dataset, 2025.
- [16] Stanford Center for Population Health Sciences. Medicaid 100% [2011-2019] other services dataset, 2025.
- [17] Department of Health Care Services State of California. Medi-cal certified eligibles by month of eligibility, zip code, and sex. [data.chhs.ca.gov](https://data.chhs.ca.gov), 2025.
- [18] Department of Finance State of California. E-4 population estimates for cities, counties, and the state, 2011-2020, with 2010 and 2020 census benchmarks. <https://dof.ca.gov/forecasting/demographics/estimates/e-4-population-estimates-for-cities-counties-and-the-state-2011-2020-with-2010-and-2020-census-benchmark-new/>, May 2025.
- [19] Steven Manson, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles. National historical geographic information system: Version 19.0, 2024.
- [20] California Department of Housing and Community Development. Mobile-homes and rv parks database, 2024.
- [21] Daniele Zanaga, Ruben Van De Kerchove, Wanda De Keersmaecker, Niels Souverijns, Carsten Brockmann, Ralf Quast, Jan Wevers, Alex Grosu, Audrey Paccini, Sylvain Vergnaud, Oliver Cartus, Maurizio Santoro, Steffen Fritz, Ivelina Georgieva, Myroslava Lesiv, Sarah Carter, Martin Herold, Linlin Li, Nandin-Erdene Tsendbazar, Fabrizio Ramoino, and Olivier Arino. Esa worldcover 10 m 2020 v100, October 2021.
- [22] Andrea Riebler, Sigrunn H Sørbye, Daniel Simpson, and Håvard Rue. An intuitive bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4):1145–1165, August 2016.
- [23] Extreme heat and community resilience program round 1 final program guidelines: Fy 2023 – 2024. Technical report, California Governor’s Office of Planning and Research, 2024.
- [24] Governor’s Office of Land Use, Climate Innovation (LCI) Extreme Heat, and Community Resilience Program (EHCRP), 2025.
- [25] Calenviroscreen 4.0. <https://oehha.ca.gov/calenviroscreen/maps-data>, April 2025.

- 
- [26] California Natural Resources Agency, Four Twenty Seven, Argos Analytics, Habitat Seven, and Public Health Institute. Chat - california heat assessment tool, March 2025.
- [27] Centers for Disease Control and Prevention. Heat and health index. <https://ephtracking.cdc.gov/Applications/heatTracker/>, May 2025.
- [28] Laura A Bakkensen, Cate Fox-Lent, Laura K Read, and Igor Linkov. Validating resilience and vulnerability indices in the context of natural disasters. *Risk analysis*, 37(5):982–1004, 2017.
- [29] Jihoon Jung, Christopher K. Uejio, Kristina W. Kintziger, Chris Duclos, Keshia Reid, Melissa Jordan, and June T. Spector. Heat illness data strengthens vulnerability maps. *BMC Public Health*, 21(1):1999, December 2021.
- [30] Frank C Curriero, Karlyn S Heiner, Jonathan M Samet, Scott L Zeger, Lisa Strug, and Jonathan A Patz. Temperature and mortality in 11 cities of the eastern united states. *American journal of epidemiology*, 155(1):80–87, 2002.
- [31] Sharon L Harlan, Gerardo Chowell, Shuo Yang, Diana B Petitti, Emmanuel J Morales Butler, Benjamin L Ruddell, and Darren M Ruddell. Heat-related deaths in hot cities: estimates of human tolerance to high temperature thresholds. *International journal of environmental research and public health*, 11(3):3304–3326, 2014.
- [32] Hana Putnam, David M Hondula, Aleš Urban, Vjollca Berisha, Paul Iñiguez, and Matthew Roach. It’s not the heat, it’s the vulnerability: attribution of the 2016 spike in heat-associated deaths in maricopa county, arizona. *Environmental research letters*, 13(9):094022, 2018.
- [33] California Office of Environmental Health Hazard Assessment. Cal-heatscore. Web application, 2025. Accessed: 2025-11-14.
- [34] National Oceanic and Atmospheric Administration National Weather Service. Heatrisk. Web application, 2025. Accessed: 2025-11-14.
- [35] Centers for Disease Control and Prevention. Heat health tracker. <https://ephtracking.cdc.gov/Applications/heatTracker/>, 2025.
- [36] Stanford Center for Population Health Sciences. Medicaid 100% [2011-2019] personal summary and demographic eligibility dataset, 2025.
- [37] Kimberly J. O’Malley, Karon F. Cook, Matt D. Price, Kimberly Raiford Wildes, John F. Hurdle, and Carol M. Ashton. Measuring diagnoses: Icd code accuracy. *Health Services Research*, 40(5p2):1620–1639, October 2005.
- [38] Chronic Conditions Warehouse Virtual Research Data Center. T-msis analytic files (taf) research identifiable files (rifs) user guide version 1.10, December 2023.

- 
- [39] Redivis. Redivis, 2024.
- [40] California Department of Health Care Services. Medi-cal certified eligible counts, by month of eligibility, zip code, and sex, 2024.
- [41] Lance A. Waller and Carol A. Gotway. *Applied spatial statistics for public health data*. Wiley series in probability and statistics. John Wiley Sons, Hoboken, N.J, 2004.
- [42] Julian Besag, Jeremy York, and Annie Molli. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1):1–20, March 1991.
- [43] M.M. Thornton, R. Shrestha, Y. Wei, P.E. Thornton, and S-C. Kao. Daymet: Daily surface weather data on a 1-km grid for north america, version 4 r1, 2022.
- [44] Robert I. McDonald, Tanushree Biswas, T. C. Chakraborty, Timm Kroeger, Susan C. Cook-Patton, and Joseph E. Fargione. Current inequality and future potential of us urban tree cover for reducing heat-related health impacts. *npj Urban Sustainability*, 4(1):18, April 2024.
- [45] Robert I. McDonald, Tanushree Biswas, Cedilla Sachar, Ian Housman, Timothy M. Boucher, Deborah Balk, David Nowak, Erica Spotswood, Charlotte K. Stanley, and Stefan Leyk. The tree cover and temperature disparity in us urbanized areas: Quantifying the association with income across 5,723 communities. *PLOS ONE*, 16(4):e0249715, April 2021.
- [46] Alina M. McIntyre, Madeleine K. Scammell, Maria Pilar Botana Martinez, Leila Heidari, Abgel Negassa, Roseann Bongiovanni, and M. Patricia Fabian. Facilitators and barriers for keeping cool in an urban heat island: Perspectives from residents of an environmental justice community. *Environmental Justice*, page env.2022.0019, July 2022.
- [47] Matthew Perry and Contributors. rasterstats, 2024.
- [48] Stanford Center for Population Health Sciences. Us zip codes to census tracts crosswalk, December 2019.
- [49] Ron Wilson and Alexander Din. Understanding and enhancing the u.s. department of housing and urban development’s zip code crosswalk files. *Cityscape*, 20(2):277–294, 2018.
- [50] SciPy 1.0 Contributors. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, 17(3):261–272, March 2020.
- [51] Chico Department of Geography California State University, Planning, and Geographic Information Center. The central valley historic mapping project, 2003.

- 
- [52] Industrial Economics Incorporated and California Department of Insurance. *Impacts of extreme heat to California's people, infrastructure, and economy*. 2024.
- [53] James G MacKinnon and Halbert White. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3):305–325, 1985.
- [54] Skipper Seabold and Josef Perktold. Statsmodels: Econometric and statistical modeling with python. page 92–96, Austin, Texas, 2010.
- [55] M. T. Stoecklin-Marois, T. E. Hennessy-Burt, and M. B. Schenker. Engaging a hard-to-reach population in research: Sampling and recruitment of hired farm workers in the micasa study. *Journal of Agricultural Safety and Health*, 17(4):291–302, 2011.
- [56] Alicia Bugarin and Elias S. Lopez. Farmworkers in california. Technical Report 4950, California Research Bureau, California State Library, Sacramento, CA, 1998.
- [57] Patrick Rogers and Matthew K. Buttice. Farmworkers in california: A brief introduction. Technical report, California Research Bureau Short Subjects, October 2013.
- [58] Alexander Gershunov and Kristen Guirguis. California heat waves in the present and future. *Geophysical Research Letters*, 39(18), 2012.
- [59] Scott C Sheridan, Cameron C Lee, Michael J Allen, and Laurence S Kalkstein. Future heat vulnerability in california, part i: projecting future weather types and heat events. *Climatic Change*, 115(2):291–309, 2012.
- [60] G. Brooke Anderson and Michelle L. Bell. Heat waves in the united states: Mortality risk during heat waves and effect modification by heat wave characteristics in 43 u.s. communities. *Environmental Health Perspectives*, 119(2):210–218, February 2011.
- [61] Scott C Sheridan, Michael J Allen, Cameron C Lee, and Laurence S Kalkstein. Future heat vulnerability in california, part ii: projecting future heat-related mortality. *Climatic Change*, 115(2):311–326, 2012.
- [62] Amy Heinzerling, Rebecca L. Laws, Matt Frederick, Rebecca Jackson, Gayle Windham, Barbara Materna, and Robert Harrison. Risk factors for occupational heat-related illness among california workers, 2000–2017. *American Journal of Industrial Medicine*, 63(12):1145–1154, December 2020.
- [63] Caitlin G Jones Ngo, Rebecca J Schmidt, Jason Vargo, Mohammad Al Hamdan, Erwan Monier, and Kathryn C Conlon. Joint effects of wildfire smoke and extreme heat on hospitalizations in california, 2011–2020. *ISEE Conference Abstracts*, 2024(1):isee.2024.1634, August 2024.

- 
- [64] Carina J Gronlund. Racial and socioeconomic disparities in heat-related health effects and their mechanisms: a review. *Current epidemiology reports*, 1:165–173, 2014.
- [65] Patricia Solís, Katsiaryna Varfalameyeva, and Carlos Aguiar Hernandez. Heat resilience among mobile home owners in arizona: towards a multi-scale approach to address spatial incongruence and accountable decision making. *GeoJournal*, 88(5):4991–5012, 2023.
- [66] C.J. Gabbe, Jamie Suki Chang, Morayo Kamson, and Euichan Seo. Reducing heat risk for people experiencing unsheltered homelessness. *International Journal of Disaster Risk Reduction*, 96:103904, October 2023.
- [67] Margaret V. Du Bray, Rhian Stotts, Richard Southee, and Amber Wutich. Beyond extreme: Heat emergency and water insecurity for people experiencing houselessness in phoenix, arizona, usa during and after the heatwave of 2023. *Human Ecology*, 51(5):799–808, October 2023.
- [68] Philip Agee, Leila Nikdel, Andrew McCoy, Simin Kianpour Rad, and Xinghua Gao. Manufactured housing: Energy burden outcomes from measured and simulated building performance data. *Energy Policy*, 186:113985, March 2024.
- [69] Lora A. Phillips, Patricia Solís, Chuyuan Wang, Katsiaryna Varfalameyeva, and Janice Burnett. Engaged convergence research: An exploratory approach to heat resilience in mobile homes. *The Professional Geographer*, 73(4):619–631, October 2021.
- [70] Diane M. Gubernot, G. Brooke Anderson, and Katherine L. Hunting. Characterizing occupational heat-related mortality in the united states, 2000–2010: An analysis using the census of fatal occupational injuries database. *American Journal of Industrial Medicine*, 58(2):203–211, February 2015.
- [71] S. Claire Slesinski, Franziska Matthies-Wiesler, Susanne Breitner, Geronimo Gussmann, and Alexandra Schneider. Social inequalities in exposure to heat stress and related adaptive capacity: A systematic review. *Environmental Research Letters*, February 2025.
- [72] Maria C Mirabelli and David B Richardson. Heat-related fatalities in north carolina. *American Journal of Public Health*, 95(4):635–637, 2005.
- [73] Quirina M Vallejos, Sara A Quandt, Joseph G Grzywacz, Scott Isom, Haiying Chen, Leonardo Galván, Lara Whalley, Arjun B Chatterjee, and Thomas A Arcury. Migrant farmworkers’ housing conditions across an agricultural season in north carolina. *American Journal of Industrial Medicine*, 54(7):533–544, 2011.

- 
- [74] Centers for Disease Control and Prevention. Heat-related deaths among crop workers—united states, 1992-2006. *JAMA*, 300(9):1017, September 2008.
- [75] Diane M Gubernot, G Brooke Anderson, and Katherine L Hunting. The epidemiology of occupational heat exposure in the united states: a review of the literature and assessment of research needs in a changing climate. *International journal of biometeorology*, 58:1779–1788, 2014.
- [76] Community, Labor Center: UC Merced, and California Department of Public Health. *Farmworker Health in California*. 2022.
- [77] California Department of Industrial Relations. California code of regulations, title 8, section 3395: Heat illness prevention in outdoor places of employment. <https://www.dir.ca.gov/title8/3395.html>. Accessed: November 4, 2025.
- [78] The division of occupational safety and health process deficiencies and staffing shortages limit its ability to protect workers. Audit Report 2024-115, California State Auditor, Sacramento, CA, 2025. Available online at [www.auditor.ca.gov](http://www.auditor.ca.gov).
- [79] State osha annual report (soar) ffy 2024. Annual report, State of California Department of Industrial Relations, California, March 2025. Evaluation Period: October 1, 2023 - September 30, 2024.
- [80] Jimmy Gomez. Ca assembly bill 1550, 2016.
- [81] Alexander Fekete. Social vulnerability (re-) assessment in context to natural hazards: Review of the usefulness of the spatial indicator approach and investigations of validation demands. *International Journal of Disaster Risk Science*, 10(2):220–232, 2019.
- [82] Shira Mitchell, Eric Potash, Solon Barocas, Alexander D’Amour, and Kristian Lum. Algorithmic fairness: Choices, assumptions, and definitions. *Annual review of statistics and its application*, 8(1):141–163, 2021.
- [83] Irene Y Chen, Emma Pierson, Sherri Rose, Shalmali Joshi, Kadija Feryman, and Marzyeh Ghassemi. Ethical machine learning in healthcare. *Annual review of biomedical data science*, 4(1):123–144, 2021.
- [84] Connie Leyva. Senate bill 1000, 2016.
- [85] Extreme Heat and Community Resilience Program. Extreme heat and community resilience program round 1 grant awards, 2025.
- [86] Stella Beckman, Debbie Shrem, Cassie Marquez, Joyce Veal, and Robert Harrison. Occupational heat-related illness emergency department visits, california, 2016-2023. Technical report, California Department of Public Health, April 2025.



## REFERENCES

---

- [87] Paulette Cha and Joy Collins. *Health Care Access among California's Farmworkers*. April 2022.
- [88] Kristen Hansen, Armin Schwartzman, Lara Schwarz, Anais Teyton, Rupa Basu, and Tarik Benmarhnia. The spatial distribution of heat related hospitalizations and classification of the most dangerous heat events in california at a small-scale level. *Environmental Research*, 261:119667, November 2024.