ClimaWATCH

Mathematica

Overview

Mathematica's ClimaWATCH (Climate and Weather Analytics, Trends, and Community Health) tool helps officials conduct heat vulnerability assessments by exploring how heatwaves have impacted health and magnified inequity across the country. Dynamic data summaries and maps—provided by county, demographic group, care setting, and diagnosis—clarify where heatwaves have concentrated, how socio-environmental factors differ in counties with and without heatwaves, and which communities have faced higher rates of heat-related health issues among Medicaid beneficiaries—one of the most vulnerable groups across the country.

Data and Methods

Environmental exposure

We obtained daily temperature and dew point data from the PRISM Climate Group (2020) for the 48 contiguous U.S. states and from the National Centers for Environmental Information (2021) for Alaska and Hawaii. For the contiguous U.S., we applied a population-weighted approach to 4 km-by-4km grid-level data to calculate daily county-level temperature estimates during the warm season of May 1 to September 30 between 1981 and 2020 (Anderson et al 2013; Gronlund et al 2014). For Alaska and Hawaii, we averaged data from individual weather stations within a county to calculate county-level estimates.[1] Since we relied on observed data for Alaska and Hawaii (versus modeled data, which was used for the lower 48 states), we needed to impute missing temperature values in the time series from a single weather station. To do so, we used linear interpolation of available temperature observations.

We chose apparent temperature (AT) as the primary heat exposure metric to account for the joint effects of temperature and humidity on human health. We calculated AT (in °C), as follows, based on the methods of Kalkstein and Valimont (1986):

$$AT = -2.653 + (0.994 \times \text{mean temperature}) + 0.0153 \times (\text{dew point}^2)$$

Using this measure, we identified heatwaves for each county and year. There is currently no consensus definition for a heatwave, which can be defined as two or more consecutive days with temperatures above a physiologically based absolute threshold or a location-based relative threshold (Robinson 2001). We examined the following two definitions:

- Method 1: Two or more consecutive days when the daily mean AT was above 90°F (32.2°C) and exceeded the 95th percentile of historical temperatures (based on daily average ATs during the warm seasons of 1981-2020).
- Method 2: Two or more consecutive days when the daily mean AT exceeded the 95th percentile of historical temperatures (based on daily average ATs during the warm seasons of 1981-2020).

In the Exposure panel of the tool, the number of counties affected by a heatwave may exceed the number of heatwaves because a single heatwave can span multiple counties. Likewise, in the Sortable table, a heatwave that spans multiple counties will be reflected in each affected county's tally, and so the sum of the heatwave counts across all counties may exceed the state-level heatwave count.

[1] In 2019, Alaska County FIPS 02261 (Valdez-Cordova) split into 02063 (Chugach) and 02066 (Copper River). To maintain consistency across years in ClimaWATCH, we remapped FIPS 02063 and 02066 to FIPS 02261.

Community features and vulnerability

To examine how heat-related excess health issues distribute by level of urbanicity, we used data from the 2017-2018 Area Health Resource File (AHRF) provided by the U.S. Department of Health and Human Services' Health Resources and Services Administration. Specifically, we classified counties using the AHRF Core Based Statistical Area Indicator Code, as follows:

- **Metropolitan:** Counties in Metropolitan Statistical Areas have at least one urbanized area of 50,000 people or more and adjacent territory with a high degree of social and economic integration with the core (as measured by commuting ties).
- **Micropolitan:** Micropolitan Statistical Areas have at least one urban cluster of at least 10,000 but less than 50,000 people and adjacent territory with a high degree of social and economic integration with the core (as measured by commuting ties).
- **Other:** Non-Statistical Areas are counties that do not meet either of the criteria above.

We classified counties based on the percentage of the resident population who are minorities (all persons except non-Hispanic Whites), using data from the 2014-2018 American Community Survey (ACS), in line with the approach used in the CDC 2018 SVI. We also characterized counties' social vulnerability using the 2018 Social vulnerability index (SVI) from the Centers for Disease Control and Preventions' (CDC) Agency for Toxic Substances and Disease Registry. The SVI uses 15 U.S. Census variables to identify communities that may need support before, during, or after natural disasters or public health crises. We classified counties based on quintiles of their SVI score for each of the following four themes (and corresponding U.S. Census variables):

- **Minority or Language vulnerability (**Hispanic or Latino [of any race]; Black and African American, Not Hispanic or Latino; American Indian and Alaska Native, Not Hispanic or Latino; Asian, Not Hispanic or Latino; Native Hawaiian and Other Pacific Islander, Not Hispanic or Latino; Two or More Races, Not Hispanic or Latino; Other Races, Not Hispanic or Latino)
- **Socioeconomic vulnerability** (below 150% poverty, unemployed, housing cost burden, no high school diploma, no health insurance)
- **Household composition or Disability vulnerability** (aged 65 or older, aged 17 or younger, civilian with a disability, single-parent households, English language proficiency)
- **Housing or Transportation vulnerability** (multi-unit structures, mobile homes, crowding, no vehicle, group quarters)

To characterize counties' vulnerability related to social and environmental conditions, we used the following data sources and metrics:

- **Heat Vulnerability Index** (HVI) is a county-level population-weighted average of the vulnerability index calculated by Manware et al. (2022), which includes measures of underlying health condition prevalence, demographic composition, land cover, temperature exposure, and housing. Data comes from the CDC, ACS, and National Aeronautics and Space Administration. Higher values of the index score (which ranges from 5 to 30) represent higher heat vulnerability.
- **National Risk Index** (NRI), modeled by the Federal Emergency Management Agency, uses data on Expected Annual Loss due to natural hazards, Social Vulnerability, and Community Resilience to create an index that captures the potential for negative impacts due to a natural hazard. The NRI ***Risk Index Score*** (which is on a 0-100 scale) considers the composite risk of 18 types of hazards occurring (Avalanche, Coastal Flooding, Cold Wave, Drought, Earthquake,

Hail, Heat Wave, Hurricane, Ice Storm, Landslide, Lightning, Riverine Flooding, Strong Wind, Tornado, Tsunami, Volcanic Activity, Wildfire, and Winter Weather). The ***Risk Index Rating*** is a qualitative rating (with five categories ranging from "Very Low" to "Very High") that describes a county's relative risk index score compared to all other counties.

- **Impervious Surface Cover** is a population-weighted average of the percentage of urban area within a county that is covered by impervious surfaces (that is, hard surfaces, such as pavement, that absorb and store heat in areas where people live). Estimates for the continental U.S. come from the Environmental Protection Agency, based on 2016 data from the National Land Cover Database (NLCD), while those for Hawaii were derived from the NLCD directly (from 2001). Similarly, data for Alaska was obtained from NLCD directly using 2011 data and updating any change to impervious surfaces with their data from 2016 data.
- **Tree Inequity Score** is based on the Tree Equity Score (TES) developed by American Forests using data from the NLCD, ACS, and USDA Forest Service. The TES identifies counties that (1) are not meeting tree coverage targets (established based on counties' natural biome type and population density), and (2) have high socioenvironmental vulnerability (based on demographics, socioeconomics, the prevalence of mental and physical health issues, and urban heat island severity). The measure inverts the TES (100 – TES), such that higher scores indicate more vulnerable communities with a wider gap between existing and target tree canopy cover.

Health outcomes

Prior to estimating heat-related health service use and spending, we reviewed the literature to identify conditions known to be caused or worsened by excessive heat. We identified four acute diagnosis categories: heat-related illnesses, electrolyte imbalance, acute renal failure, and acute myocardial infarction.[2] Table 1 displays International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) codes for these diagnosis categories. We also aggregated these codes into a broad acute composite diagnosis category.

Table 1. ICD-10-CM codes used to define diagnosis categories.

Diagnosis	ICD-10-CM code	Description
Heat related illnesses	T67.x	Effects of heat and light
	V93.2	Heat exposure on board watercraft
	X30	Exposure to excessive natural heat
	X32	Exposure to sunlight
Electrolyte imbalance	E87.x	Other disorders of fluid, electrolyte and acid-base balance
Acute renal failure	N17.x	Acute kidney failure
Acute myocardial infarction	I21.x	Acute myocardial infarction
	I22.x	Subsequent ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction
Acute composite	*Any of the above conditions*	

Medicaid data. We obtained Medicaid program data on health service use and spending from the Transformed Medicaid Statistical Information System Analytic Files (TAF) Research Identifiable Files. For the years 2016 through 2020, we retained claims in the Inpatient or Other Services files that had a diagnosis code (primary or otherwise) for any of the diagnoses shown in Table 1 (Morano and Watkins 2017) and a service date within one month of the warm season (that is, from April 1 through October 31).

We calculated two key outcomes of interest using the Medicaid data:

(1) **Health service use**, based on the number of Medicaid beneficiaries with a claim for the diagnoses of interest during the months of interest;

[2] The diagnosis categories are modified versions of categories used by Knowlton et al. (2009).

(2) **Health care spending** based on the expenditures reflected in such claims. The Medicaid data include fee-for-service (FFS) expenditures only, with no data available on managed care expenditures.

ClimaWATCH also displays the number of unique beneficiaries enrolled in Medicaid for the selected geography and year. This number includes those with a heat-related claim during the warm season as well as beneficiaries who were enrolled but did not have any claims for heat-related health service use during that time. The information is based on the TAF Annual Demographic and Eligibility file.

County-level case-crossover design to attribute health issues to heatwaves

To estimate health service use and spending *attributable* to excessive heat, we adopted a case-crossover study design applied at the county level. This design enables adjustment for known and unknown time-stable confounders and is therefore ideally suited to estimate health service use and spending due to heatwave exposure. Within this framework, we analyzed health outcomes during three periods of time: days during which a heatwave (HW) occurred, days within a 7-day buffer period after each heatwave (BF), and days during a reference period (RF) during which there were no heatwaves. To account for any natural time trend effects, we identified two reference periods for each heatwave – one before the heatwave and one after. Both reference periods were chosen to be as close in time to the heatwave as possible and were always within the same year and heat season as the heatwave.

- To identify *pre-heatwave* RF periods, we developed a sequential matching algorithm to identify a period that had the same length and a similar number of weekends and holidays (+/- 1 day) as the heatwave period, was free of excessive heat, and was not within a 14-day washout period of another heatwave.
- To identify *post-heatwave* RF periods, we adjusted the matching algorithm such that, when the algorithm failed to identify a matching period for a given heatwave, we lightened the matching requirements. Specifically, we allowed the reference period to be shorter than the heatwave (for heatwaves that were longer than ten days), we removed the requirement to include a similar number of weekend/holiday days as the heatwave, and we allowed a given reference period to be "reused" for another heatwave, if necessary.

For each of these three periods, we summed daily health service use and spending for all beneficiaries within a county (for reference period days, we averaged outcomes during the pre-heatwave and post-heatwave reference periods). Because these periods may have varied in length, we standardized our outcome metrics by calculating service use and spending *per day.*

Lastly, to determine the excesses health service use and spending attributable to heatwave exposure (that is, to parse out background, expected levels of health service use and spending from the whole), we compared outcomes that occurred during a heatwave or within a 7-day buffer period after the heatwave with outcomes that occurred during the non-heatwave reference period, as follows:

*Excess = (average use/spending per day during [(HW + BF) – RF] * number of heatwave days*

We standardized this measure by dividing by the number of unique beneficiaries with a claim during the period of interest and for the diagnosis of interest, yielding a measure of excess per beneficiary (in the units of number of beneficiaries or dollars spent). We also report this excess as a percent change from the background reference period (which was undefined if there was no service utilization or spending during the reference period).

To calculate state-level outcome and excess measures, we summed county-level daily health service use and spending for all counties across the state during each day of the three periods of interest. We also broke out summaries of Medicaid health service use and spending by:

- **Geography** U.S., state, or county.
- **Year.** 2016 through 2020.
- **Demographic group.**[3] Adults (including expansion Adults), Children (Medicaid and Children's Health Insurance Program), Blind and Disabled, Aged, and Pregnant.
- **Care setting.** Emergency department, in-patient hospital, and other (which includes care provided in any other facility or non-facility setting other than a long-term care facility).
- **Diagnosis category.** See Table 1 above.

Limitations

- Results may be somewhat sensitive to our definition of a heatwave and could vary with alternate characterizations (based on max, mean, or min temperature; relative threshold; or other criteria).
- Excess spending could not be calculated when all claims for a given year, county, diagnosis, and period were managed care claims (which did not include information on expenditures). To contextualize the excess spending measures reported, ClimaWATCH displays the proportion of Medicaid beneficiaries enrolled in managed care, which is based on the Centers for Medicare & Medicaid Services [CMS] Medicaid Managed Care Enrollment Report).
  - The proportion of beneficiaries enrolled in managed care differs from state to state and can vary greatly (from 9 to 99 percent).
  - Because the data we analyzed only include expenditures based on FFS claims, estimated spending for states with a large proportion of beneficiaries in managed care may be less reliable than estimated spending for states that primarily use FFS payment models.
- The quality of the Medicaid TAF data we analyzed can vary by state and year, as summarized in the DQ Atlas and TAF Data Quality Resources provided by CMS. Depending on the state and year selected (per this DQ Atlas link), the set of data quality issues that could affect ClimaWATCH summaries include, but are not limited, to the following:

  (1) Challenges linking claims (which contain service use and cost information) with beneficiary information (such as county of residence and eligibility group);
  (2) Poor reporting rates for the eligibility group code, which is used to define beneficiaries' demographic group;
  (3) Inconsistencies in the volume of claims reported in Inpatient or Other Services files, which may indicate issues with the comprehensiveness and accuracy of service use information;
  (4) Missing FFS payment information, which could lead to underestimating health care costs;
  (5) Missing or invalid diagnosis codes, which could lead to undercounting health care visits;
  (6) Missing or invalid procedure, place of service, or revenue codes, which could affect our ability to identify admissions to the emergency department service setting.

[3] Demographic groups are based on eligibility categories that define the Medicaid enrollment pathway. Categories are mutually exclusive, and we excluded the COVID group because most data we analyzed preceded the pandemic.

References

Anderson, G.B. and L.M. Bell (2011). Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities. *Environmental health perspectives*, 119(2), 210-218.

Gronlund, C.J., A. Zanobetti, J.D. Schwartz, G.A. Wellenius, M.S. O'Neill (2014). Heat, heat waves, and hospital admissions among the elderly in the United States, 1992–2006. *Environmental health perspectives*, 122(11), 1187-1192.

Kalkstein L.S. and K.M. Valimont 1986. An evaluation of summer discomfort in the United States using a relative climatological index. *Bull Am Meteorol Soc* 7:842–848.

Knowlton, K., M. Rotkin-Ellman, G. King, H.G. Margolis, D. Smith, G. Solomon, G., et al. (2009). The 2006 California heat wave: impacts on hospitalizations and emergency department visits. Environmental health perspectives, 117(1), 61-67.

Manware, M., R. Dubrow, D. Carrión, Y. Ma, and K. Chen (2022). Residential and race/ethnicity disparities in heat vulnerability in the United States. *GeoHealth*, 6, e2022GH000695.

Morano, L.H. and S. Watkins (2017). Evaluation of Diagnostic Codes in Morbidity and Mortality Data Sources for Heat-Related Illness Surveillance. Public Health Reports, 132(3), 326-335.

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, accessed December 2020.

Robinson P.J. 2001. On the definition of a heatwave. *J Appl Meteorol* 40:762–775.

Swanwick, R.H., Q.D. Read, S.M. Guinn, M.A. Williamson, K.L. Hondula, A.J. Elmore. Dasymetric population mapping based on US census data and 30-m gridded estimates of impervious surface. *Sci Data* 9, 523 (2022).