Technical Appendix

All code and data used in the model are publicly available on GitHub (github.com/sigalm/Climate-Health). This document is intended to complement the codebase by explaining assumptions, discussing modeling decisions, and providing information on data sources.

# Generating microdata

A microdata set from the 2021 American Community Survey was obtained through [IPUMS USA](https://usa.ipums.org/usa/) with select variables of interest, including county and household identifiers, and sociodemographic variables.

Current asthma prevalence data by age group and county was obtained from the [California Department of Public Health](https://www.cdph.ca.gov/Programs/CCDPHP/DEODC/EHIB/CPE/Pages/CaliforniaBreathingCountyAsthmaProfiles.aspx), sourced from the 2019-2020 California Health Interview Survey. Age- and county-specific asthma data for age groups 0-4 and 5-17 was largely unavailable. Missing data was imputed using the age-specific California average for age group 0-4 (91% missing), and age-specific population-weighted residual average for age group 5-17 (53% missing). Population weights were derived from the 2010 census, which was the most recent data available.

We recognize that asthma is not evenly distributed within counties. A major factor is proximity to sources of pollution, which are typically in or near lower-income neighborhoods. We therefore used data on poverty (household income as a percentage of the federal poverty line) to adjust the probability of having asthma for each individual in the dataset (Table ##). A binary asthma variable was created using these probabilities.

Table ##. Poverty and asthma prevalence

|  |  |  |
| --- | --- | --- |
| **Household income** | **Proportion of sample** | [**Risk ratio for asthma**](https://www.lung.org/research/trends-in-lung-disease/asthma-trends-brief/current-demographics#:~:text=In%202018%2C%20current%20asthma%20rates,to%20above%20the%20poverty%20threshold) |
| 200% or more of the poverty line | 0.75 | (reference) |
| 100% or more but less than 200% of the poverty line | 0.14 | 1.32 |
| Less than 100% of the poverty line | 0.11 | 1.62 |

Lastly, for individuals who were assigned as having asthma, we defined an asthma control variable by randomly sampling from the vector of asthma control states given the probabilities of having each level of control. These control probabilities were calculated during model calibration and are described elsewhere.

# Data sources

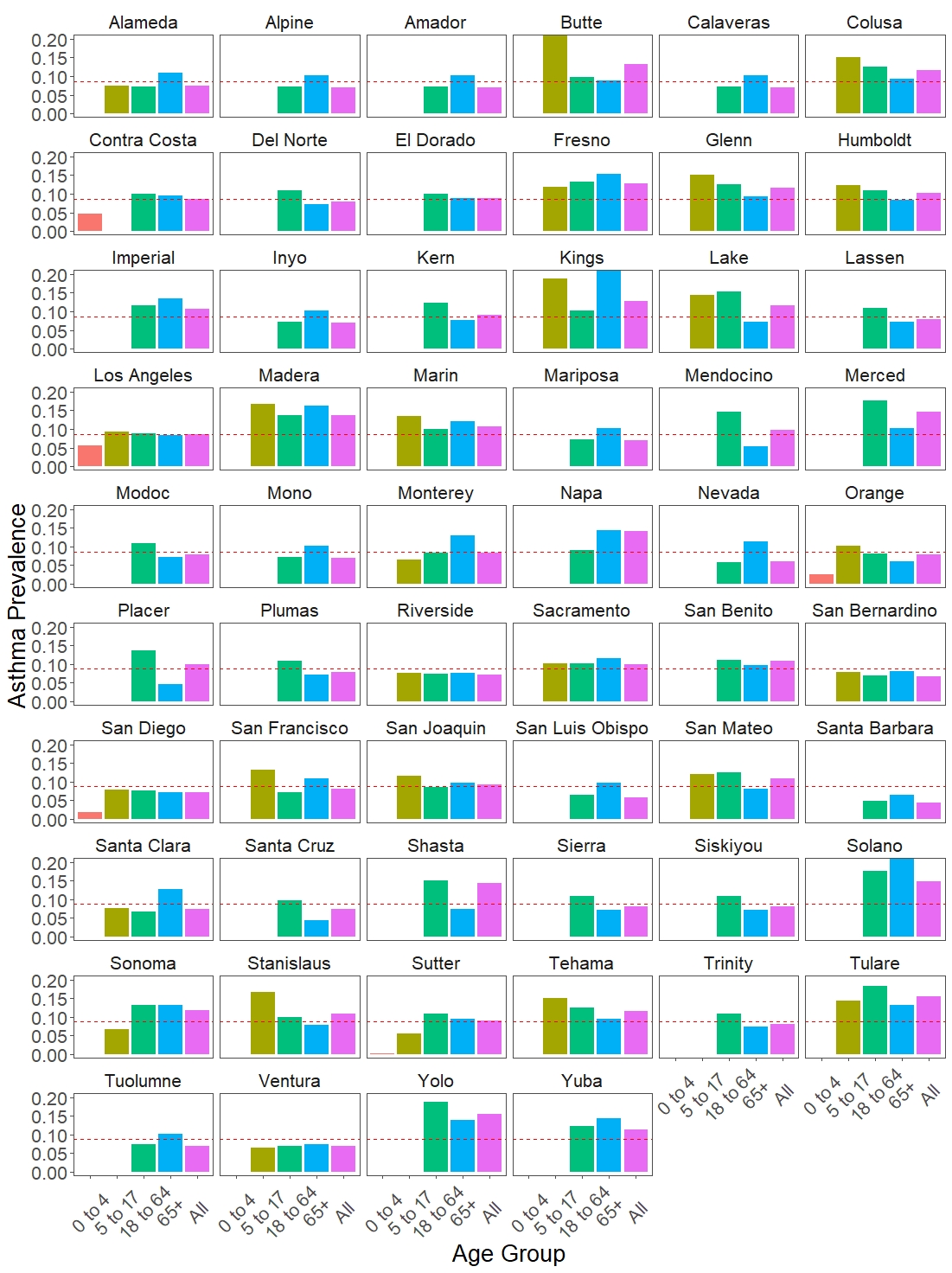
## Census data

The data from which the model population is created were 2021 American Community Survey data, obtained via [IPUMS USA](https://usa.ipums.org/usa/about.shtml), who provide and maintain a database of public use microdata drawn from numerous federal censuses and national surveys. IPUMS does not identify places with a population fewer than 100,000. As a result, the California data included individuals from 35 of California’s 58 counties. These 35 counties are representative of 96.4% of the California population. Extracted variables include county FIPS code, total household income, sex, age, race, Hispanic origin, any health insurance coverage, private health insurance coverage, public health insurance coverage, and poverty status (as a percentage of the federal poverty level).

## Asthma prevalence

Data on current asthma prevalence was obtained on 6/21/2023 from the [CDPH California Asthma Dashboard](https://www.cdph.ca.gov/Programs/CCDPHP/DEODC/EHIB/CPE/Pages/CaliforniaBreathingCountyAsthmaProfiles.aspx), which report California Health Interview Survey data. Detailed notes on the data can be found [here](https://www.cdph.ca.gov/Programs/CCDPHP/DEODC/EHIB/CPE/CDPH%20Document%20Library/Notes_About_the_Data_ADA.pdf). The dataset included asthma prevalence by age group and county. Across all age groups, county-specific prevalence ranged from 0.3% to 33.4%, with an overall California prevalence of 8.7% (Figure ##). Prevalence tended to be higher in suburban and rural counties and was especially high among 5–17-year-olds in Butte County (33.4%), and those over age 65 in Solano (23.8%) and Kings (21.4%) counties.

As with the census data, prevalence data were suppressed when sample sizes were small in order to protect confidentiality. Younger age groups and smaller counties were therefore missing. The youngest age group of 0-4-year-olds had data for only five counties. We therefore used the California prevalence for that age group (3.3%) when the county-specific prevalence was not available. The next age group of 5-17-year-olds also had 53% missing data. Given greater data availability, we imputed the “residual” average for the missing counties weighted by the population size of 5-17-year-olds in each county. Population size by age were obtained from the 2010 Census via Social Explorer tables (SE\_T008) to compute county population weights.



**Figure ##. Current asthma prevalence by age group and county in California, 2019-2020.** Dotted horizontal line indicate California average prevalence (8.7%). Prevalence capped at 20%. Source: 2019-2020 California Health Interview Survey.

## Fire and smoke

We chose the 2018 Camp Fire as given wide availability of evidence for both the smoke pollution it caused as well as health outcomes the event led to. Daily average PM2.5 was selected as the smoke variable. While PM2.5 is among the most dangerous pollutants associated with wildfires, there are limitations to its use as a predictor variable. PM2.5 is not a single substance, rather it refers to a mixture of different chemical compounds. The composition of wildfire PM2.5 greatly differs from ambient pollution and is considered to contain more toxic substances. Measurements of PM2.5 concentration in the air does not provide information about its composition, and therefore does not allow to distinguish differences in potential health impacts. However, the Camp Fire caused substantial increases in PM2.5 concentration in the areas it affected, therefore any PM2.5-related health outcomes over the period of the fire could reasonably be attributed to wildfire smoke rather than ambient pollution.

Daily average PM2.5 measurements between 2018-2019 were obtained from the [Environmental Protection Agency’s Air Data portal](https://www.epa.gov/outdoor-air-quality-data/download-daily-data). Improbable measurements of non-positive PM2.5 concentration were removed as potential measurement errors (0.6% of all observations). For counties with multiple monitors, measurements from multiple monitors were averaged. We then calculated a background PM2.5 value for each county by calculating the median daily PM2.5 during 09/01/2018 and 11/01/2018 (this date range was selected after visualizing trends in PM2.5 over time). To calculate the specific impact of the Camp Fire on PM2.5 pollution, we subtracted this background pollution value from the daily PM2.5 for fire days, [as previously published](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8745685/). Dates outside of the fire period were assigned 0 for fire-specific PM2.5.

## Impact of age on asthma control

Risk ratios for age on asthma control transition probabilitieswere calculated from [Bloom et al](https://thorax.bmj.com/content/thoraxjnl/73/4/313.full.pdf).’s 2017 study of asthma in the UK population. The study evaluated rates of asthma exacerbation across age groups using electronic health record data. While the dependent variable (exacerbation) in this study is not necessarily a direct representation of asthma control, however it is closely related with asthma control. This was the closest available estimate we could find, so we applied them in our model to the probabilities of moving into poorer asthma control states.

|  |  |  |  |
| --- | --- | --- | --- |
| **Age group** | **Reported rate per 10 person-years (95% CI)** | **Calculated probability (weekly)** | **Calculated risk ratio** |
| **Under 5** | 4.27 (4.18 – 4.38) | 0.00818 | (ref) |
| **5 to 17** | 1.48 (1.47 – 1.50) | 0.00284 | 0.348 |
| **18 to 54** | 3.22 (3.21 – 3.24) | 0.00617 | 0.755 |
| **55 and above** | 9.49 (9.37 – 9.42) | 0.01808 | 2.211 |

## Impact of wildfire smoke exposure on asthma control

We utilized the risk ratios estimated by [Wu et al.](https://bmcpediatr.biomedcentral.com/articles/10.1186/s12887-019-1530-7) for the association between smoke exposure and asthma exacerbation. Similar to above, the dependent variable of asthma exacerbation is not necessarily a direct representation of asthma control; however, it was the closest available estimate we could find, and the study design matched well with our model in terms of measurement timepoints. This empirical study found an odds ratio for asthma exacerbation of 1.047 per day of exposure in a one-week model, and 1.049 per day of exposure in a two-week model. We assumed the increased risk of poor control is distributed evenly across asthma control states (i.e., the *relative* risk of transitioning from well controlled to somewhat controlled is the same as the relative risk of transitioning from well controlled to poorly controlled, if exposed), because we did not have any evidence that would suggest otherwise.

# Model Calibration

1. **Calibration of transition probabilities:** Previously published models differed in how they depicted the natural history asthma. [Most](https://doi.org/10.1111/j.1398-9995.2008.01724.x) [models](https://pubmed.ncbi.nlm.nih.gov/22352412/) [used](https://pubmed.ncbi.nlm.nih.gov/32917353/) various healthcare use outcomes as Markov states (e.g., severe exacerbation, emergency department visit). However, we preferred to model asthma control, rather than acute healthcare use events, as we wanted to be able to distinguish small changes in control in response to climate exposures. With expert input, we decided to use a 5-level asthma control framework (ranging from “completely controlled” to “not controlled at all”), based on [GINA classifications](https://ginasthma.org/wp-content/uploads/2019/04/GINA-2019-main-Pocket-Guide-wms.pdf). Further support for this framework was found in [several](https://pubmed.ncbi.nlm.nih.gov/16522452/) [publications](https://pubmed.ncbi.nlm.nih.gov/28669892/) in which 5-category physician assessment scores were used to validate the 3-level Asthma Control Test (ACT).

We therefore needed a transition probability matrix for transitions between 5 states. Starting from a population with “somewhat controlled” asthma, we ran the model through several iterations until equilibrium was reached. We then compared the equilibrium distribution of asthma control to the published distribution of 5-level asthma control. We repeated this until a transition probability matrix was achieved that produced predictions consistent with empirical data on prevalence of asthma control. The calibrated model approached equilibrium within 5 iterations, with a negligible cumulative 0.044 error in asthma control distribution (Table ##).

**Table ##. Calibration for distribution of asthma control**

|  |  |  |  |
| --- | --- | --- | --- |
| **Control status** | **Literature distribution** | **Model at equilibrium** | **Delta** |
| Completely controlled | 0.091 | 0.088 | -0.0026 |
| Well-controlled | 0.391 | 0.390 | -0.0014 |
| Somewhat controlled | 0.285 | 0.301 | 0.0160 |
| Poorly controlled | 0.194 | 0.176 | -0.0182 |
| Not controlled at all | 0.039 | 0.045 | 0.0062 |
| *Sum* | *1.000* | *1.000* | *0.0443* |

We were able to identify [one empirical study](https://www.sciencedirect.com/science/article/pii/S0091674909017709) which reported a Markov transition probability matrix for 3-levels of asthma control. However, this model predicted a high proportion of uncontrolled asthma at equilibrium, with a substantially higher probability of becoming and remaining uncontrolled compared to our calibrated model (the probability of remaining controlled was similar). These 3-state Markov probabilities were based on several clinical trials assessing various asthma therapies, which might suggest poorer control and/or more severe asthma in the study sample that is not comparable to the general population which was the focus of our analysis. As such, we did not do further calibration on transition probabilities.

1. **Calibration of fire-related increases in asthma health resource utilization (HRU):** While the model evaluates the probability of HRU as a function of only disease state (asthma control), having smoke exposure may also lead to seeking health care without necessarily impacting underlying asthma control (e.g., by triggering an acute exacerbation). To account for this increase in acute HRU risk given smoke presence, we generated a second set of HRU probabilities, calibrated to empirical HRU data from prior California wildfires. These adjusted values were applied for individuals who have had at least one day of smoke exposure in that week, whereas the unadjusted HRU probabilities were used for those without exposure. We then tested whether the predicted asthma HRU outcomes were comparable to empirical data by calculating the modelled increases in HRU (urgent/outpatient care, emergency department visits, and inpatient hospitalizations) during model cycles in which smoke was present. The calculated increases were similar to those reported by [Hutchinson et al.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6038982/) and [Malig et al.](https://www.sciencedirect.com/science/article/pii/S004896972102578X?via%3Dihub), and are provided in Table ## below.

**Table ##. Increase in health resource utilization without and without wildfire smoke exposure**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Type of encounter sought** | **None** | **OCS burst** | **Urgent care / Outpatient visit** | **Emergency department visit** | **Hospitalization** |
| **Simulated proportion without smoke** | 0.9200 | 0.0380 | 0.0338 | 0.0052 | 0.0032 |
| **Simulated proportion following 2 weeks of smoke days (cutoff 35 mg/m3)** | 0.8862 | 0.0530 | 0.0440 | 0.0114 | 0.0054 |
| **Simulated % change** | *-4%* | 39% | 30% | 119% | 69% |
| **Empirical evidence from Malig et al.a** | n/a | n/a | n/a | 112% | 53% |
| **Empirical evidence from Hutchinson et al.b** | n/a | n/a | ~30% | 112%\* | ~50% |

aStudy of 2017 North San Francisco Bay wildfires. Reports rates of ED visits and hospitalizations during the fire period (10/9/17 – 10/18/17) and non-fire period (10/1/15 – 10/18/15; 10/1/16 – 10/18/16; 10/1/17 – 10/8/17).

bStudy of 2007 San Diego wildfires. Reports risk ratios of ED visits, hospitalizations, and outpatient visits during periods of fire exposure (10/22/2007 – 10/26/2007) and compared with occurrences during six 5-day periods matched on the day of the week starting 3-9 weeks before the exposed period.

\*The two studies calculated identical increases in emergency department visits.

**TO ADD:**

* Residual probability calculation for “stay-put” transitions
* Adjusting costs for inflation?
* Adjusting hsu for delayed return to health