It is made available under a [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) . **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2023.01.31.23285059;](https://doi.org/10.1101/2023.01.31.23285059) this version posted June 11, 2024. The copyright holder for this preprint

# 1 Long-term exposure to wildland fire smoke  $PM<sub>2.5</sub>$  and mortality in the 2 contiguous United States

3 Yiqun Ma<sup>1,2</sup>, Emma Zang<sup>3</sup>, Yang Liu<sup>4</sup>, Jing Wei<sup>5</sup>, Yuan Lu<sup>6,7</sup>, Harlan M. Krumholz<sup>6,7</sup>, Michelle L. 4 Bell<sup>8</sup>, Kai Chen<sup>1,2,\*</sup>

- 5 <sup>1</sup>Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT, 6 USA
- 7 <sup>2</sup>Yale Center on Climate Change and Health, Yale School of Public Health, New Haven, CT, USA
- 8 <sup>3</sup>Department of Sociology, Yale University, New Haven, CT, USA
- 9<br>10 <sup>4</sup>Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA
- 11 <sup>5</sup>Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary
- 12 Center, University of Maryland, College Park, MD, USA<br>13 <sup>6</sup>Center for Outcomes Research and Evaluation. Yale N
- <sup>6</sup>Center for Outcomes Research and Evaluation, Yale New Haven Hospital, New Haven, CT, 14 USA
- 15 <sup>7</sup>Section of Cardiovascular Medicine, Department of Medicine, Yale School of Medicine, New 16 Haven, CT, USA
- 17 <sup>8</sup>School of the Environment, Yale University, New Haven, CT, USA
- 18<br>19
- \*Corresponding author: Kai Chen
- 20 **Email:** [kai.chen@yale.edu](mailto:kai.chen@yale.edu)
- 21
- 22 **Author Contributions:** Y.M. conducted formal analyses and drafted the manuscript. K.C.
- 23 conceived of and supervised the conduct of this study and edited the manuscript. K.C. and E.Z.<br>24 applied and managed the mortality dataset. J.W. modelled the all-source PM<sub>2.5</sub> data and revised
- applied and managed the mortality dataset. J.W. modelled the all-source  $PM_{2.5}$  data and revised
- 25 the manuscript. E.Z., Y.L., Y.L., H.M.K., and M.L.B. contributed to the interpretation of results and
- 26 manuscript revision. All authors reviewed and approved the final version of this manuscript.
- 27 **Competing Interest Statement:** The authors declare no conflict of interests.
- 28 **Classification:** Biological Sciences, Environmental Sciences
- 29 **Keywords:** wildland fire, wildfire, mortality, fine particulate matter, United States
- 30

# **Abstract**

 Despite the substantial evidence on the health effects of short-term exposure to ambient fine 33 particles  $(PM_{2.5})$ , including increasing studies focusing on those from wildland fire smoke, the impacts of long-term wildland fire smoke PM<sub>2.5</sub> exposure remain unclear. We investigated the association between long-term exposure to wildland fire smoke PM2.5 and non-accidental mortality and mortality from a wide range of specific causes in all 3,108 counties in the contiguous U.S., 2007–2020. Controlling for non-smoke PM2.5, air temperature, and unmeasured spatial and temporal confounders, we found a non-linear association between 12-month moving average 39 concentration of smoke  $PM_{2.5}$  and monthly non-accidental mortality rate. Relative to a month with 40 the long-term smoke PM<sub>2.5</sub> exposure below 0.1  $\mu$ g/m<sup>3</sup>, non-accidental mortality increased by 0.16-0.63 and 2.11 deaths per 100,000 people per month when the 12-month moving average of  $P_{\text{M}_{2.5}}$  concentration was of 0.1-5 and 5+  $\mu g/m^3$ , respectively. Cardiovascular, ischemic heart disease, digestive, endocrine, diabetes, mental, and chronic kidney disease mortality were all 44 found to be associated with long-term wildland fire smoke  $PM_{2.5}$  exposure. Smoke  $PM_{2.5}$  contributed to approximately 11,415 non-accidental deaths/year (95% CI: 6,754, 16,075) in the contiguous U.S. Higher smoke PM2.5-related increases in mortality rates were found for people aged 65 above. Positive interaction effects with extreme heat (monthly number of days with daily 48 mean air temperature higher than the county's 90<sup>th</sup> percentile warm season air temperature) were also observed. Our study identified the detrimental effects of long-term exposure to wildland fire 50 smoke  $PM_{2.5}$  on a wide range of mortality outcomes, underscoring the need for public health actions and communications that span the health risks of both short- and long-term exposure.

# 52 **Significance Statement**



#### **Main Text**

#### **Introduction**

Wildland fire is a growing public health concern in the United States. As a result of the warming

climate (1), a long history of fire suppression (2), and an increase in human-caused fire ignitions

(3), the country has witnessed a marked increase in the area affected by wildland fires over the

past few decades, with the burned area roughly quadrupling (4). In recent years, wildland fire

contributed to up to 25% of total fine particulate matter (PM2.5) concentrations across the U.S.,

and up to half in some Western regions (4). Under climate change, the prevalence, frequency,

and intensity of wildland fire activities are expected to increase in the future (5).

 Wildland fire smoke is a complex mixture. Among the various air pollutants emitted by 75 wildland fires,  $PM_{2.5}$  is widely used as an indicator of exposure because it is a major component of smoke, can deeply penetrate the respiratory system, and has demonstrated links to public health (6). Previous studies on the health effects of wildland fire smoke mostly focused on the western U.S., where the majority of large fires occurred (7-9). However, the pollutants from wildland fire smoke can travel long distances from the source, potentially affecting human health 80 thousands of kilometers away outside the West (10).

 Previous studies on the health effects of wildland fire exposure predominantly focused on the effects of short-term exposure, typically examining exposure periods within one or two weeks. Most studies reported a positive relationship between short-term wildland fire smoke exposure and all-cause mortality (9, 11-13). A growing number of studies linked short-term exposure to wildland fire smoke to increased risks of respiratory mortality and presented mixed evidence regarding cardiovascular mortality (11-14). Recent studies have also documented worsened diabetic outcomes (15), higher mortality rates among patients with kidney failure (16), and impaired mental health (17, 18) associated with short-term wildland fire smoke exposure. However, given that climate change has contributed to an increase in wildland fire season length, increasing the duration of exposure, the health impacts of long-term wildland fire smoke exposure should be a concern of growing importance (19-21). To date, little is known

 about the impacts of long-term exposure to wildland fire smoke on human health (19, 21). Most previous studies focused on mental health, reporting associations between long-term wildland fire exposure and mental outcomes, such as anger problems (22), posttraumatic stress disorder (23), depression (24), and anxiety (25). However, these studies could not distinguish between the effects of smoke and the overall impacts of wildland fires. A few studies suggested that exposure to wildland fires was associated with child mortality, COVID-19 mortality, cardiovascular disease mortality, and respiratory disease morbidity (21), but most of them did not measure the uncertainties of estimates. Given that short-term exposure to wildland fire smoke has been linked to a wide range of mortality outcomes, exploring whether long-term wildland fire smoke exposure is associated with these health effects is worthwhile.

 Furthermore, the effects of wildland fire smoke PM2.5 can be heterogeneous among population subgroups due to physiological, behavioral, and socioeconomic factors (26). Previous studies indicated that demographic factors such as sex, age, and race and ethnicity may modify 105 the association between smoke  $PM_{2.5}$  and health outcomes (11, 26, 27). However, existing research mostly focused on the effects of short-term wildland fire smoke exposure and generated mixed findings (9). To better prepare communities for smoke and tackle environmental justice issues, a deeper understanding of susceptibility to wildland fire exposure, particularly long-term exposure, among specific subgroups is needed to help inform targeted public health outreach efforts.

 In the context of climate change, the co-occurrence of wildland fires and extreme heat 112 events is expected to increase (28). In addition, both extreme heat and  $PM_{2.5}$  were associated with impaired cardiopulmonary functions (29, 30). Therefore, extreme heat may interact with wildland fire smoke and further aggravate health effects. A previous study suggested synergistic effects between short-term extreme heat and wildland fire smoke exposures on daily cardiorespiratory hospitalizations in California (31). However, little is known about the potential 117 interaction between extreme heat and long-term smoke  $PM_{2.5}$  exposure nationwide. Constrained by a lack of nationwide validated data on pollutant concentrations

 attributable to wildland fire smoke, many previous studies on the impact of wildland fire smoke on health outcomes have been focusing on episodes with high wildland fire smoke exposure (or smoke wave) using binary measures of smoke concentrations (32). Recently, a machine learning 122 model was developed to estimate wildland fire smoke  $PM<sub>2.5</sub>$  concentrations for the contiguous U.S., using a combination of meteorological factors, fire variables, aerosol measurements, and 124 land use and elevation data (33). This new, high-resolution wildland fire smoke  $PM_{2.5}$  dataset (10×10 km<sup>2</sup>) enabled us to further examine the impact of wildland fire smoke, ranging from the more common, low-level smoke concentrations to the increasingly frequent extremely high concentrations, and to explore its potentially non-linear effects on mortality. Utilizing the nationwide monthly wildland fire smoke PM2.5 and mortality data from 2007 to 2020, this study aimed to (a) estimate the potentially non-linear associations of long-term smoke 130 PM<sub>2.5</sub> exposure with county-level monthly non-accidental and cause-specific mortality from a broad spectrum of diseases, (b) calculate the attributable cause-specific mortality burden in each county, (c) examine the associations in different sex, age, and racial and ethnic groups, and (d) 133 explore the interaction effect between smoke  $PM_{2.5}$  and extreme heat on mortality.

#### **Results**

## *Description of smoke PM2.5 exposure and monthly mortality in the contiguous U.S.*

 For months from January 2007 to December 2020, we calculated the moving average of 137 smoke PM<sub>2.5</sub> concentration of the current and previous 11 months for each county to represent 138 the average exposure to smoke  $PM_{2.5}$  in the previous year. The average 12-month moving 139 average concentration of smoke  $PM_{2.5}$  across all county-months during the study period was 0.4 140  $\mu$ g/m<sup>3</sup>, contributing to approximately 5% of all-source PM<sub>2.5</sub> (Table S1).

 From 2007 to 2020, all 3,108 counties in the contiguous U.S. experienced some amount 142 of smoke  $PM_{2.5}$ , with the western, north central, and southeastern counties being exposed to 143 higher long-term exposure (12-month moving average concentration of smoke PM<sub>2.5</sub>) than other regions (**Fig. 1A**). The temporal variability in this exposure was also higher in these regions compared with other regions (**Fig. 1B**). The overall temporal trend of the 12-month moving average concentration of smoke PM2.5 for all U.S. contiguous counties is displayed in **Fig. S1**. To account for the potentially non-linear health effects of smoke PM2.5, we categorized the 12-month moving average of smoke PM2.5 concentrations into 9 bins: 0-0.1 (reference), 0.1-  $0.2, 0.2$ -0.3, 0.3-0.4, 0.4-0.5, 0.5-0.7, 0.7-1, 1-5 and  $5+ \mu q/m<sup>3</sup>$ , corresponding to approximately

 9.9%, 17.8%, 17.7%, 14.4%, 11.1%, 14.9%, 9.1%, 5.0%, and 0.1% of county-months in the study period (**Fig. S2**). Only 15 county-months were not exposed to any smoke PM2.5 in the current and previous 11 months.

 A total of 33,902,722 non-accidental deaths were included in this study, including 11,514,374 deaths from cardiovascular diseases, 3,584,654 from respiratory diseases, 1,624,915 from endocrine diseases, 939,531 from genitourinary diseases, 2,465,890 from nervous diseases, 1,823,243 from mental and behavioral disorders (hereafter referred to as "mental disorders"), and 1,406,184 from digestive diseases. From 2007 to 2020, the mean age-adjusted non-accidental mortality rate was 61.9 (SD: 23.9) deaths per 100,000 people per month in all counties in the contiguous U.S. The descriptive statistics of age-adjusted mortality rate for each population subgroup and each specific cause are listed in **Table S1**.



 We also examined the associations between 12-month moving average of smoke PM2.5 concentrations and monthly mortality rates from other specific causes, including stroke, myocardial infarction, hypertensive diseases, hypertensive heart diseases, respiratory diseases, chronic obstructive pulmonary disease (COPD), nervous diseases, Alzheimer's disease and related dementias (ADRD), and genitourinary diseases (**Fig. S3**). Although some of them were 193 also found to be sensitive to long-term smoke  $PM_{2.5}$  exposure, such as mortality from stroke, myocardial infarction, respiratory diseases, and ADRD, these specific causes were not presented in the main figures or included in further analysis because their estimates were either insignificant after the Bonferroni correction for multiple comparisons, or inconsistent in direction across smoke 197 PM<sub>2.5</sub> concentration bins. Negative effect estimates were found for hypertensive heart disease, COPD, and nervous disease mortality in some concentration bins, but none of them remain statistically significant after Bonferroni correction. The detailed estimates for all examined specific causes are listed in **Table S2**.

#### *Cause-specific mortality burden attributable to long-term smoke PM2.5 exposure*

203 Assuming the homogeneity within each smoke  $PM_{2.5}$  bin and the causality of the estimated smoke PM2.5-mortality relationships, we further quantified the mortality burden 205 attributable to long-term smoke  $PM_{2.5}$  exposure. As non-smoke  $PM_{2.5}$  was adjusted for in the 206 model, the estimated smoke  $PM_{2.5}$ -attributable mortality burden is in addition to the well- recognized burden from non-smoke PM2.5. On average, approximately 11,415 non-accidental deaths (95% CI: 6,754, 16,075) were attributable to smoke PM2.5 in the contiguous U.S. per year. The spatial distribution of this attributable burden was generally consistent with the distribution of smoke PM2.5 concentration (**Fig. 3A**). The estimated attributable non-accidental mortality burden for each year is displayed in **Fig. S4**.

212 For mortality from other specific causes, each year, long-term smoke PM<sub>2.5</sub> exposure contributed approximately 4,512 deaths (95% CI: 1,922, 7,102) from cardiovascular diseases, including 3,753 (95% CI: 1,915, 5,592) ischemic heart disease deaths, 1,142 deaths (95% CI:

 285, 1,999) from endocrine diseases, including 858 (95% CI: 149, 1,566) diabetes deaths, 2,083 deaths (95% CI: 1,143, 3,022) from mental disorders, 537 deaths from digestive diseases (95% CI: -200, 1,273), and 320 deaths (95% CI: -72, 713) from chronic kidney diseases. Among the total long-term smoke PM2.5-attributable non-accidental deaths, approximately 72.5% were from cardiovascular diseases, mental disorders, endocrine diseases, and digestive diseases (**Fig. 3B; Table S3**).

#### *Association between long-term smoke PM2.5 and monthly mortality rate by subgroup*

 We examined the association between 12-month moving average of smoke PM2.5 concentrations and monthly non-accidental mortality rates across different sex, age, and race and ethnicity groups (**Table 1**). No significant difference in effect estimates was observed between males and females. Compared with people aged 0 to 64, greater increases in mortality rates were 227 observed among people who aged above 65 across all smoke PM<sub>2.5</sub> concentration bins. Among different racial and ethnic groups, compared with non-Hispanic White people, significantly higher 229 associations were observed for non-Hispanic Black and Hispanic people when smoke  $PM_{2.5}$ 230 concentration was from 1 to 5  $\mu q/m<sup>3</sup>$ . After considering multiple comparisons, the difference between racial and ethnic minorities and non-Hispanic White people became insignificant, but the significant differences between age groups remained. Subgroup analyses were also performed for major categories of mortality that showed

 consistent and significant sensitivity to long-term exposure to smoke PM2.5: cardiovascular, endocrine, digestive, and mental mortality (**Table S4**). In general, no significant difference

between sex or race and ethnicity groups was detected for these four outcomes after considering

multiple comparisons. Consistent with the findings for non-accidental mortality, greater smoke

PM2.5-related increases in mortality rates from cardiovascular diseases and mental disorders

were observed among people who aged above 65.

#### *Interaction between long-term smoke PM2.5 and extreme heat days*

242 Extreme heat commonly co-occurred with wildland fire smoke  $PM_{2.5}$  in the contiguous 243 U.S. Here, we examined the interaction effects of long-term smoke PM<sub>2.5</sub> and current-month extreme heat days. For each county, we defined extreme heat days as days with daily mean air 245 temperature higher than the county's  $90<sup>th</sup>$  percentile warm season air temperature (May to September, 2007-2020). From 2007 to 2020, a total of 657,402 extreme heat days were identified in the counties studied, spanning 126,237 county-months. We calculated the number of extreme heat days in each county for each month to represent monthly extreme heat exposure. Based on 249 the spatial distribution of the 12-month moving average of smoke PM<sub>2.5</sub> concentrations and the current-month number of extreme heat days (**Fig. 4A**), the North Central, South, Southeast, and 251 West regions experienced higher co-exposure to smoke  $PM_{2.5}$  and extreme heat than other regions. By including an interaction term for the current-month number of extreme heat days and 254 12-month moving average of smoke PM<sub>2.5</sub> concentration bins in our model, we found a 255 significantly positive interaction between extreme heat and smoke PM<sub>2.5</sub> at levels ranging from 0.1 to 5 μg/m<sup>3</sup> (**Fig. 4B**). The interaction results in general remained robust when we used 257 alternative temperature thresholds in the definition of extreme heat days (85<sup>th</sup> or 95<sup>th</sup> percentile of local warm-season temperature distribution). Among mortality of specific causes, positive

interactions were found for cardiovascular and mental disorder mortality (**Fig. S5**).

#### *Sensitivity analyses, placebo tests, and stratification by distance to fire*

 Our results generally remained robust when we included different combinations of fixed 263 effects, additionally adjusted for dew point temperature, did not adjust for non-smoke PM<sub>2.5</sub> or current-month air temperature, adjust for 12-month moving average of air temperature instead of current-month temperature, or used alternative degrees of freedoms in the natural cubic spline of air temperature in the model (**Fig. S6**). The estimated non-linear pattern also remained consistent when we used a quasi-Poisson model and used finer bins or a natural cubic spline for smoke PM2.5 in the model (**Fig. S7-S9**). The results of the spatial and temporal randomization tests

269 indicated that the estimates in our study were unlikely driven by spatial or temporal dependence 270 due to model misspecification (**Fig. S10**). When the exposure period to smoke PM2.5 was 271 extended from 12 months to 24 and 36 months, the model coefficients decreased and eventually 272 becoming null at a 36-month exposure window in most bins (**Fig. S11**). 273 In a stratified analysis by distance to fire, we observed that the association between long-274 term smoke  $PM_{2.5}$  exposure and non-accidental mortality were significantly higher in counties far 275 away from the fire point  $(≥ 150 \text{ km})$  in the current and past 11 months compared with those close 276 to an active fire  $( $150 \text{ km}$ )$  when the 12-month moving average of smoke PM<sub>2.5</sub> concentrations 277 was from 0.2 to 0.4 μg/m<sup>3</sup> and from 1 to 5 μg/m<sup>3</sup> (Fig. S12). However, when the smoke PM<sub>2.5</sub> 278 concentrations were above 5  $\mu$ g/m<sup>3</sup>, the estimate was only significantly positive in counties close 279 to fire. This result suggests that being close to a fire was unlikely to be a main driver of the

280 estimated results in our study.

# **Discussion**

 To the best of our knowledge, this is the first study comprehensively examining the 283 associations between long-term exposure to wildland fire smoke  $PM_{2.5}$  and mortality from a wide range of specific causes for all ages in the whole contiguous U.S. We found that average 285 exposure to wildland fire smoke  $PM<sub>2.5</sub>$  in the past one year was associated with increases in non- accidental, cardiovascular, ischemic heart disease, digestive, endocrine, diabetes, mental, and chronic kidney disease mortality. In addition to the well-documented mortality burden from non- smoke PM2.5, in total, we estimated that smoke PM2.5 contributed to over 10,000 non-accidental 289 deaths in the contiguous U.S. each year. Higher smoke PM<sub>2.5</sub>-related increases in mortality rates were found for people aged 65 above. In addition, positive interaction effects between 12-month 291 moving average of smoke  $PM_{2.5}$  concentrations and current-month number of extreme heat days on non-accidental, cardiovascular, and mental disorder mortality were observed.

 The impacts of long-term wildland fire smoke exposures on mortality are understudied in existing literature. However, our results are in general consistent with a few previous findings that long-term exposure to wildland fire smoke is associated with increased premature mortality (21). For example, a study in Brazil observed a significant association between long-term exposure to 297 forest fire PM<sub>2.5</sub> (measured by percentage hours of PM<sub>2.5</sub> concentrations > 25 µg/m<sup>3</sup> divided by 298 the total number of estimated hours of  $PM_{2.5}$  in 2005) and increases in cardiovascular disease 299 mortality rates in older adults ( $\geq 65$  years) (34). Controlling for non-fire-sourced PM<sub>2.5</sub>, a case-300 control study in low-income and middle-income countries reported that each 1  $\mu$ g/m<sup>3</sup> increment of 301 monthly mean fire-sourced  $PM_{2.5}$  concentration was associated with a 2.31% increased risk of child mortality, but this estimate became statistically insignificant after extending the exposure time window from 1 month to 12 months (35). However, given the differences in study population and location, exposures (e.g., the intensity, frequency, and duration of wildfire smoke), study design, and outcome measures, estimates are not directly comparable across studies. More studies investigating the long-term impacts of wildland fire smoke exposure on mortality from a comprehensive spectrum of specific causes and conditions are needed in the future.

308 Although the health impacts of wildland fire smoke  $PM_{2.5}$  could be different from urban background PM2.5 due to differences in chemical composition and particle size, and the episodic nature of smoke (7, 36), the biological mechanisms are likely to align with those documented for 311 all-source (i.e., total mass)  $PM_{2.5}$ . PM<sub>2.5</sub> can travel into the respiratory tract and bloodstream and trigger oxidative stress and inflammation, leading to impaired lung and vascular function (37). PM<sub>2.5</sub> may deposit in the kidney, contributing to the development of kidney diseases (38). PM<sub>2.5</sub> can also enter the gastrointestinal tract, causing imbalances in the intestinal microecology (39). In 315 addition, PM<sub>2.5</sub> exposure has been associated with insulin resistance, which may progress to diabetes and other endocrine diseases (40). Furthermore, the oxidative stress, systemic and 317 neuroinflammation, and hypothalamic-pituitary-adrenal axis dysregulation triggered by PM<sub>2.5</sub> have been linked to psychological diseases (41). The aggravation of physical health conditions could also worsen mental health (42). A multi-level conceptual framework has also been proposed for understanding the pathways connecting wildland fire smoke with mental health and well-being, which involves loss of nature, reduced access to livelihoods, reduced outdoor activities, and many other social and behavioral factors (17). More research is needed in the future to better understand the underlying mechanisms of the health impacts of wildland fire smoke. In our study, higher effect estimates were found for older adults compared with people 325 aged 0-64, which is consistent with the literature of all-source PM<sub>2.5</sub>. The greater increases in 326 mortality rates associated with smoke  $PM_{2.5}$  among older adults reflect both higher baseline mortality rates and higher susceptibility to air pollution in the older population as found in previous studies, due to decreased physiological, metabolic and compensatory processes, and a higher prevalence of comorbidities (43). In addition, we found higher effect estimates in counties far away from the fire point compared with those close to an active fire when the 12-month moving 331 average of smoke PM<sub>2.5</sub> concentrations was below 5  $\mu$ g/m<sup>3</sup>. This finding suggests a substantial public health burden from smoke, given that the population residing near the fires is likely much smaller compared to the larger number of people affected downwind due to smoke transport.

 Furthermore, we observed a positive interaction effect between long-term exposure to 335 wildland fire smoke PM<sub>2.5</sub> and current-month extreme heat. To date, although research has highlighted the synergistic health impacts of heat and air pollution (44), only a limited number of 337 studies have specifically examined the interaction with  $PM_{2.5}$  from wildland fire smoke. A recent study in California found evidence of synergistic effects between extreme heat and wildland fire 339 smoke (31); however, since this study focused on short-term smoke PM<sub>2.5</sub> exposure, its results are not directly comparable to those of our study. The positive interaction between wildland fire smoke PM2.5 and extreme heat indicates an increasing mortality burden for U.S. populations in the future given that the co-exposure to both hazards is expected to increase under the changing climate (28). Further studies on the compounding effects of wildland fires and other climate-related events, such as heatwaves and droughts, are warranted.

 Our study estimated that over 10,000 non-accidental deaths per year resulted from wildland fire smoke PM2.5 in the contiguous U.S., which is over 1,000 times higher than the recorded wildfire deaths in the U.S. Billion-dollar Weather and Climate Disasters report by the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (10 deaths/year due to the fire itself) (45). This indicates a great number of deaths brought by wildland fires that could not be captured by official tolls. According to the Global Burden of Disease Study, ambient particulate matter pollution contributed to an average of approximately 67,800 deaths annually in the contiguous U.S. from 2007 to 2020 (46, 47). Using this number as a reference, our results indicate that smoke PM2.5-related deaths account for 354 about 16.8% of the deaths associated with all-source  $PM_{2.5}$ . In addition, the U.S. Billion-dollar Weather and Climate Disasters report estimated that wildfire events cost about 3.1 billion dollars per year in the U.S. (45), but this estimate does not take into account the health care related losses or values associated with loss of life due to the fire itself or smoke (48). A recent study reported that the economic value of the health impacts of wildland fire smoke could be in the tens 359 to hundreds of billions of U.S. dollars, but the exposure-response function for the  $PM<sub>2.5</sub>$ -mortality 360 relationship they used was for all-source  $PM_{2.5}$ , not wildland fire-specific PM<sub>2.5</sub> (49). Our study

 suggests a tremendous wildland fire smoke-related mortality burden, and our effect estimates for the relationship between wildland fire PM2.5 and mortality could be applied in future estimates of the costs of wildland fires to more fully account for the fire and smoke impacts.

 The findings of our study have several key implications. First, wildland fire smoke is a national concern in the U.S. Its health effects extend beyond the western regions where wildland fires mostly occur, impacting the entire country. Second, the health impacts of wildland fire smoke are not limited to those in response to short-term (daily) exposures. Long-term exposure to smoke PM2.5 contributed to substantial mortality burden in the U.S. and will become increasingly important in the future due to the prolonged wildland fire seasons under climate change. Third, the health impacts of wildland fire smoke cover a wide spectrum of causes of death ranging from cardiovascular and endocrine diseases to mental disorders and digestive diseases. Furthermore, in addition to the well-recognized detrimental effects of extremely high concentrations of smoke PM2.5, even relatively low levels can be harmful. Finally, the relationship between wildland fire smoke and mortality appears to be non-linear, underscoring the necessity for further research into the modifying factors of this association and highlighting the importance of utilizing concentration- specific exposure-response functions in future health impact assessments. Some limitations of this study should be noted. First, the wildland fire smoke PM2.5 concentrations were modelled and subject to uncertainty. Because direct measurements of the 379 smoke contribution to PM<sub>2.5</sub> pollution are not available, the smoke PM<sub>2.5</sub> prediction was based on PM2.5 anomalies at monitoring stations, which may be an imprecise estimate of the 381 concentrations of smoke (33).  $PM<sub>2.5</sub>$  from prescribed fire smoke is under-represented in the model output; therefore, the exposure data we used does not fully capture the total smoke exposure people experience from wildland fires (i.e., both wildfire and prescribed fire). Additionally, wildland 384 fire smoke has a unique spatiotemporal pattern and measuring long-term smoke  $PM_{2.5}$  exposure 385 using average concentrations may overlook its episodic nature. Average smoke  $PM_{2.5}$ 

 concentrations primarily reflect intensity but do not adequately capture the frequency and duration aspects of long-term smoke PM2.5 exposure (50). Future studies using different exposure metrics

388 that better reflect various aspects of long-term smoke  $PM_{2.5}$  exposure are warranted.

Furthermore, we assumed that the non-smoke PM2.5 was the difference between the all-source

PM<sub>2.5</sub> and the smoke PM<sub>2.5</sub> concentrations, but they were generated using different methods and

may introduce measurement errors.

 Second, lacking detailed location information, this county-level ecological study is susceptible to ecological fallacy. We were unable to capture the within-county heterogeneity of smoke PM2.5 exposure among different subgroups or analyze the influence of wildfire evacuation in this study. Besides, the estimates of smoke PM2.5 in our study may partially capture the health effects of other pollutants within the smoke mixture. Due to the absence of publicly available, full- spatial coverage data updated to 2020, we were unable to adjust for other wildland fire-related air pollutants, such as nitrogen dioxide, ozone, carbon monoxide, and polycyclic aromatic 399 bydrocarbons. In addition, we used the smoke  $PM_{2.5}$  concentration range of 0-0.1  $\mu$ g/m<sup>3</sup> as the 400 reference due to the limited number of county-months with 0  $\mu$ g/m<sup>3</sup> long-term exposure. This 401 choice of reference could result in an underestimation of both the smoke PM<sub>2.5</sub>-related mortality changes and the attributable mortality burden. Furthermore, although we examined mortality from a wide range of plausible specific causes, we did not cover all possible causes of death (e.g., infectious diseases and cancer were excluded). Future studies are warranted to explore the association between wildland fire smoke and these diseases. Finally, the calculation of attributable mortality burden was based on the assumptions of homogeneity and causality of the estimated smoke PM2.5-mortality relationships (see details in Methods). Violation of these assumptions may bias the results. In conclusion, our study identified the detrimental effects of long-term wildland fire smoke

 PM2.5 exposure on a wide range of mortality outcomes in the U.S. With wildland fire intensity and frequency anticipated to increase in the future driven by climate change (5), more effective public health actions and communications that span the health risks of short- and long-term exposure are urgently needed both in and outside the areas where the wildland fires occur.

# **Materials and Methods**

# *Mortality and population data*



Ambient wildland fire smoke PM2.5 estimates for the contiguous U.S. was provided by a

441 recent study by Childs et al. (33). In brief, smoke days were identified as days when smoke was 442 overhead based on satellite imagery, and station-based ground smoke PM2.5 on those days was 443 calculated as anomalies above the median on non-smoke days. Then, a model was trained to 444 predict the station-based smoke PM2.5 using meteorological factors, fire variables, aerosol 445 measurements, and land use and elevation data. Finally, the trained model was applied to 446 produce daily estimates of smoke  $PM_{2.5}$  over the contiguous U.S. at a resolution of 10×10 km<sup>2</sup> 447 (33). This model performed well over the entire range of observed smoke PM<sub>2.5</sub> ( $R^2 = 0.67$ ), but 448 the model performance was lower on days with station-based smoke  $PM_{2.5}$  concentrations above 449  $\,$  50 µg/m<sup>3</sup> compared to days with concentrations below 50 µg/m<sup>3</sup> (33). We additionally validated 450 this model against a recently published wildland fire-specific PM<sub>2.5</sub> model in California, which 451 applied a novel ensemble-based statistical approach to isolate wildland fire-specific  $PM_{2.5}$  from 452 other sources of emissions (52). This external validation showed a great consistency between the 453 monthly county-level predictions from these two models in California, 2006–2020, with an Rsquared (R<sup>2</sup> ) value of 0.92 and a root-mean-square error (RMSE) of 1.14 µg/m<sup>3</sup> 454 (**Fig. S13**). The 455 daily smoke PM<sub>2.5</sub> concentrations were aggregated into monthly county-level average using 456 population-weighted averaging to match with the mortality data. For months from January 2007 to 457 December 2020, we calculated the 12-month moving average of smoke PM<sub>2.5</sub> concentration for 458 each county to represent the average exposure to smoke PM<sub>2.5</sub> in the previous year. To account 459 for the potentially non-linear effects, we divided the 12-month moving average smoke  $PM_{2.5}$ 460 concentrations into 9 bins: 0-0.1 (reference), 0.1-0.2, 0.2-0.3, 0.3-0.4, 0.4-0.5, 0.5-0.7, 0.7-1, 1-5 and 5+ μg/m<sup>3</sup> 461 . The distribution of samples across the bins is shown in **Fig. S2.** 462 **Data of daily total all-source PM**<sub>2.5</sub> concentrations at 1×1 km<sup>2</sup> resolution were obtained 463 from the USHighAirPollutants (USHAP) dataset (53). This daily surface PM<sub>2.5</sub> concentration data 464 was derived via a deep learning model that integrated big data from satellites, models, and 465 surface observations (53). Similar to smoke  $PM_{2.5}$ , we averaged the daily all-source  $PM_{2.5}$ 466 concentrations into monthly county-level data. Non-smoke PM<sub>2.5</sub> concentrations were then 467 calculated by subtracting the smoke  $PM_{2.5}$  from the all-source  $PM_{2.5}$  concentrations. For negative



# 495 *Statistical analysis*

 To estimate the association between long-term exposure to wildland fire smoke PM2.5 and monthly mortality rates, we applied a panel fixed effects model which exploits local temporal variation in both exposure and outcome. Panel fixed effects models have been increasingly applied in environmental epidemiology in recent years (57, 58). In our study, the main model can be expressed as

501 MortalityRate<sub>c.v.m</sub>

$$
= \alpha_c + \theta_{y,m} + \delta_{s,y} + \eta_{c,m}
$$

$$
+\sum_{b=1}^{B}(\beta_b \text{SmokePM2.5Bins}_{c,y,m}^b)+\beta \text{NonSmokePM2.5}_{c,y,m}
$$

$$
+ ns(Temperature_{c,y,m}, df = 5) + \varepsilon_{c,y,m}
$$

505 where Mortality Rate<sub>cym</sub> represents the non-accidental or other cause-specific age-adjusted 506 mortality rates in county  $c$ , year  $y$ , and month  $m$ .  $SmokePM2.5Bins_{c,y,m}^b$  is a dummy for whether 507 the moving average of smoke  $PM<sub>2.5</sub>$  concentration of the current and previous 11 months in 508 county c, year y, and month m falls into the concentration range of bin  $b$ .  $\alpha_c$  refers to time-509 invariant county effects and  $\theta_{v,m}$  refers to time-varying effects that are common in all counties. By 510 introducing the indicators for each county and each year-month, this model can potentially control 511 for all spatial confounders that only vary across counties (e.g., urbanicity) and all temporal 512 confounders that only vary by time (e.g., seasonality), either measured or unmeasured (59). 513 Fixed effects at the state by year level  $(\delta_{s,y})$  and at the county by month of year level  $(\eta_{c,m})$ 

514 account for regional long-term trend and seasonality in exposure and outcomes.

515 NonSmokePM2.5<sub>c.v.m</sub> is the 12-month moving average of non-smoke PM<sub>2.5</sub> concentrations in

516 county  $c$ , year  $y$ , and month  $m$ . Current-month air temperature was controlled by a flexible natural

517 cubic spline with five degrees of freedom (df).  $\varepsilon_{c,v,m}$  is the error term. We weighted models using

518 the population size in each county to improve the precision of our estimates (60).

519 Heteroskedasticity-robust standard-errors were used to compute 95% confidence intervals.

520 Associations were considered statistically significant at α of 0.05 with a Bonferroni correction

521 given as  $\alpha/m$ , where m is the number of examined specific causes of death (17 specific causes). 522 Only specific causes with estimates consistent in direction across smoke PM<sub>2.5</sub> concentration bins 523 and statistically significant after Bonferroni correction were included in further analysis. 524 In subgroup analyses, to estimate the association between long-term smoke PM2.5 525 exposure and monthly mortality rates by sex, age, and race and ethnicity, we used an expanded 526 dataset nesting the subgroups, interacted the subgroup variable with all terms in the model, and 527 reported the statistical significance of the interaction term between the subgroup variable and 528 smoke PM<sub>2.5</sub>. To account for multiple comparison, Bonferroni correction was also performed in 529 subgroup analyses ( $m = 3$  subgroups  $\times$  5 major causes of mortality). Furthermore, we 530 explored the interaction between extreme heat and smoke  $PM_{2.5}$  by interacting the current-month 531 number of extreme heat days and bins of 12-month moving average concentration of smoke 532 PM2.5 in the model. In addition, we also conducted stratified analysis by distance to fire (close, 533 far) to investigate its potential modification effect. We tested the lag pattern in the effects of 534 smoke PM<sub>2.5</sub> by extending the 12-month moving average of smoke PM<sub>2.5</sub> concentration to 24 and 535 36, representing longer-term smoke PM2.5 in the past up to 3 years. 536 Based on the estimated coefficients of smoke PM<sub>2.5</sub> bins  $(\beta_b)$  for each cause, which can 537 be interpreted as the changes in mortality rate associated with being in those bins of smoke PM<sub>2.5</sub> 538 concentration compared to the reference bin  $(0-0.1 \,\mu g/m^3)$ , we calculated the number of deaths 539 attributable to smoke PM<sub>2.5</sub> (attributable number, AN) in each county in each month by  $AN_{c,m}$  = 540  $\beta_b \times Population_{c,m}$ , where  $\beta_b$  is the estimated coefficient of the corresponding smoke PM<sub>2.5</sub> bin 541 (b) in that month and Population<sub>c,m</sub> is the total population in county c, month m. This calculation 542 relied on two key assumptions. First, this approach assumes homogeneity within each smoke 543 PM2.5 bin, disregarding the variation within the bin. In this study, the 12-month moving average 544 concentrations of smoke  $PM_{2.5}$  were divided into nine narrowly defined bins, each with a relatively 545 small concentration range. Therefore, the estimate of each bin is likely to well represent the 546 average effects of smoke  $PM_{2.5}$  within that specific concentration range. This assumption was 547 further tested by using finer bins and a non-linear curve to smoke  $PM_{2.5}$  in sensitivity analyses

 (see below; **Fig. S8-S9**). In addition, by using the regression coefficients to estimate attributable 549 deaths, we assume a causal relationship between smoke  $PM_{2.5}$  concentration and mortality rates. Traditionally, health impact assessments for long-term exposure use exposure-response functions estimated from cohort studies. However, to the best of our knowledge, no published 552 cohort studies have reported the association between long-term wildland fire smoke PM<sub>2.5</sub> exposure and cause-specific mortality for the total population across the entire contiguous U.S. Similar to a previous study (61), we applied the estimates from a panel fixed effects model to the calculation of attributable mortality burden. Our statistical model accounted for unmeasured temporal and spatial confounders by the year-month and county fixed effects and unmeasured confounders that vary both over time and space by the space-time interaction terms. Therefore, the observed association is unlikely to be primarily driven by unmeasured confounding factors. Spatial and temporal randomization tests were performed to test this assumption (see below; **Fig. S10**). In addition, unlike cohort studies that commonly focus on people identified by specific characteristics (e.g., older adults), our study design allows us to cover the entire population across all age groups. This makes our estimates more suitable for health impact assessments of the general population.

 Several sensitivity analyses were performed to test the robustness of our results: (a) we used different choices of fixed effects in the model; (b) we additionally adjusted for dew point 566 temperature in the model; (c) we removed non-smoke PM<sub>2.5</sub> or current-month air temperature from the model; (d) an alternative four or six dfs was used in the natural cubic spline of air temperature; (e) we adjusted for 12-month moving average of air temperature instead of current- month air temperature; (f) we used mortality count as the outcome variable and performed a 570 quasi-Poisson model; (g) we used finer bins of smoke  $PM_{2.5}$  concentration; and (h) we used a 571 natural cubic spline with knots at 0, 0.1, 0.3, 0.5, 1, and 5  $\mu$ g/m<sup>3</sup> to model wildland fire smoke PM2.5 concentration.

 In addition, to assess the likelihood of model misspecification influencing our main results, we performed a spatial randomization test and a temporal randomization test. In the

- 575 spatial randomization test, we randomized the smoke PM<sub>2.5</sub> exposure for 2,000 times across
- county while keeping their de facto year-month; in the temporal randomization test, we
- randomized the exposure variable for 2,000 times across year-month while keeping the
- corresponding counties. Such placebo tests are commonly used to detect spatial and temporal
- dependence due to model misspecification in panel models (62, 63).

# **Acknowledgments**

- Research reported in this publication was supported by the National Heart, Lung, And Blood
- Institute of the National Institutes of Health under Award Number R01HL169171. Dr. E. Zang
- received support from the National Institute on Aging (R21AG074238-01), the Research
- Education Core of the Claude D. Pepper Older Americans Independence Center at Yale School
- of Medicine (P30AG021342), and the Institution for Social and Policy Studies at Yale University.
- Y. Ma, Dr. K. Chen, and Dr. Y. Lu acknowledge support from the Yale Planetary Solutions seed
- grant. Dr. Y. Liu received support from the National Institute of Environmental Health Sciences
- (1R01ES034175). The content is solely the responsibility of the authors and does not necessarily
- represent the official views of the National Institutes of Health.

# **References**









 $\Box$  0.09-0.15  $\Box$  0.15-0.25  $\Box$  0.25-0.35  $\Box$  0.35-2.74

- 742 743 **Fig. 1. Spatial distribution of 12-month moving average of smoke PM2.5 concentration in**  744 **the contiguous U.S., 2007–2020. A**: The distribution of the mean 12-month moving average concentration of smoke PM<sub>2.5</sub> in the contiguous U.S. counties (µg/m<sup>3</sup>). **B**: The distribution of the 746 standard deviation (SD) of the 12-month moving average concentration of smoke PM<sub>2.5</sub> in the 747 contiguous U.S. counties ( $\mu q/m^3$ ). contiguous U.S. counties  $(\mu g/m^3)$ .
- 748





750 751 **Fig. 2. Association between 12-month moving average concentration of smoke PM2.5 and**  752 **monthly mortality rate.** This figure displays the estimated changes in monthly cause-specific 753 mortality rate (per 100,000 people) associated with each bin of 12-month moving average of 754 smoke PM<sub>2.5</sub> concentration, relative to a month in the same county with the long-term smoke 754 smoke PM<sub>2.5</sub> concentration, relative to a month in the same county with the long-term smoke 755 PM<sub>2.5</sub> exposure below 0.1  $\mu$ g/m<sup>3</sup>. The error bars indicate 95% confidence intervals. PM<sub>2.5</sub> exposure below 0.1  $\mu$ g/m<sup>3</sup>. The error bars indicate 95% confidence intervals.



- **Fig. 3. Average annual non-accidental mortality burden attributable to long-term smoke**
- **PM2.5 exposure. A**: The spatial distribution of average annual non-accidental deaths (per
- 100,000 people) attributable to long-term smoke PM2.5 exposure (12-month moving average). **B**:
- Estimated average annual cause-specific deaths attributable to long-term smoke PM<sub>2.5</sub> exposure.



764<br>765 765 **Fig. 4. Interaction effects between extreme heat and smoke PM2.5 on non-accidental** 766 **mortality rate. A**: This bivariate choropleth map shows the spatial distribution of the 12-month 767 moving average of smoke PM<sub>2.5</sub> concentrations and the average number of current-month 767 moving average of smoke  $PM_{2.5}$  concentrations and the average number of current-month 768 extreme heat days in each county, 2007-2020. Darker blue indicates higher average smok 768 extreme heat days in each county, 2007-2020. Darker blue indicates higher average smoke PM<sub>2.5</sub><br>769 concentrations: darker red indicates more average extreme heat days: and darker purple 769 concentrations; darker red indicates more average extreme heat days; and darker purple<br>770 indicates higher values of both variables. **B**: The interaction effects between the number o 770 indicates higher values of both variables. **B**: The interaction effects between the number of 771 current-month extreme heat days (a continuous variable) and 12-month moving average of 772 smoke PM<sub>2.5</sub> concentration bins (a categorial variable). The error bars indicate 95% confide 772 smoke PM $_{2.5}$  concentration bins (a categorial variable). The error bars indicate 95% confidence<br>773 intervals. Extreme heat davs were defined as davs with daily mean air temperature higher than 773 intervals. Extreme heat days were defined as days with daily mean air temperature higher than<br>774 the county's 90<sup>th</sup> (main analysis), 85<sup>th</sup>, or 95<sup>th</sup> (sensitivity analyses) percentile warm season air 774 the county's 90<sup>th</sup> (main analysis), 85<sup>th</sup>, or 95<sup>th</sup> (sensitivity analyses) percentile warm season air 775 temperature (May to September, 2007-2020). 776

# 777 **Table 1. Associations between 12-month moving average of smoke PM2.5 concentration**

778 **bins and monthly non-accidental mortality rate (per 100,000 people) in population** 

subaroups





\*The *P* value indicates the statistical significance of between-group difference, with males, people aged 0-64, and non-Hispanic White people as the reference group. We added an interaction term of the subgroup variable and smoke PM2.5 variable into the main model and reported the *P* value of this interaction term.