**Introduction:**

Ambient particulate matter pollution, especially PM2.5 (particulate matter with a diameter smaller than 2.5 microns, as opposed to coarse particulate matter with a diameter smaller than 10 microns), has been robustly linked to a wide variety of negative health effects. Due to their small size, these particles are able to travel deeper into the lungs and enter the bloodstream, at which point they can cause inflammation, affect the cardiovascular system, and even affect the central nervous system after crossing the blood-brain barrier.1 Demonstrated health impacts of ambient PM2.5 include asthma, reduced lung function, increased risk for respiratory infections, cardiovascular disease, diabetes, cancer, and premature mortality.2,3

PM2.5 pollution due to wildfire smoke is a significant issue in the United States and has grown worse in recent years even as there have been overall reductions in all-source PM2.5.4 Over the past decade, wildfire smoke has accounted for roughly a quarter of PM2.5 pollution, and as much as half of all PM2.5 pollution in western states.5 This comes even as the US has made overall reductions in PM2.5 levels. Wildfire smoke PM2.5 is expected to continue harming these efforts, and under climate modeling projections, worsening wildfire seasons may lead to more than a doubling of current smoke PM2.5 levels, which could fully offset successes in PM2.5 reduction in certain regions.6,7 Smoke PM2.5 is not just an issue for the Western US, as smoke plumes can extend far into the midwestern and eastern regions. Some projections estimate that up to three quarters of smoke-related mortality in the US is outside the West, in part due to differences in population density.8

PM2.5 associated with wildfire smoke is understudied compared to all-source PM2.5, and there is evidence that smoke-related PM2.5 may be particularly harmful, as several studies have examined the impact of both smoke and all-source PM2.5 on hospitalizations and found smoke PM2.5 to be associated with larger increases.9,10 Aguilera et al. explain that this association could be due to wildfire smoke’s higher concentration of organic compounds compared to all-source PM2.5, which increases its ability to cause oxidative stress and inflammation.9

Given the possibility that smoke PM2.5 presents a different and perhaps heighted risk profile, and is an increasingly significant problem in the US, it is important to further investigate its effects. Several studies have looked at the impact of smoke PM2.5 on mortality by analyzing the effect of acute wildfire events on daily mortality. Jegasothy et al. and Morgan et al. both studied bushfire smoke PM in Sydney, Australia, at the daily level to find mixed results, as Jegasothy et al. reported a significant positive association in adults over the age of 65, while Morgan et al. did not find an association without mortality despite a significant positive association with hospitalizations.11,12 Chen et al. and Ye et al. used similar approaches, both with daily data, to studying this association in other international contexts, and found significant, positive associations.13,14 To our knowledge, no peer-reviewed papers have studied this relationship at the monthly or yearly level. Compared to the daily level, this wider temporal lens captures medium- and long-term exposures, minimizes concerns around displacement (i.e. smoke days precipitating deaths that would have occurred shortly after, instead of causing deaths in individuals who would have otherwise lived), and perhaps most importantly, provide a template for further research studies that can are more accessible to run given the relative simplicity of accessing less granular data compared to daily-level statistics.

A number of papers, summarized in Table 1, have used this ecological approach to studying the effect of all-source PM2.5 on mortality. In brief, this approach involves using PM data aggregated to the geographic unit level through area- or population-weighting, using annual instead of daily outcome data, and utilizing an analytic strategy based on a quasi-Poisson regression with two-way fixed effects (TWFE) for geographic unit and year. Many other studies in this space use concentration response functions (CRFs), which are based on previous literature that directly estimate the statistical association between air pollution and health, to extrapolate from existing data and calculate the societal impact of air pollution in terms of mortality and economic cost, among other metrics.15,16,17 Those studies are not fundamentally comparable, though they rely on this type of research and are an important way this research can contribute to the field.

TWFE models are a traditionally econometric approach that have increasingly been applied in environmental and social epidemiology.18 These models compare each group to itself across multiple years (achieved by the spatial FE), adjusting for secular trends (the time FE), and therefore are hypothetically able to adjust for both unit-level observable and unobservable confounders that may be time-varying or time-invariant. As a result, they are thought to produce more plausibly causal estimates than traditional models that rely on total confounder adjustment. It is worth noting that these fixed effects come from econometric and not epidemiologic methodology, and are quite different from random effects that are common in epidemiologic mixed models.19 The spatial FE, for example, creates a dummy variable for each county such that only within-unit variation is modeled, as opposed to mixed models using random effects, which leverage both within- and between-unit variation.

Included in Table 1 is Wang et al., who used TWFE models to study the effect of annual PM2.5 on mortality at the census tract level in New Jersey, and clearly explained this analytic approach.20 Many other studies in Table 1 are explicitly modeled after Wang et al., with similar methodology. Though the vocabulary used is different, their approach is functionally identical to the quasi-Poisson regression with TWFE. Differences in terminology may have arisen due to the relative newness of TWFE models in epidemiology. Many of the papers in Table 1 source their methodology to Armstrong et al., which demonstrated how conditional Poisson models run using the `gnm’ package in R can be used to analyze panel data in epidemiologic contexts.21 As confirmed by our analyses, this approach is functionally identical to running a quasi-Poisson regression with TWFE using the `fixest` package in R, which is a newer package designed specifically for this approach and is more computationally efficient, though `fixest` is for now more commonly used in econometric studies. A major drawback of the `gnm` package is the lack of transparency around which standard errors (SEs) are used, unlike `fixest.` SE choice has an enormous impact on the statistical inference of model results, and the implications of this will be discussed later in this paper.

The studies in Table 1 are fundamentally quite similar, but they do frequently make different design choices at several stages. Aggregating environmental data to the chosen spatial level of analysis can be done through different methods, typically either weighted by area or by population, which can create nontrivial differences. TWFE models minimize the number of covariates that should be included, but they cannot control for variables that vary across time and space, such as temperature, which is frequently included as a covariate in these models. Temperature can be modeled in a few different ways and is mostly modeled linearly or using natural cubic splines with different degrees of freedom. Other common covariates include precipitation and economic output. Other modeling choices include FE specifications, weighting the regression by population, and SE choice, among other possibilities. All these choices can be reasonably made *a priori*, though they involve different assumptions and may lead to nontrivial differences in the research finding. We believe that the research methodology on air pollution and mortality studied at the non-daily level is well characterized by Gelman and Loken’s “garden of forking paths,” which describes how researcher degrees of freedom can lead to differing conclusions even as reasonable decisions are made at each step.22

This paper analyzes the effect of wildfire smoke PM2.5 on all-cause mortality in the contiguous US at the county-month level by systematically testing different model specifications. Our goal is to analyze how modeling choices affect the observed association and visualize all so-called forking paths. We use a quasi-Poisson regression with FEs as has been previously done, though we use the `fixest` package, which is newer and more flexible for these analyses than the `gnm` package. Given the variety of different model specifications that have been used in similar analyses, we systematically vary these choices to determine the range of estimates that can be obtained through a priori reasonable modeling approaches. We then recommend a set of specifications and several best practices for future studies in this area.

**Methods:**

**Estimating wildfire smoke PM2.5 exposure:**

Wildfire smoke PM2.5 data come from Childs et al. 2022, which produced 10 km x 10km gridded estimates for the contiguous US for all days, 2006-2020.23 In brief, their methods involved determining which days had wildfire smoke plumes overhead based on satellite imagery. For such “smoke days,” PM2.5 anomalies from EPA ground stations were attributed to wildfire smoke PM2.5, and the smoke-attributable PM2.5 value was calculated as the difference between the anomalous value and the 3-year non-smoke day median. A machine learning model was used to predict ground station values, including data on meteorological measures, aerosol optical depth predictions, HYSPLIT trajectory points, and topological data, among other sources. After training, this ML model was used to generate gridded, smoke-attributable PM2.5 predictions for the entire country. These data have already been used to measure exposure in environmental epidemiology studies such as Wen & Burke.24

Population-weighted aggregation was used to take the daily, 10km2 gridded smoke PM2.5 estimates from Childs et al. to the county-month level. Population data came from WorldPop’s 2013 estimates, which provide values of population density at the 10km2 level. The proportion of each county that overlapped with a grid cell was calculated, population densities were combined with these areas to determine the total population of each county-grid cell overlap, and then the mean smoke PM2.5 estimate across all overlapping grid cells was calculated, weighted by grid cell population. Population weighting was used because counties often cover large tracts of land and do not have an evenly distributed population, so area-weighting estimates may misrepresent the population’s true exposure to smoke PM2.5 or other environmental variables.

**Mortality and population data:**

Mortality data were obtained from the National Center for Health Statistics National Vital Statistics Service’s 1999-2020 multiple cause of death files (MCOD).25 These are decedent-level data derived from all death certificates in the US, with information on primary cause of death, additional non-primary causes, and demographic data. Aggregated versions of these data are publicly available through the CDC’s Wide-ranging Online Data for Epidemiology Research (WONDER) web portal; however, restricted-access data are necessary for these analyses, which require granular county-month-level counts, including cells with fewer than 10 events.

The primary outcome in this study is all-cause mortality. We included months between January 2006 and February 2020, excluding months after that point to avoid data irregularities due to the COVID-19 pandemic. This 15-year period includes many major wildfire events and captured the trend of increasing wildfire smoke exposure in the US, which is expected to continue in the coming decade.

**Environmental control data:**

Through its use of FEs, our modeling approach controls for all spatially invariant time-varying and time-invariant spatially-varying confounders, but does not control for confounders that vary across both space and time simultaneously. Consistent with previous literature, we assume only temperature and precipitation meet these criteria and thus require direct adjustment. Monthly mean temperature and precipitation data at the 4km x 4km level were obtained from Oregon State’s PRISM Climate group and aggregated to the county level by both population- and area-weighting. Other potential confounders do not need to be controlled for due to the nature of the analytic approach, and the relative infrequency with which existing studies directly control for additional covariates.

**Estimating the effect of wildfire smoke PM2.5 on mortality:**

Our models utilized a quasi-Poisson regression with two-way fixed effects. Quasi-Poisson regression was chosen because our total mortality is a count variable, though the model results can be interpreted as an effect on the mortality rate due to the offset term. We did not choose negative binomial regression because, although both models are capable of handling overdispersed data, negative binomial regression gives outsized weight to smaller counties, while quasi-Poisson regression weights counties by size more evenly.26 Our analyses involved running many regressions with slight variations in specification, so it would be impractical to enumerate them all. Equations 1-3 show the three sets of FEs we studied, along with our recommended approach to modeling covariates, though the covariate selection is varied in our main analysis.

Text, letter

Description automatically generated

All three equations model the natural logarithm of mortality counts for each county *c*, month *m*, and year *y*. The difference between the three equations is in the fixed effects and their implicit assumptions.

Equation 1 uses a county-calendar month FE, ˙˙ƒ

*ηc-m*,and a year FE, *ƛy*. The county-calendar month FE creates dummy variables for each county-calendar month in our sample (in other words, 3,083 counties \* 12 calendar months = 36,996 intercepts). The potential value of this FE instead of the more common county-only FE is that different counties may have different seasonal effects for variables unmeasured in our regression. Like with other FEs that incorporate county, this FE means the model compares each county’s mortality counts to themselves over time, theoretically eliminating the need to control for both measured and unmeasured year-invariant county-month-level confounders such as socioeconomic status, urbanicity, and so forth. The year FE *ƛy* creates dummies for each year of our sample (14 dummies total for years 2006-2020), controlling for nationwide secular trends.

Equation 2 is substantially similar to Equation 1 but uses a county FE, *ηc* (creating 3,083 intercepts, one for each county), and a year-month FE, *ƛy-m* (creating 12 calendar months \* 14 years + 2 months of 2020 = 170 intercepts). This set of FEs compares each county to itself across all months in all years, adjusting for secular trends by each month of each year. This approach does not adjust for seasonality.

Equation 3 uses a county FE, *ηc* (creating 3,083 intercepts, one for each county), and a year FE, *ƛy* (creating 15 intercepts, one for each year). It is conceptually the simplest, comparing each county to itself across all months in all years, adjusting for secular trends at the yearly level.

Other parameters in Equations 1-3 vary slightly across the regressions we ran in our analysis, but they are substantially similar, so they are explained here. *SmokePMc, m, y* represents the mean daily wildfire smoke-attributable PM2.5 pollution in county *c* during month *m* of year *y*, and is our main variable of interest. We used no lag terms, so this regression models the effect of same-month smoke PM2.5 on mortality. *Tempc,m,y*represents the mean monthly temperature of each county, and a natural cubic spline with 3 degrees of freedom is applied to account for nonlinear effects of temperature on mortality. Similarly, *Precipc,m,y* represents the mean precipitation for each county, modeled linearly. The offset term *ln(Popc,y)* represents the total population of each county *c*  in each year *y*, and it is included so that the model results can be interpreted as an effect on mortality rates instead of on raw mortality counts. εc,m,y represents the error term. β*1* represents the average effect of an additional µg/m3 of average monthly smoke PM2.5, though it must be exponentiated to be interpretable as a relative increase. The model was weighted by county population size. Robust standard errors were used, and they were clustered at the county level to account for within-unit correlation. Modeling was done using the `fixest` package in R 4.2.0. Tests for significance were two-tailed.

None of these models can control for confounders that vary across both time and county. We chose to directly adjust for temperature and precipitation because they clearly vary across both county and time, and are known to be associated with wildfire risk, ambient PM2.5, and mortality.27–30 Other potential confounders are not directly measured, and represent a possible source of bias.

While the offset term does allow the model coefficients to be interpreted as affecting the mortality rate instead of the raw death counts, this effect on the mortality rate is not age-adjusted, strictly speaking. However, the county-month fixed effect accounts for county-level population age structures, and the time fixed effect accounts for nationwide trends in aging. These combined effects minimize the concern of confounding by population age structure, because in order for it to be a confounder, county age structure would have to vary year-to-year in a way that is correlated with changes in smoke PM2.5 and is not captured in national trends, which is not highly plausible.

**Results:**

Our analysis using the CDC MCOD data included 36,775,902 all-cause deaths from January 2006 through February 2020. The distribution of mortality counts and age-adjusted rates, along with wildfire smoke PM2.5, temperature, and precipitation data are summarized in Table 2. These values were calculated by averaging data across all monthly timepoints for each county, such that each number represents the average per-month value across all months in the sample. The mean monthly county-level age-adjusted death rate was 69 deaths per 100,000 individuals. The mean county-level smoke PM2.5 value was 0.41 µg/m3, though this distribution was highly right-skewed. The mean county-level temperature was 12.79 oC, and the mean precipitation level was 86.18 mm.

Figure 1 shows the spatial and temporal distributions of wildfire smoke pollution and mortality in the contiguous US during the study period. Wildfire smoke pollution is widespread throughout the US and is not localized just to the Western states, as both the Midwest and Southeast are shown to experience many months with average smoke PM2.5 values over 0.5 µg/m3 (Fig. 1a). The distribution of smoke PM2.5 is highly right-skewed, with over 98% of county-months experiencing less than 5 µg/m3, though there is no known safe level of PM2.5 (Fig. 1e). Smoke pollution varies significantly both within years and across years (Fig. 1b). It tends to be lower during the fall and winter months while worsening during the spring and summer. In our sample, the summers of 2007-08, 2011-12, and 2017-18 were among the most severe wildfire seasons. Mortality rates also vary significantly by region and by season, though less so across years (Fig. 1c, 1d). Counties with high mortality rates tend to be clustered in the South and Southeast. Temporally, mortality rates tend to worsen in the fall and winter months.

Our main analysis tested 75 total regressions with both IID SEs and heteroskedasticity-robust SEs clustered at the county level (Fig. 2). 56 out of the 75 models had negative point estimates, suggesting that mortality rates decrease as smoke PM2.5 increases at the county-month level. When using IID SEs, 59 out of 75 models found a statistically significant association. However, when using robust SEs, eight out of 75 models produce statistically significant results. The regressions with the 5 lowest point estimates (Fig. 2b) were an order of magnitude lower than the remaining 70 estimates (Fig. 2a). All five of these regressions used FEs for county and year, and did not control for monthly temperature. Other models with FEs for county and year that controlled for temperature in some form produced results consistent with models using other FEs. Regressions using county-calendar month and year FEs all produced negative point estimates, while regressions using county and year-month FEs produced 21 negative estimates and four positive estimates. Regressions using county and year FEs produced 10 negative and 15 positive point estimates, including the five lowest and 12 highest estimates.

Controlling for temperature made a large difference in estimated effect size and direction, though this effect was altered by choice of FEs. For all three sets of FEs, models that did not control for temperature tended to produce lower point estimates, and this effect was most pronounced in the models with county and year FEs. The specific modeling choice for temperature did not appear to make a major difference as long as it was included in some form. Modeling choices for precipitation did not appear to have a major impact on model results. Even complete exclusion of precipitation as a covariate did not make a major impact on point estimates or SEs for any model tested.

Our preferred model, indicated with a red point estimate in Fig. 2a, produces a point estimate of -.000519, or -.0519%, indicating that a 1µg/m3 increase in monthly smoke PM2.5 is associated with a -.0519% decrease in the all-cause mortality rate. This result is nonsignificant with robust SEs (p = .36) and highly significant with IID SEs (p < .001).

Our second analysis ran our preferred model specification while varying the range of years included to test the sensitivity of these analysis to the window of the study period (Fig. 3). The first part of this analysis ran the same regression using all 11 possible four-year spans within our study window (Fig. 3a). All but the two most recent time periods are negative point estimates. Eight of the 11 are statistically significant under the IID assumption, while two are significant when using robust SEs. The size of the SEs also varies substantially by time window, narrowing considerably in the later periods. The second part of this analysis iteratively added years to the initial four-year span (Fig. 3b). All point estimates are negative, and trend toward zero as years are added. All estimates are statistically significant under the IID assumption, though only the first 3 estimates are significant when using robust SEs. The last part of this analysis stratified these regressions by age (Fig. 3c,d). In Fig. 3c, 9 of the 11 regressions in the 65 plus group produce negative point estimates, compared to 4 of the 11 regressions in the under 65 group. In Fig. 3d, all regressions in the 65 plus group produce negative point estimates compared to seven in the under 65 group.

**Discussion:**

This study tested the sensitivity of quasi-Poisson TWFE models to changes in FEs, covariates, and study window. Our primary analysis found these models to be highly sensitive to choices in FEs and inclusion vs. exclusion of temperature as a covariate, and mostly insensitive to modeling choices regarding temperature and inclusion vs. exclusion of precipitation as a covariate. Choice of SEs also makes a major difference, as robust SEs were observed to be much larger than IID SEs. Our secondary analysis found that the years included in ecological studies can also have a significant impact on the association estimated.

The results of our main analysis show that different yet seemingly reasonable choices in FEs can lead to widely disparate results, especially when combined with other modeling decisions. The estimates from regressions using FEs for county and year, without temperature controls, are the clearest example of this (Fig. 2b). Regressions with county and year FEs were also less stable than other FE choices even when temperature was accounted for, as they produced all of the highest point estimates. This instability likely reflects the comparative lack of precision of the time FE compared to the other choices, which each accounted for sub year-level effects.

Models using county and year-month FEs or county-calendar month and year FEs produced more consistent estimates across the range of covariate specifications, though a large degree of variability remained. County-calendar month and year FE models produced the most consistent estimates of the three FE sets tested, and all estimates were negative. County and year-month FEs were next most stable, but still produced both positive and negative estimates depending on covariate specifications. County and year-month FEs are perhaps the most straightforward choice of FEs because it is the simplest choice of a spatial FE and a time FE when using monthly data, and is most closely analogous to the studies in Table 1 that use a year FE for their annual data. Ma et al. was released in preprint in medRxiv in February 2023, and studied the effect of wildfire smoke PM2.5 on mortality using the same CDC WONDER mortality data and wildfire smoke PM2.5 data from Childs et al. as our study, though they only analyzed the 2006-2016 range.31 They chose to use a year-month time FE and found a positive point estimate for the smoke PM2.5-mortality association, in line with our findings. The limitation of this straightforward time FE is that it fails to account for seasonality, which is potentially important to account for given the demonstrated seasonality of wildfire smoke exposure. It is especially important to adjust for seasonality when analyzing a large geographic region like the US because different areas are known different seasonal effects for all-source PM2.5, and this may plausibly be the case for other variables as well. Therefore, a county-calendar month FE prevents these differential effects from biasing the overall estimate.

Choosing the correct SEs were shown to be quite important for calculating the correct p-value. Our analyses presented in Fig. 2 and Fig. 3 display both the IID SEs and robust SEs, and IID SEs are consistently shown to be dramatically smaller, with major implications for the statistical significance of the regression’s effect estimate. These IID SEs are calculated under the assumption of independent and identically distributed observations. The IID assumption is almost certainly an incorrect one in the context of using panel data such as ours. Even at the monthly level there is likely to be correlation between months, and there is almost certainly correlation between nearby counties. Still, the IID assumption often underpins the default SE estimates in statistical packages, and can be a common oversight for researchers.32

Given that the IID assumption is often violated, especially in panel data settings, robust SEs are typically a better choice. Our analyses show that robust SEs were much larger than IID SEs such that statistical significance was frequently achieved with near-zero p-values when IID SEs were used, while use of robust SEs suggested the effect estimate did not even approach statistical significance. This difference is crucial, and not simply an academic distinction, as it is not uncommon for research studies to mistakenly use IID SEs. Indeed, our replication of the main findings in Ma et al. suggest that they used the IID assumption to calculate their SEs and p-values. When we used robust SEs instead, several of the effect estimates were no longer statistically significant, and all of the p-values were much larger. Furthermore, when we used the `gnm` package instead of `fixest`, we were able to replicate several of our point estimates exactly, but the SEs were significantly narrower. We failed to exactly replicate the SEs produced by the model in `gnm` by using IID SEs in `fixest`, though they were substantially similar. Given that the `gnm` package appears to default to using IID SEs without a clear choice to specify robust SEs, and that the papers highlighted in Table 1 use `gnm` without specifying robust SEs, it appears likely that many if not all of these studies relied on the IID assumption. None made their code or data publicly available, so we were unable to replicate findings outside of Ma et al. Based on the large difference in SE size that we found, it is potentially likely that the results of these studies are sensitive to using robust, clustered SEs.

The range of years included in a study also appears to be a major factor in what findings it generates. Our analyses that varied the window of time we used to run our regressions suggest there may not be a consistent, year-to-year association between wildfire smoke and mortality. If this is the case, then it follows that studies covering only a handful of years cannot reliably produce a generalizable effect estimate. Our analysis in Fig. 3b appears to suggest that adding years to the sample refines the effect estimate, as the initially strongly negative estimates increase toward zero as more years are included. However, this dynamic is not uniformly true, and can depend on model specifications. When we sought to replicate the findings in Ma et al., which used county and year-month FEs, we found an initially statistically significant result using robust SEs, which became nonsignificant after including the entire 2006-2019 range (Ma et al. only included 2006-2016). These findings highlight the difficulty of pinning down a consistent effect estimate using non-daily ecological data, and may help explain challenges to reproducing similar effect estimates across the literature partially captured in Table 1.

The following is our recommendation for future studies using monthly or yearly ecological data to estimate the effect of wildfire smoke PM2.5 on mortality with TWFE models. First, regarding FEs, the spatial FE should be chosen based on the most granular geographic unit available, in our case the county level. The time FE should incorporate seasonality, especially if the geographic range studied is wide enough that there may be different seasonal effects by region. Second, regarding environmental covariates, temperature should be controlled for in some manner, but the specific way it is represented is relatively unimportant. We lightly recommend using population-weighted values along with a natural cubic spline to account for a nonlinear response function. Controlling for precipitation did not appear to make a meaningful impact, though it could theoretically affect the relationship. Third, robust SEs should be used. Using IID SEs can lead to confidence intervals that are far too narrow and p-values that are far too low. We recommend using the `fixest` package due to the simplicity of specifying SE methods, in addition to its overall computational efficiency. Fourth, code should be published, and data made publicly available when possible. Few studies in this area of research are readily replicable, and code sharing would help increase transparency.

Our preferred model, using FEs for county-calendar month and year, controlling for temperature with population-weighted values in a natural cubic spline with three degrees of freedom, and controlling for precipitation with a linear term, is indicated in red in Fig. 2a. It produced a negative point estimate that was not statistically significant when using robust SEs. This model, when applied over varying year ranges in the analyses presented in Fig. 3, produced both positive and negative point estimates but mostly found nonsignificant, negative associations. Finding a negative association between wildfire smoke pollution and mortality is somewhat surprising and inconsistent with literature that studied this association using daily time series data and reliably found positive associations. We theorize several possible explanations for our findings. First, they may be explained through behavioral modifications made in the presence of wildfire smoke. People may choose to remain indoors and avoid exposure to smoky air, though pollution infiltrating indoor spaces is an issue that somewhat undermines the credibility of this explanation.33,34 Still, previous research has demonstrated that wildfire smoke leads to observable decreases in short-term ED visits, drives increased internet activity related to air quality, and leads to more time spent at home.35,36 Second, our findings may reflect a lack of sufficient temporal granularity. The largest studies of the association between wildfire smoke pollution and mortality, namely Chen et al. and Ye et al., used daily data on smoke PM2.5 and mortality. Major studies on all-source PM2.5 and mortality have also relied on daily data, or tracked individuals through a cohort design.37,38 It is possible that studying this association at the monthly level obscures the true effect, perhaps because smoke PM2.5 may only be significantly elevated for a handful of days per month. Lastly, it is possible that our results reflect random chance. Our confidence intervals are wide and therefore we cannot rule out the possibility that the true association is positive.

Our study has several important strengths. First, we applied a straightforward and accessible study design to a pertinent public health issue. Many of the foundational papers establishing the association between air pollution and mortality have relied on cohort designs or daily time-series data, but these designs are more time- and resource-intensive than studies that rely on monthly or yearly data across a longer period. If shown to be efficacious, this design is potentially faster and simpler for studying not just air pollution and mortality, but potentially other health outcomes as well. Other potential advantages of this approach are that it captures the entire population, unlike smaller-scale cohort studies that may have issues adequately representing minority groups, or larger-scale Medicare cohort studies that only include people 65 and over; it can capture a wide geographic range, since many studies using daily mortality data are only able to access said data for a select group of municipalities or regions; and it reduces concerns that displacement is driving the observed association, a common issue with daily-level studies.

Second, our analysis addressed the concern that different modeling choices make a large impact on the observed association between wildfire smoke PM2.5 pollution and mortality by systematically varying model specifications that have been observed in published literature or have reasonable conceptual grounding. We tested all possible combinations of chosen model specifications and displayed the results. This approach highlights the instability of these TWFE models when applied to data at the monthly level, and the potential inadequacy of sensitivity analyses that only adjust modeling choices slightly. For example, few studies include sensitivity analysis that adjust FEs, though we demonstrated that conceptually similar FEs can greatly change results.

Third, our study period of 15 years was long compared to similar papers. In addition to increasing statistical power simply due to access to more years of data, this allowed us to observe how studying different temporal windows can also change the observed association.

Lastly, the quality of our data was high. The CDC’s MCOD data are a census of mortality counts from US residents and there is minimal concern about measurement error. For our analysis, the main source of measurement error attributable to the MCOD data is that there may be individuals for whom their county of death is not where they spent most of the month, and thus their smoke PM2.5 estimate is misattributed. However, this issue is unlikely to arise on any major scale, and in any case would bias results toward the null. The wildfire smoke PM2.5 estimates from Childs et al. were rigorously developed and validated, have been used in several published analyses to this point, and have been independently validated by Ma et al.

Our study also has several important limitations. First, the TWFE approach has been increasingly scrutinized in recent years, in part over the key assumption that there are no confounders that are both time- and spatially-varying. In our context, this assumption is clearly violated by both temperature and precipitation, which we directly adjusted for, though we assumed there were no other such confounders. This is unlikely to be the case, as other variables, like flu trends, can also vary across both time and space while impacting mortality. Other criticisms of TWFE models include concerns over their true ability to isolate both temporal and spatial variation instead of simply combining them (and thus their intepretability), and whether they are appropriate to apply in panel data contexts outside of the traditional difference-in-difference setting that involves two groups and two time periods.39,40

Second, mortality is an outcome far downstream from initial health impacts of wildfire smoke pollution, and it is a rare outcome compared to other observable health effects. It is therefore more difficult to find an association between wildfire smoke PM2.5 and mortality even though there are demonstrable health impacts of wildfire smoke. More sensitive health outcomes like ED visits may be better suited to studying the health impacts of smoke pollution. However, time series studies using daily mortality data have shown an association between wildfire smoke pollution and mortality, albeit in different global contexts, so it is possible that our mixed findings are more reflective of our temporal granularity than of mortality as an outcome.

Third, our analysis of how varying model specifications affects measured associations did exclude several possible covariates, and we cannot definitively conclude that they would not have also made a difference. For example, some studies have chosen to adjust for all-source PM2.5, or GDP per capita. We chose not to include all-source PM2.5 in part because we were able to replicate Ma et al.’s main regression, which does include all-source PM2.5 as a covariate, without including it ourselves. Additionally, Ma et al. ran a sensitivity analysis that excluded all-source PM2.5 and concluded it was not substantially different than their main regression. We did not include GDP per capita because the spatial FE should adjust for the most pertinent regional differences in economic output. While we believe these choices were reasonable, we are still limited in the model specifications and covariates we could study and thus the conclusions we can draw regarding the sensitivity of this modeling approach to different choices.

**Conclusions:**

` Wildfire smoke PM2.5 pollution is already a major issue in the US that is not limited to the Western states, and it is projected to worsen in future years, threatening improvements in all-source PM2.5. There is a robust literature on the negative health effects of all-source PM2.5, including its effects on mortality, but smoke-specific PM2.5 has been understudied. Existing research on the association between PM2.5 (smoke-specific or all-source) and mortality has largely utilized daily-level mortality data, though a number of studies have also sought to estimate the association using yearly data. This non-daily approach has the potential benefits of being faster and less resource-intensive to utilize, and easier to cover a broad study population in terms of demographics like age and race/ethnicity, geographic area, and time period, which increases generalizability. However, it is unclear the extent to which effect estimates are impacted by different modeling choices under this research strategy.

Our study therefore sought to systematically vary modeling choices to determine the sensitivity of this research approach to diverging yet reasonable assumptions. We found that choice of FEs matters a great deal, as does controlling for temperature, though the specific way temperature was modeled was relatively trivial, as was controlling for precipitation. Our recommended model uses FEs for county-calendar month, in order to adjust for region-specific seasonal effects, and year; it models temperature using a natural cubic spline with three degrees of freedom, and precipitation with a linear term. Smoke PM2.5, temperature, and precipitation are all aggregated to the county level using population weighting. Robust SEs are used to estimate the confidence interval and statistical significance, because the IID assumption is not met.

The estimate preferred model was negative, suggesting that higher values of monthly smoke PM2.5 decrease monthly mortality, but not statistically significant, and the confidence interval was too wide to draw definitive conclusions. In general, the wide range of estimates produced by our analysis implies, first of all, that this research strategy is sensitive to different modeling decisions and thus reproducibility may be challenging. Second, given the well-studied positive association between all-source PM2.5 and mortality, and the small but strong body of evidence for a positive association between wildfire smoke PM2.5 and mortality, it is most plausible that the lack of a definitive finding from our analysis is reflective of a lack of temporal granularity. Simply put, mortality at the monthly level is likely too insensitive to monthly mean smoke PM2.5 to capture the true association with any sort of precision. Even in bad months, wildfire smoke PM2.5 is often only severely elevated for a few days, so a true uptick in deaths due to those high-smoke days can be diluted by random variation throughout the rest of the month.

Our study set out to investigate the impact of different reasonable modeling choices on effect estimates for the association between wildfire smoke PM2.5 and mortality at the monthly level and recommend a set of best practices based on those results. We stand by our recommendations, though perhaps our most important conclusion is that daily level data are needed to estimate the effect of air pollution on mortality, even with high quality sources and a wide spatiotemporal range.

**References:**

1. Shou Y, Huang Y, Zhu X, Liu C, Hu Y, Wang H. A review of the possible associations between ambient PM2.5 exposures and the development of Alzheimer’s disease. *Ecotoxicol Environ Saf*. 2019;174:344-352. doi:10.1016/j.ecoenv.2019.02.086

2. Kim KH, Kabir E, Kabir S. A review on the human health impact of airborne particulate matter. *Environ Int*. 2015;74:136-143. doi:10.1016/j.envint.2014.10.005

3. Reid CE, Brauer M, Johnston FH, Jerrett M, Balmes JR, Elliott CT. Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environ Health Perspect*. 2016;124(9):1334-1343. doi:10.1289/ehp.1409277

4. US EPA O. Particulate Matter (PM2.5) Trends. Published July 19, 2016. Accessed April 20, 2023. https://www.epa.gov/air-trends/particulate-matter-pm25-trends

5. Burke M, Driscoll A, Heft-Neal S, Xue J, Burney J, Wara M. The changing risk and burden of wildfire in the United States. *Proc Natl Acad Sci*. 2021;118(2):e2011048118. doi:10.1073/pnas.2011048118

6. Ford B, Val Martin M, Zelasky SE, et al. Future Fire Impacts on Smoke Concentrations, Visibility, and Health in the Contiguous United States. *GeoHealth*. 2018;2(8):229-247. doi:10.1029/2018GH000144

7. Neumann JE, Amend M, Anenberg S, et al. Estimating PM2.5-related premature mortality and morbidity associated with future wildfire emissions in the western US. *Environ Res Lett*. 2021;16(3):035019. doi:10.1088/1748-9326/abe82b

8. O’Dell K, Bilsback K, Ford B, et al. Estimated Mortality and Morbidity Attributable to Smoke Plumes in the United States: Not Just a Western US Problem. *GeoHealth*. 2021;5(9). doi:10.1029/2021GH000457

9. Aguilera R, Corringham T, Gershunov A, Benmarhnia T. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. *Nat Commun*. 2021;12(1):1493. doi:10.1038/s41467-021-21708-0

10. Liu JC, Wilson A, Mickley LJ, et al. Wildfire-specific Fine Particulate Matter and Risk of Hospital Admissions in Urban and Rural Counties. *Epidemiology*. 2017;28(1):77-85. doi:10.1097/EDE.0000000000000556

11. Jegasothy E, Hanigan IC, Van Buskirk J, et al. Acute health effects of bushfire smoke on mortality in Sydney, Australia. *Environ Int*. 2023;171:107684. doi:10.1016/j.envint.2022.107684

12. Morgan G, Sheppeard V, Khalaj B, et al. Effects of Bushfire Smoke on Daily Mortality and Hospital Admissions in Sydney, Australia. *Epidemiology*. 2010;21(1):47-55. doi:10.1097/EDE.0b013e3181c15d5a

13. Chen G, Guo Y, Yue X, et al. Mortality risk attributable to wildfire-related PM2·5 pollution: a global time series study in 749 locations. *Lancet Planet Health*. 2021;5(9):e579-e587. doi:10.1016/S2542-5196(21)00200-X

14. Ye T, Xu R, Yue X, et al. Short-term exposure to wildfire-related PM2.5 increases mortality risks and burdens in Brazil. *Nat Commun*. 2022;13(1):7651. doi:10.1038/s41467-022-35326-x

15. Matz CJ, Egyed M, Xi G, et al. Health impact analysis of PM2.5 from wildfire smoke in Canada (2013–2015, 2017–2018). *Sci Total Environ*. 2020;725:138506. doi:10.1016/j.scitotenv.2020.138506

16. Johnston FH, Borchers-Arriagada N, Morgan GG, et al. Unprecedented health costs of smoke-related PM2.5 from the 2019–20 Australian megafires. *Nat Sustain*. 2020;4(1):42-47. doi:10.1038/s41893-020-00610-5

17. Fann N, Alman B, Broome RA, et al. The health impacts and economic value of wildland fire episodes in the U.S.: 2008–2012. *Sci Total Environ*. 2018;610-611:802-809. doi:10.1016/j.scitotenv.2017.08.024

18. Strumpf EC, Harper S, Kaufman JS. FIXED EFFECTS AND DIFFERENCE-IN-DIFFERENCES.

19. Gunasekara FI, Richardson K, Carter K, Blakely T. Fixed effects analysis of repeated measures data. *Int J Epidemiol*. 2014;43(1):264-269. doi:10.1093/ije/dyt221

20. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating Causal Effects of Long-Term PM 2.5 Exposure on Mortality in New Jersey. *Environ Health Perspect*. 2016;124(8):1182-1188. doi:10.1289/ehp.1409671

21. Armstrong BG, Gasparrini A, Tobias A. Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis. *BMC Med Res Methodol*. 2014;14(1):122. doi:10.1186/1471-2288-14-122

22. Gelman A, Loken E. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “ﬁshing expedition” or “p-hacking” and the research hypothesis was posited ahead of time.

23. Childs ML, Li J, Wen J, et al. Daily Local-Level Estimates of Ambient Wildfire Smoke PM 2.5 for the Contiguous US. *Environ Sci Technol*. 2022;56(19):13607-13621. doi:10.1021/acs.est.2c02934

24. Wen J, Burke M. Lower test scores from wildfire smoke exposure. *Nat Sustain*. 2022;5(11):947-955. doi:10.1038/s41893-022-00956-y

25. Multiple Cause of Death 1999-2020. Accessed February 28, 2023. https://wonder.cdc.gov/wonder/help/mcd.html

26. Ver Hoef JM, Boveng PL. QUASI-POISSON VS. NEGATIVE BINOMIAL REGRESSION: HOW SHOULD WE MODEL OVERDISPERSED COUNT DATA? *Ecology*. 2007;88(11):2766-2772. doi:10.1890/07-0043.1

27. An H, Gan J, Cho S. Assessing Climate Change Impacts on Wildfire Risk in the United States. *Forests*. 2015;6(12):3197-3211. doi:10.3390/f6093197

28. Basu R. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environ Health*. 2009;8(1):40. doi:10.1186/1476-069X-8-40

29. Basu R. Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence. *Epidemiol Rev*. 2002;24(2):190-202. doi:10.1093/epirev/mxf007

30. Tai APK, Mickley LJ, Jacob DJ. Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to climate change. *Atmos Environ*. 2010;44(32):3976-3984. doi:10.1016/j.atmosenv.2010.06.060

31. Ma Y, Zang E, Liu Y, et al. *Wildfire Smoke PM 2.5 and Mortality in the Contiguous United States*. Occupational and Environmental Health; 2023. doi:10.1101/2023.01.31.23285059

32. Inlow M. Teaching The Iid Assumption In Engineering Statistics I. In: *2006 Annual Conference & Exposition Proceedings*. ASEE Conferences; 2006:11.1230.1-11.1230.7. doi:10.18260/1-2--821

33. Shrestha PM, Humphrey JL, Carlton EJ, et al. Impact of Outdoor Air Pollution on Indoor Air Quality in Low-Income Homes during Wildfire Seasons. *Int J Environ Res Public Health*. 2019;16(19):3535. doi:10.3390/ijerph16193535

34. Liang Y, Sengupta D, Campmier MJ, Lunderberg DM, Apte JS, Goldstein AH. Wildfire smoke impacts on indoor air quality assessed using crowdsourced data in California. *Proc Natl Acad Sci*. 2021;118(36):e2106478118. doi:10.1073/pnas.2106478118

35. Heft-Neal S, Gould C, Childs M, et al. *Behavior Mediates the Health Effects of Extreme Wildfire Smoke Events*. National Bureau of Economic Research; 2023:w30969. doi:10.3386/w30969

36. Burke M, Heft-Neal S, Li J, et al. Exposures and behavioural responses to wildfire smoke. *Nat Hum Behav*. 2022;6(10):1351-1361. doi:10.1038/s41562-022-01396-6

37. Zanobetti A, Schwartz J. The Effect of Fine and Coarse Particulate Air Pollution on Mortality: A National Analysis. *Environ Health Perspect*. 2009;117(6):898-903. doi:10.1289/ehp.0800108

38. Schwartz J, Wei Y, Yitshak-Sade M, Di Q, Dominici F, Zanobetti A. A national difference in differences analysis of the effect of PM2.5 on annual death rates. *Environ Res*. 2021;194:110649. doi:10.1016/j.envres.2020.110649

39. Kropko J, Kubinec R. Interpretation and identification of within-unit and cross-sectional variation in panel data models. Al-Ameri T, ed. *PLOS ONE*. 2020;15(4):e0231349. doi:10.1371/journal.pone.0231349

40. Imai K, Kim IS. On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data. *Polit Anal*. 2021;29(3):405-415. doi:10.1017/pan.2020.33

41. Renzi M, Forastiere F, Schwartz J, Davoli M, Michelozzi P, Stafoggia M. Long-Term PM10 Exposure and Cause-Specific Mortality in the Latium Region (Italy): A Difference-in-Differences Approach. *Environ Health Perspect*. 2019;127(6):067004. doi:10.1289/EHP3759

42. Yu P, Xu R, Li S, et al. Exposure to wildfire-related PM2.5 and site-specific cancer mortality in Brazil from 2010 to 2016: A retrospective study. Hanna L, ed. *PLOS Med*. 2022;19(9):e1004103. doi:10.1371/journal.pmed.1004103

43. Fan Z, Li Y, Wei J, et al. Long-term exposure to fine particulate matter and site-specific cancer mortality: A difference-in-differences analysis in Jiangsu province, China. *Environ Res*. 2023;222:115405. doi:10.1016/j.envres.2023.115405

44. Yu W, Guo Y, Shi L, Li S. The association between long-term exposure to low-level PM2.5 and mortality in the state of Queensland, Australia: A modelling study with the difference-in-differences approach. Kirrane E, ed. *PLOS Med*. 2020;17(6):e1003141. doi:10.1371/journal.pmed.1003141

45. Leogrande S, Alessandrini ER, Stafoggia M, et al. Industrial air pollution and mortality in the Taranto area, Southern Italy: A difference-in-differences approach. *Environ Int*. 2019;132:105030. doi:10.1016/j.envint.2019.105030

46. Yu P, Xu R, Li S, et al. Loss of life expectancy from PM2.5 in Brazil: A national study from 2010 to 2018. *Environ Int*. 2022;166:107350. doi:10.1016/j.envint.2022.107350

47. Han C, Xu R, Gao CX, et al. Socioeconomic disparity in the association between long-term exposure to PM2.5 and mortality in 2640 Chinese counties. *Environ Int*. 2021;146:106241. doi:10.1016/j.envint.2020.106241

48. Nyadanu SD, Tessema GA, Mullins B, Kumi-Boateng B, Ofosu AA, Pereira G. Ambient particulate matter air pollution and stillbirth in Ghana: A difference-in-differences approach. *Atmospheric Pollut Res*. 2022;13(7):101471. doi:10.1016/j.apr.2022.101471

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 1: Ecological research studies using Poisson regression with TWFE to study particulate matter pollution’s effect on health** | | | | | | | | |
| **Article** | **Journal** | **Exposure** | **Outcome** | **Spatial & temporal levels** | **Modeling approach** | **FEs** | **Covariates** | **Main findings** |
| Wang et al. 2016 | Environmental Health Perspectives | Area-weighted PM2.5 from 1km2 initial grid | AC mortality | Census tract – year level, 2004-2009 | **Poisson regression with overdispersion** | Census tract and year | linear spline (df=1) | 3% (.2, 5.9%) increase in mortality per 2µg/m3 annual increase |
| Renzi et al. 201941 | Environmental Health Perspectives | Area-weighted PM10 from 1km2 initial grid | AC mortality for ages 35+ | 378 municipalities in Latium (IT) at the year level, 2006-2012 | **Conditional Poisson regression (based on Wang et al. 2016)** | District and year | Mean summer temp, mean winter temp, std. dev of summer temp, std. dev of winter temp | .8% (.2, 1.3%) increase in mortality per 1 µg/m3 annual increase |
| Yu et al. 202242 | PLOS Medicine | Wildfire smoke PM2.5 from .25 degree^2 initial grid | Cancer mortality (1,332,526 total deaths) | Municipality-level in Brazil at the year level, 2010-2016 | Quasi-Poisson regression | Municipality and year | Temperature, GDP | RR = 1.02 (1.01, 1.03) for all-cancer mortality per 1 µg/m3 annual increase |
| Fan et al. 202343 | Environmental Research | Pop-weighted PM2.5 from .01 degree^2 initial grid | Cancer mortality (947,337 total deaths) | 53 districts in Jiangsu Province at the year level, 1998-2013 | **Conditional Poisson regression (based on Wang et al. 2016)** | District and year | Air temp, relative humidity | 2.7% (2.0, 3.4%) increase in cancer mortality per 1 µg/m3 annual increase |
| Yu et al. 202044 | PLOS Medicine | Population-weighted PM2.5 from 1km2 initial grid | AC mortality (217,510 total deaths) | Postcode region (449 total) at the year level, 1990-2013 | **Conditional Poisson regression (based on Wang et al. 2016)** | Postcode and year | Mean summer and winter temp, sd of summer and winter temp; economic development | 2.02% (1.41, 2.63%) per 1µg/m3 annual PM2.5 increase |
| Leogrande et al. 201945 | Environmental International | Population-weighted exposure to industrial PM10 | Mortality in 11 areas in Taranto (IT) | Cohort-level, n=262,375 individuals | **Conditional Poisson regression (based on Wang et al. 2016)** | Year, area, age group | -- | 1.86% (-0.06, 3.83%) increase per 1µg/m3 industrial PM10 |
| Yu et al. 202227 | Environmental International | Population-weighted PM2.5 from initial .05 degree^2 grid | Loss of life expectancy | Municipalities in Brazil (5,565 total) at the year level, 2010-2018 | Conditional Poisson regression (based on Yu et al. 2020) | Municipality, year | mean summer and winter temps, and their SDs; GDP per capita | RR=1.18 (1.15, 1.21) for all-mortality for each 10µg/m3 increase in annual PM2.5 |
| Han et al. 202136 | Environmental International | Population-weighted PM2.5 from 11km2 grid | AC mortality | 2,869 counties in China, data from 2000 and 2010 censuses | **Conditional Poisson regression (based on Wang et al. 2016)** | Municipality, year | mean summer and winter temps, and their SDs (population weighted) | 3.8% (3.0-5.0) increase in ACM per 10µg/m3 annual PM2.5 increase |
| Nyadanu et al. 202248 | Atmospheric Pollution Research | Zonal statistics aggregation from .01 degree2 initial grid | Stillbirths (81,611 stillbirths out of 5,229,338 total births) | District level (260 districts in Ghana) at the year level, 2012-2019 | **Conditional Poisson regression (based on Wang et al. 2016)** | District and year | Temperature (same aggregation as PM2.5)--season-specific mean and SD values | RR = 1.03 (.97, 1.09) per 10µg/m3 annual avg all-source PM2.5 |

A picture containing text

Description automatically generated

**Figure 1: Spatiotemporal variation in wildfire smoke PM2.5 and mortality.** **a)** Spatial variation in monthly wildfire smoke PM2.5 over 0.5 µm/m3 at the county level, January 2006 – February 2020. **b)** Temporal variation in mean monthly wildfire smoke PM2.5 over the same time period as **a.** Spikes tend to occur during the spring and summer months. **c)** Spatial variation in monthly age-adjusted mortality rate per 100,000 at the county level, January 2006-February 2020. **d)** Temporal variation in monthly age-adjusted mortality rate per 100,000. Spikes tend to occur during the fall and winter months. **e)** Distribution of monthly wildfire smoke PM2.5 values at the county level. A log scale is used for the y axis because the vast majority of county-months experience near-zero levels of smoke PM2.5.

Table

Description automatically generated

Graphical user interface, chart

Description automatically generatedt

**Figure 2: Varying regression specifications leads to large differences in point estimates and standard errors.** We ran 75 regressions with varied fixed effects and covariate specifications. Shown here for each regression are point estimates, standard errors using the IID assumption in light blue, robust standard errors clustered at the county level in black, and model parameters below. The red point estimate in **a** indicates our recommended model specification. Figures **a** and **b** are separated to clearly display the point estimates and standard errors, since the 5 results shown in **b** are an order of magnitude lower than the remaining 70 results.

Chart

Description automatically generated

**Figure 3: Varying inclusion years leads to large differences in point estimates and standard errors.** All analyses used our recommended regression over different sets of years. **a)** We test all fixed 4-year windows in our sample. **b)** We iteratively add years from a base 4-year window. **c,d)** We run the same regressions as **a** and **b** while stratifying by age.

Table

Description automatically generated