# Characterizing Exposure and Health Impacts of Ultrafine Particles and Wildfire Smoke

Annie Doubleday

A dissertation
submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee:
Lianne Sheppard, Chair
Tania Busch Isaksen, Co-chair
Sara Adar

Program Authorized to Offer Degree:
Department of Environmental and Occupational Health Sciences

University of Washington

## Abstract

Characterizing Exposure and Health Impacts of Ultrafine Particles and Wildfire Smoke

Annie Doubleday

Chair of the Supervisory Committee:

Lianne Sheppard
Department of Environmental and Occupational Health Sciences
Department of Biostatistics

Tania Busch Isaksen
Department of Environmental and Occupational Health Sciences

Fine particulate matter ($PM_{2.5}$) air pollution is associated with excess morbidity and mortality and estimated to be the most important environmental risk factor for mortality globally. Ultrafine particles (UFPs) and $PM_{2.5}$ from wildfire smoke (WFS), are both thought to be more toxic than ambient $PM_{2.5}$. The goal of this dissertation is to deepen our understanding of these understudied pollutants and to estimate health effects of their exposures in distinct applications.

In the first aim, I developed an approach to use ultrafine particle mobile measurements from traffic-related air pollution mobile monitoring campaigns for epidemiology. I developed an exposure surface that is more representative of residential locations by lessening the contribution of localized on-road UFP plumes. Compared to an exposure surface that uses unadjusted mobile data, the exposure surface with predictions derived from on-road plume adjusted data performs better, with an MSE-based $R^2$ of 0.71 compared to 0.60. Further, I found that UFP predictions at residential locations that use mobile measurements versus stationary measurements are more spatially resolved and pick up more information at known hotspot

locations. We recommend correcting mobile measurements for use in epidemiology. Future research is needed to understand the impact of using predictions from on-road plume adjusted measurements in epidemiologic inference.

In the second aim, I examined the impact of UFP, among several other traffic-related air pollutants, on physical functioning in the Adult Changes in Thought study, a cohort of older adults in the Puget Sound region of Washington state. Physical functioning is an important measure of aging and is measured by three assessments that make up the modified Performance-Based Physical Function Assessment, giving a score from 0-12: gait speed, chair-stand time, and grip strength. We used exposure predictions derived from the mobile monitoring campaign discussed in Aim 1, and examined UFP (several size bins), nitrogen dioxide ($NO_2$), black carbon (BC), and $PM_{2.5}$. We used linear mixed effect models to evaluate the association at baseline and longitudinally. There was no evidence of a longitudinal effect of UFPs on physical performance. Exposure to higher levels of BC, UFP in the range of 56-178nm, and non-time varying $PM_{2.5}$ resulted in faster decline in physical performance score compared to exposure at lower levels, equivalent to an additional 3-5 months of aging over 5 years.

In the third aim, I examined the association between wildfire smoke and emergency department (ED) visits in Washington state during June-September, 2017-2020. I examined the impact of wildfire smoke presence versus absence on ED visits, and the impact of each 10 $\mu g/m^3$ increase in $PM_{2.5}$ from wildfire smoke. I used a time-stratified case crossover design with conditional logistic regression. I observed an increased odds of asthma ED visits immediately after and in five days following initial wildfire smoke exposure (lag 0 OR: 1.13; 95% CI: 1.10, 1.17; lag 1-5 ORs all 1.05 or greater with a lower CI of 1.02 or higher), and an increased odds of respiratory ED visits across all five days following exposure. I observed mixed results for cardiovascular ED visits, with evidence of increased odds several days following initial exposure. I also found increased odds across all visit categories for each 10 $\mu g/m^3$ increase in

wildfire smoke impacted $PM_{2.5}$. Finally, I observed increased odds of respiratory visits among ages 19-64, and among ages 5-64 for respiratory visits. The results of this study provide evidence for an increase in respiratory ED visits immediately following exposure and in the several days following, and a lagged effect on cardiovascular ED visits. These increased risks are observed particularly among children and younger to middle-aged adults.

This dissertation advances our understanding of exposure characterization and health impacts of two toxic components of fine particulate matter, ultrafine particles and wildfire smoke, and provides support for further investigation.

**Acknowledgments**

I want to thank my primary advisor, Dr. Lianne Sheppard, for her incredible support and mentorship throughout my doctoral studies. I also want to thank my co-chair, Dr. Tania Busch Isaksen, for her support and mentorship throughout my academic career, and investment in my long-term growth and development. And, to my other committee members, Dr. Elena Austin, Dr. Sara Adar, and Dr. Anjum Hajat for their support and feedback throughout my graduate training.

To the Sheppard lab: Dr. Magali Blanco, Amanda Gassett, Dr. Jianzhao Bi, Nancy Carmona; and the Busch Isaksen lab: Claire Schollaert, Daaniya Iyaz, for their support and assistance.

To the DEOHS staff members who provided support and resources: Brian High, Trina Sterry, Veronica James.

To my amazing family and friends for their continued support during my graduate training. To my partner, Megumi, for her incredible support and confidence in me throughout my PhD, and to my family for their constant support.

To my funding sources during my doctoral training: BEBTEH: Biostatistics, Epidemiologic, & Bioinformatic Training in Environmental Health (NIEHS, T32ES015459), and Department of Environmental and Occupational Health Sciences research assistant and teaching assistant funding.

# Chapter 1: Introduction

Fine particulate matter ($PM_{2.5}$) air pollution is associated with excess morbidity and mortality (1–3) and estimated to be the most important environmental risk factor for mortality globally (4). Recent evidence suggests that not all fine particulate matter is equally toxic. Ultrafine particles (UFPs), defined as particulate matter less than 0.1 microns in diameter, and $PM_{2.5}$ from wildfire smoke (WFS), may be more toxic than total ambient $PM_{2.5}$ (5,6). However, exposure variability and health impacts of these components of total $PM_{2.5}$ are not fully understood. The goal of this dissertation is to deepen our understanding of these understudied particulate matter pollutants and to estimate health effects of their exposures in distinct applications.

Aim 1: UFPs are composed of fine particulate matter less than 0.1 microns in diameter. Sources of UFPs include combustion, vehicle traffic, industry, and air transportation (7). Thus, exposure to UFPs is of concern predominantly in urban areas with peak concentrations near roadways and in neighborhoods under aircraft landing paths (8,9). However, UFPs are not monitored by regulatory monitors, they decrease rapidly at distances from their primary emission source (i.e., over a scale of 100-300m) (10,11), and they vary significantly over space and time. Thus, it is challenging to understand spatial differences in UFP exposure that may be associated with chronic health outcomes.

Mobile monitoring campaigns, where air pollution measurements are collected from a moving platform, are increasingly employed to understand localized spatial differences in exposure to traffic-related air pollutants (TRAP), including UFPs (12–17), and to predict exposure for epidemiologic cohorts (18,19). Many mobile monitoring studies use on-road data, or measurements collected while the mobile platform is in motion (non-stationary), to estimate annual averages and create exposure prediction surfaces for many applications (12–17,20,21).

Only a few studies have investigated the impact of predictions from mobile measurements on epidemiologic inference or considered whether mobile measurements are representative of residential exposure (20–23). We designed and conducted a mobile monitoring campaign in the greater Seattle area that included measurements at 309 short-term stationary locations as well as on road (24). These 309 locations were chosen based on their representation of Adult Changes in Thought (ACT) cohort locations, and act as our gold standard comparison for this study. Thus, our campaign is uniquely designed to investigate differences in UFP exposure predictions from mobile measurements versus short-term stationary measurements. We add importantly to the literature by evaluating how to utilize mobile data into exposure prediction modeling for epidemiology to better characterize fine-scale exposure, and ultimately to better characterize health impacts.

Aim 2: Toxicology studies suggest that UFPs may be more toxic than larger particles (7,25). UFPs may enter the circulatory system through the lungs (26) and can cross the blood-brain barrier via the olfactory bulb (7). Through these mechanisms, UFPs are hypothesized to contribute to systemic inflammation and oxidative stress. These impacts may contribute to the development of chronic diseases (6). Recent epidemiologic studies find evidence for an association between UFP exposure and cardiovascular and inflammatory impacts, including changes to blood pressure, heart rate variability, and systemic inflammation (6). As individuals age, chronic diseases, including cardiovascular and respiratory diseases, often take a toll on individual physical functioning (27). However, few studies have examined the impact of air pollution on physical disability or physical functioning, (28–31) and none have examined the impact of UFP exposure.

Physical function is an important measure of aging (32,33). Diminished physical function is common among older adults, and of growing concern due to the increasing population of older adults worldwide (34). Poor or worsening physical function may be detectable much earlier

than the onset of other adverse outcomes, and thus may aid in earlier identification of individuals at high risk of hospitalization or mortality, for example (35).

We examine the association between long-term exposure to UFPs and changes in physical function, leveraging a well-characterized cohort of older adults in the Puget Sound region (Aim 2).

Aim 3: Studies examining the impact of wildfire smoke exposure on morbidity are increasingly abundant (36–42) often examining respiratory and cardiovascular hospitalizations in many regions including California (36), Washington (43), and Australia (44). These studies consistently find significant associations between short-term wildfire smoke exposure and respiratory hospitalizations, however, the results for cardiovascular hospitalizations are mixed. Additionally, a few studies examine emergency department (ED) visit outcomes (41,45,46). Findings from these studies are similar to those from hospitalization studies, with significant associations reported for respiratory visits, and mixed results for cardiovascular visits.

Syndromic surveillance is a real-time monitoring system that collects Emergency Department and Urgent Care visit data to help rapidly identify outbreaks and other public health emergencies (49). Utilization of syndromic surveillance data increases the power to detect an association due to the high frequency of ED visits, as compared to more severe outcomes such as hospitalization or mortality. However, there are few studies explicitly examining syndromic surveillance data in a wildfire smoke context (47,48). Further, to date, there are only two statewide wildfire smoke epidemiologic studies in Washington, and only one examining non-mortality outcomes (43,50). By examining the impact of wildfire smoke exposure during the summers of 2017-2020 on syndromic surveillance outcomes in Washington, we gain a better understanding of the burden on the healthcare system due to recent smoke events. Our results may help inform to state and local agencies in crafting more targeted risk communication and interventions.

# Chapter 2: Characterizing ultrafine particle mobile monitoring data for epidemiology[1]

[1]Reprinted with permission from: Annie Doubleday, Magali Blanco, Elena Austin, Julian Marshall, Timothy Larson, Lianne Sheppard. Characterizing ultrafine particle mobile monitoring data for epidemiology. *Environmental Science & Technology* 2023. https://doi.org/10.1021/acs.est.3c00800. Copyright 2023 American Chemical Society.

**Abstract**

Mobile monitoring is increasingly used to assess exposure to traffic-related air pollutants (TRAPs), including ultrafine particles (UFPs). Due to the rapid decrease in the concentration of UFPs and other TRAPs with distance from roadways, mobile measurements may be non-representative of residential exposures, which are commonly used for epidemiologic studies. Our goal was to develop, apply, and test one possible approach for using mobile measurements in exposure assessment for epidemiology. We used an absolute principal component score (APCS) model to adjust the contribution of on-road sources in mobile measurements to provide exposure predictions representative of cohort locations. We then compared UFP predictions at residential locations from mobile on-road plume-adjusted versus stationary measurements to understand the contribution of mobile measurements and characterize their differences. We found predictions from mobile measurements are more representative of cohort locations after down-weighting the contribution of localized on-road plumes. Further, predictions at cohort locations derived from mobile measurements had more spatial variation compared to those from short-term stationary data. Sensitivity analyses suggest additional spatial information from the mobile measurements capture features in the exposure surface not identified from the stationary data alone. We recommend correction of mobile measurements to create exposure predictions representative of residential exposure for epidemiology.

## Introduction

Mobile monitoring of air pollution dates back to the 1950s and 1960s (51). Over the past decade, mobile monitoring has increasingly been used to characterize fine-scale spatial variation in exposure to traffic-related air pollution (TRAP) in urban environments (12–17). One such pollutant, ultrafine particles (UFPs; particles ≤ 100 nm in diameter), varies over space and time with a rapid decrease in concentrations at increasing distances from their primary emission sources over a scale of 100-300 m (10,11). UFPs come from many sources in urban areas, including vehicular traffic, aircraft emissions, industrial sources, and biomass burning, among others (7). Because of their small size and low mass per particle, UFPs are typically measured as counts (particle number concentration, PNC) rather than as mass.

Many mobile monitoring platforms can collect data while in motion ("mobile" or "non-stationary") or while stopped ("stationary" or "fixed site") at a near-road location; purposes include to estimate vehicle emissions, to measure on-road concentrations, and for air pollution epidemiology. For mobile monitoring, exhaust plumes from nearby vehicles have an important influence on measurements, particularly when the monitoring platform is in motion (52,53). Many studies have observed a decrease in UFP concentration with distance from roadways depending on wind speed, wind direction, and atmospheric stability, often reporting an exponential decrease in PNC with distance from roadway (8,10). Thus, mobile monitoring platforms likely capture elevated UFP concentrations relative to near-roadway locations.

Many mobile monitoring studies use data collected on roads to estimate annual averages and create exposure prediction surfaces (12–17,20,21). Multiple drive passes of the same locations are essential to capture good long-term averages. Blanco et al. (54) showed that a minimum of 25 drive passes is needed to estimate a location's annual average with up to 25% error of the true annual average. In addition, to achieve unbiased annual average estimates, the design of drive passes is important. Mobile campaigns should strive for designs that are temporally balanced by time of day, day of week, and season of year (54,55). However, the

expense and logistical challenges of mobile monitoring studies increase with the number of drive passes and the attention to a temporally balanced sampling design.

A few studies compared UFP estimates from mobile measurements to estimates from stationary site measurements to understand the integrity and usability of these short duration mobile data (20,21,23) (see also Table S9 rows 1, 2, and 4). One study compared median concentrations and predictions from mobile measurements across the campaign to 100-minute stationary measurements at sidewalk locations for UFP and black carbon (BC) (21); a second compared the distribution of averaged 30-minute stationary and averaged 12s road segment concentrations for UFP and BC, and compared predictions from these two data sources at residential locations (20); and a third study compared UFP predictions from mobile measurements to both measurements at short-term stationary sites and to predictions at residential locations from these stationary measurements (23). All three studies found that predictions from mobile measurements overestimated predictions from stationary data, with predictions from mobile UFP data 30-60% larger than those from stationary data (20,21,23). However, none leveraged stationary data or background sites to identify and adjust localized on-road plumes in the mobile data to yield exposure predictions that are more representative of residential exposure for epidemiology. Our campaign measured TRAPs at 309 stationary near-roadway locations that are representative of residential exposure (based on the distribution of geographic covariates) in addition to mobile on-road measurements. This provides an opportunity to advance understanding of how to best use mobile and stationary data in exposure assessment for epidemiology.

Our goal was to develop an approach to optimize the use of mobile measurements for epidemiology by developing an exposure surface that is representative of residential locations. To achieve this goal, we used an intentionally designed multi-pollutant mobile monitoring campaign to characterize TRAP exposures for the Seattle-based Adult Changes in Thought (ACT) cohort. Focusing on UFPs from our campaign, we first compared mobile to stationary

measurements. We then used stationary data to guide the development of an approach to lessen the contribution of localized on-road plumes from mobile data. Lastly, we characterized our model based on on-road plume-adjusted mobile measurements.

## Materials and Methods

### Study design

We conducted a multi-pollutant mobile monitoring campaign over approximately 1,200 $km^2$ in the greater Seattle area, with 309 stationary sampling locations representative of ACT cohort locations, and 34,355 kilometers of mobile sampling between stationary locations on the same drive days (56). The measurements were collected on 288 drive days in a temporally balanced approach that included all four seasons, weekend and weekdays, and most hours of the day, between March 2019 and March 2020 (56). Mobile measurements were collected along nine fixed driving routes for multiple air pollutants, including particle number concentration (PNC) to capture UFPs, black carbon (BC), fine particulate matter ($PM_{2.5}$), nitrogen dioxide ($NO_2$), and carbon dioxide ($CO_2$). PNC was measured by four instruments: a NanoScan, DiSCmini, and two P-Traks (screened (36-1,000 nm) and unscreened (20-1,000 nm)). In this analysis, we use the unscreened P-Trak (20-1,000 nm) to characterize PNC, as it included the largest range of particle diameters at a 1-second time resolution. All nine fixed routes in the campaign were driven about 28 times (visits). The quality control procedures prioritized make-up visits at stationary locations in the case of equipment failure to ensure complete multi-pollutant data at stationary locations, but typically did not make up full drive routes (56). All 1-second measurements were lagged by 10 seconds to account for the residence time of the manifold.

### Overview of approach

Our approach consisted of several steps, outlined in Figure 1, and described below. We first aggregated and summarized the mobile measurements, then compared them to the stationary measurements. Next, we partitioned the multipollutant data using PCA to identify components dominated by on-road UFP sources. Then, we went through a series of steps to decrease the contribution of on-road sources in the mobile UFP measurements, and finally evaluated our model.

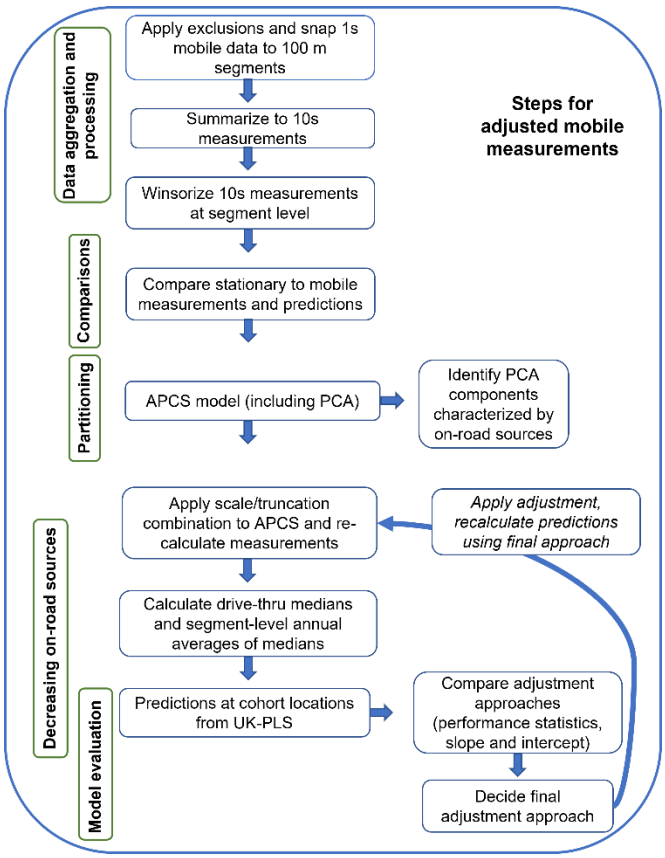


**Figure 1.** Data processing steps on-road plume-adjusted mobile UFP data.

Data aggregation and processing

We divided the driving routes into 5,887 100m road segments after exclusions (Note S1; Table S1). We chose 100m due to its relatively short length and to obtain data with adequate sampling duration for most segments (median of 10 seconds). Additionally, prior research showed little difference in model performance based on varying segment length between 50 and 200 m (15). We assigned each 1-second measurement to the nearest 100m road segment and used the midpoint of each segment as its spatial location (12,17,57).

We then averaged all 1-second measurements into 10-second averages to align with the BC data time stamps, where the 10 seconds prior to each BC time stamp was averaged. We winsorized all measurements at the segment level to the 2.5% and 97.5% quantiles to reduce the influence of outliers on our multi-pollutant PCA analysis while maintaining the spatial distribution of the data (Figure S1). As in Blanco et al., 2022, we employed a “mean of medians”

approach: for each visit through a road segment (i.e., drive pass), we calculated the median 10s concentration. Then for each road segment, we calculated the mean of all visits (56).

We first obtained geographic covariates ("geo-covariates", n=350) at cohort locations and at each road segment midpoint using previously documented methods (56). Geo-covariates included, among others, proximity variables (e.g., distance to major road, distance to airport), land-use variables, and population density variables. As in Blanco et al. 2022 and Keller et al. 2015, we excluded variables that lacked sufficient variability or coverage, had insufficient overlap in values between measurement locations and cohort locations, or contained too many outliers (Note S2) (56,58); the final geo-covariate dataset contains 191 variables (Table S2).

We developed PNC prediction models using universal kriging with partial least squares regression (UK-PLS), as we have previously reported (56). Briefly, we first fit a partial least squares (PLS) regression to summarize the geo-covariates at mobile locations. Then, we conditioned on 3 PLS components to fit a linear regression assuming spatially correlated errors (universal kriging with an exponential variogram). We obtained predicted PNC concentrations at cohort locations by using PLS score covariates and PNC measurements from nearby segments through the spatially correlated errors. Models were developed to estimate exposures using either the PNC mobile location averages described above, or the adjusted PNC mobile location averages described below.

Comparisons between stationary and mobile locations

In all analyses, the measurements, annual averages, and predictions from the 309 stationary locations were considered the reference data for our study, as these locations were designed to optimize exposure assessment for our cohort (56). To compare annual averages from stationary and mobile locations, we developed two comparison sets: one based on geographic proximity of mobile and stationary locations, and the second based on geo-covariate

similarity. We defined geographic proximity as within 100m of a stationary location, and geo-covariate proximity as mobile locations with similar geo-covariates as stationary locations. We determined this by conducting a PCA of all mobile and stationary location geo-covariates and finding the geo-covariate distance between each stationary and mobile location, defined as the Euclidean distance between the PC 1 and PC 2 scores (59). We found the mobile location closest to each stationary location based on this PCA distance. We compared annual averages from these 309 mobile locations to those from the 309 stationary locations (i.e., we selected on matched mobile location for each stationary location). We then compared differences in PNC predictions generated using stationary measurements (reference) to either those generated using geographically proximal mobile data or geo-covariate proximal mobile data at the 11,904 ACT cohort locations.

To understand if the mobile data were generally higher than the stationary data, we conducted sign tests, testing whether the probability of points on a scatterplot comparing mobile and stationary annual averages were above the 1:1 line (Note S4).

Partition of multi-pollutant data

Based on the above comparisons, and prior research showing that UFPs are present in tailpipe emissions (60) as well as rapid decreases in UFP concentration with distance from roadways (8,10), we hypothesized that concentrations may be systematically different on-road versus near-road. To correct for that difference (i.e., concentrations on-road versus near-road), we aimed to reduce the contribution of localized on-road plumes to the PNC mobile monitoring data, thereby creating an adjusted PNC measure that is less impacted by localized plumes. If that step is successful, we posited that the adjusted PNC measure would be more comparable to a near-road exposure.

To accomplish this goal, we employed the absolute principal component score (APCS) model, first described by Thurston and Spengler (61). We prepared all pollutant measurements as described in the 'Data aggregation and processing' section, then centered all variables at their mean, and scaled them by their standard deviation. We then partitioned the multi-pollutant data (*N*=305,243 10-second measurements) into component sources using principal components analysis (PCA). We applied a Varimax rotation to these components to improve interpretability, following previously developed methods (53,62,63). We included four pollutants and seven measurements at the mobile locations (56) in our PCA (Table S3). We included three PNC instruments in addition to the difference between the screened and unscreened P-Trak as they each measure different PNC size bins, thus capturing different and important sources of UFP. We kept *p*=3 components based on the scree plot showing the cumulative proportion of the variance explained. We then derived the APCS for the components:

$$APCS_{j,k} = S_{j,k} - (S_0)_j (j = 1, \dots p) \quad (1)$$

where $S_{j,k}$ are the scores for the $k^{th}$ observation of the $j^{th}$ component derived from the scaled and centered pollutants, and $(S_0)_j$ is the predicted value of the zero vector from the rotated PCA, corresponding to the scores for zero concentration values (62). $(S_0)_j$ is subtracted from $S_{j,k}$ to yield positive scores (62,63). The scores are on the scale of the standardized data and represent the orthogonal decomposition of the standardized data. Next, we regressed the APCS against $PNC_k$, the total PNC from the unscreened P-Trak:

$$PNC_k = b_0 + \sum_{j=1}^{p} b_j (APCS_{j,k}) + \varepsilon_k \quad (2)$$

The intercept in Equation 2 is the contribution to $PNC_k$ of sources not accounted for by the *p*=3 components in the PCA. Then, $b_j(APCS_{j,k})$ represents the predicted concentration of PNC contributed by component *j* to the *k*th sample (62). We utilized this model to decrease the contribution of on-road sources to the PNC, described below. We used the *psych* (v 2.2.9) (64) and *GPA rotation* (v. 2022.10-2) (65) packages in R (66) for the PCA analyses.

Decreasing the contribution of on-road sources

To meet our goal of creating adjusted PNC data that are more representative of near-road exposures, we optimized the slope (closest to 1) and then the intercept (closest to 0) of the regression line of the predictions at selected cohort locations based on the adjusted mobile PNC data, compared to predictions based on the stationary PNC data on the x-axis (see Note S3 for additional details). Based on prior research, we hypothesized that on-road sources would have higher contributions of $NO_2$, $CO_2$, and BC (63), particularly compared to PNC, and that the contributions from these pollutants to the PNC measurements should be reduced to better represent near-road exposures. Locations with higher contributions of PNC relative to other TRAPs may be more typical of non-road sources such as aircraft emissions, biomass burning, and industrial emissions (7,9,63). Based on the multi-pollutant PCA described above, we defined the PCs with large contributions of TRAPs as on-road source PCs (PCs with loadings of $NO_2$, $CO_2$, or BC greater than 0.7) from the top three PCs.

To ensure that the mobile measurements were not overly impacted by localized plumes, we scaled the on-road source PCs (PC 2 and 3) by a multiplier close to 1 to dampen their contribution overall and truncated individual values of these same components at an upper percentile threshold to mute the contribution of larger localized plumes at locations with the highest values (Note S3). We adjusted the PNC measurements accordingly:

$$AdjPNC_k = b_0 + b_1 APCS_{1,k} + \sum_{j=2}^{p=3} scaleb_j(truncateAPCS_{j,k}) + \varepsilon_k \quad (3)$$

where scale = 0.99 and truncate = 0.75 in our selected model. Using the on-road plume-adjusted mobile PNC data ($AdjPNC_k$) as input into a new UK-PLS model, we obtained PNC predictions at all cohort locations. Finally, for comparison, we considered an additional

adjustment approach, adjusting mobile measurements based on the best-fit linear regression line between the unadjusted mobile and stationary predictions.

Model evaluation

To characterize prediction performance, we used the stationary data as a reference dataset as they were selected to be representative of residential exposures. Prediction at stationary locations allowed us to externally validate performance of models based on the mobile data. For comparison, we also reported 10-fold cross-validated (CV) results for models developed from the stationary data; by fitting the model to a "training" dataset of 90% of the data and testing it on the left-out 10% "validation" dataset, repeated across all 10 folds, cross-validation is an out-of-sample validation that approximates an external evaluation. For the external validations, we considered predictions based on adjusted and unadjusted mobile PNC data at stationary locations. We report both regression-based $R^2$ and the MSE-based $R^2$ because multiple studies have documented that (as is the case with our data) mobile data tend to be higher on average than nearby stationary data (20,21,23).

Next, we evaluated maps of the predictions, showing those from the adjusted mobile and stationary data, as well as the difference between the adjusted mobile and stationary predictions. Focusing on the largest prediction differences, we additionally identified cohort locations which were outside the bulk of the data. We mapped these outlier locations to evaluate the plausibility of their predicted values, and we characterized the distribution of several geo-covariates at these locations compared to all other cohort locations. We hypothesized that if these outlier locations were closer to known PNC sources (e.g., airports, major roads, truck routes) and/or were farther from the nearest stationary location compared to the other cohort locations, the prediction differences at these locations were more likely to be large. This could be due to the large number of mobile locations providing more information in the prediction model, supporting the idea that they may be predicted better by the mobile data.

Finally, we evaluated how different measurements affected the variability of predictions at the cohort locations. Because annual averages from mobile data are inherently noisier than annual averages from stationary data, owing to the greater uncertainty of the mobile data spatial locations as well as their shorter sampling duration, we considered additional subsamples of our data and estimated the variance of the predictions from models that used: 1) all mobile data; 2) all stationary data; 3) time-sampled stationary data restricted to 10 consecutive 1-second measurements so they had a similar sampling duration to the mobile data; and 4) location-sampled mobile data based on the same 309 mobile locations selected to be the most comparable to the stationary locations, described previously. We included options 3) and 4) to foster understanding of how sampling duration and number of locations impact variability. Comparing the variance of predictions from these models allows insight into how the mobile model may be picking up additional spatial information.

All analyses were conducted in R programming software version 4.1.3 (66).

## Results and Discussion

### Data description

Table 1 describes the number of repeat visits, median number of 1-second measurements per visit, and total measurement duration for stationary, mobile, arrival, and departure data. For mobile measurements, the median number of repeat visits to each segment was 28 (IQR: 27, 29), the median number of 1-second measurements per visit was 10 (IQR: 8, 17), and the median PNC concentration was 7,130 (IQR: 4,400; 11,400) pt/cm$^3$. Table S4 gives the five-number summary of the distribution of the 1-second and 10-second PNC measurements, showing the data have a similar median and IQR at the 10-second level.

**Table 1**. Summary of 1-second PNC measurements for mobile and stationary locations

| Category | Locations (N) | N repeat visits | N 1-second measurements | Cumulative measurement time | PNC median (pt/cm$^3$) (IQR) |
|---|---|---|---|---|---|

| | | (Median (IQR)) | per visit (Median (IQR)) | (min) per location (Median (IQR)) | |
|---|---|---|---|---|---|
| Stationary | 309 | 28 (27, 29) | 120 | 62.4 (60, 66) | 5,860 (3,640; 9,260) |
| Mobile | 5,887 | 28 (27, 28) | 10 (8, 17) | 5.3 (3.9, 8.7) | 7,130 (4,400; 11,400) |
| Arrival[1] | 309 | 28 (27, 29) | 23 (18, 32) | 12.1 (10.6, 14.4) | 6,320 (3,930; 10,100) |
| Departure[2] | 309 | 28 (27, 29) | 29 (22, 42) | 14.5 (11.7, 18.6) | 6,110 (3,790; 9,770) |
| Geo-covariate matched[3] | 309 | 28 (27, 29) | 19 (11, 35) | 5.6 (3.9, 9.7) | 7,080 (4,290; 12,000) |
| Mobile, no arrival[1] or departure[2] | 5,269 | 28 (27, 28) | 10 (8, 14) | 4.9 (3.7, 7.5) | 7,310 (4,530; 11,700) |

[1]Measurements in the 100m on-road before arriving at each stationary location
[2]Measurements in the 100m on-road after departing each stationary location
[3]Measurements at 309 mobile locations matched to the 309 stationary locations by geo-covariate proximity.

The 5,887 mobile locations and 309 stationary locations have fairly similar distributions of several key geographic covariates, including distance to A1 roads (interstates and major highways) and distance to Seattle-Tacoma International Airport (Table S5). However, the stationary locations were primarily near A4 roads (local, neighborhood, and rural roads), and the median population density was higher, whereas the mobile locations were split more evenly between A3 (secondary and connecting roads) and A4 roads, with ~5% of segments on A2 roads (primary roads without limited access) (67).

Comparisons between stationary and mobile locations

Generally, annual average PNC estimates obtained from mobile data were higher than annual averages generated from stationary data (Table 1; Note S4). This was true comparing annual averages from mobile locations near stationary locations (“arriving”, “departing”), as well as when comparing annual averages from mobile locations that are similar to stationary locations in geographic covariate space. This pattern was confirmed by sign tests showing that mobile annual averages were significantly higher than stationary averages (Note S4).

A similar though less extreme pattern emerged comparing predictions at cohort locations using annual averages derived from all mobile locations versus stationary data predictions (Figure 2, top). We found an intercept greater than zero (315 pt/cm$^3$; 95% CI: 233, 397) and a slope greater than one (1.12; 95% CI: 1.11, 1.13) comparing predictions plotted at ACT cohort locations that were derived from the mobile versus the stationary measurements. The regression-based $R^2$ is 0.77.

Our findings that the mobile data and predictions were on average, higher than the stationary data and predictions, further motivated our choice to adjust the mobile measurements to reduce the contribution of localized on-road plumes as our goal was to obtain mobile measurements and predictions that are more representative of exposure at cohort locations.

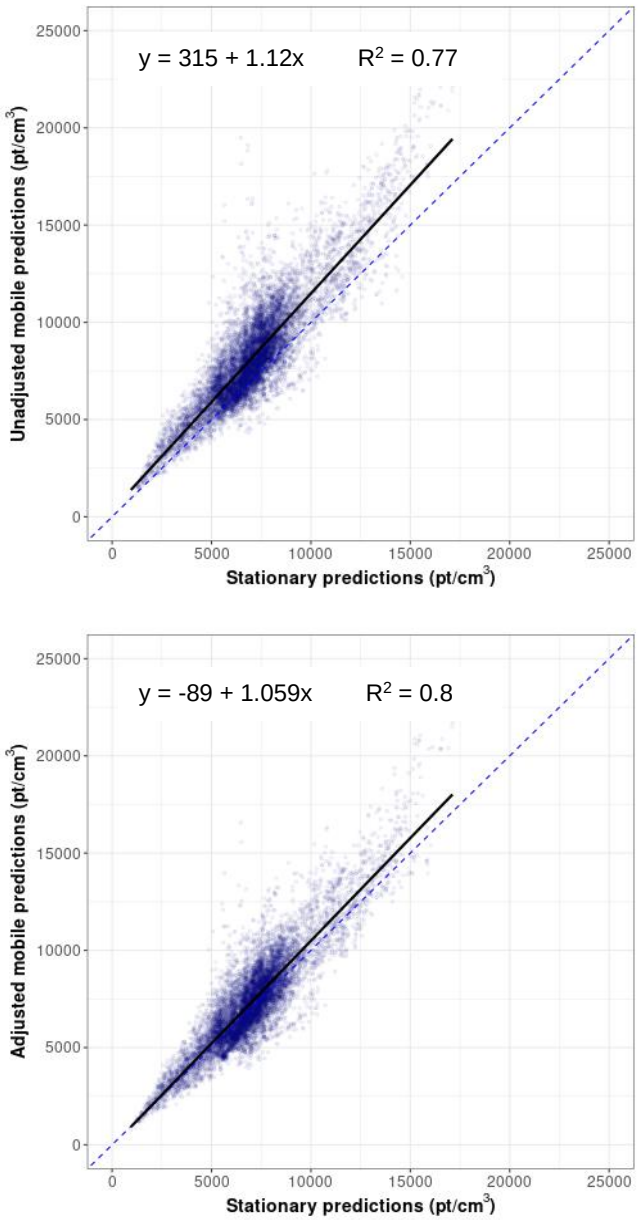


**Figure 2.** Unadjusted (top) and adjusted (bottom) PNC mobile predictions (y-axis) from UK-PLS at cohort locations compared to those from stationary data (x-axis). Adjusted predictions are scaled 0.99 and truncated 0.75 (see description in the text).

<u>On-road plume adjustment and re-estimation of predictions using mobile data</u>

In the multi-pollutant PCA (Tables S6 and S7; Figure S2), the first principal component, PC 1, is dominated by high UFP concentrations given its large loadings for all PNC measures

(Table S6), and its highest scores are concentrated around the Seattle-Tacoma International Airport and industrial areas south of downtown (Figure S2). The second component, PC 2, is characterized by large contributions from vehicle emissions ($CO_2$ and $NO_2$) and is clustered in the urban core and along major freeways and highways (9,68). The third component, PC 3, is dominated by a large BC contribution, which likely represents emissions from diesel engines (trucks) (68). Thus, we muted the contributions of the second and third components (PC 2 and PC 3), as these had high loadings of pollutants (TRAPs) that are typical of on-road sources. We evaluated a combination of multipliers (scale = 0.95, 0.99) applied to the PC 2 and PC 3 coefficients in the APCS model, and four truncating percentile thresholds (truncate = 0.75, 0.80, 0.85, 0.90) applied to the PC 2 and PC 3 scores (Eq. 3). We chose the final adjustment based on the best fit between the PNC predictions from on-road plume-adjusted mobile data and the stationary PNC predictions at selected locations: scaling the on-road source PCs (PC 2 and 3) by 0.99, and truncating values of PC 2 and 3 at the 75$^{th}$ percentile (Note S3).

We summarized the distribution of the PNC predictions from on-road plume-adjusted mobile data, following the plume adjustment method described above, versus the unadjusted predictions (Table S8 and Figures S3-S5). Like the results comparing annual averages, predictions derived from unadjusted mobile measurements are higher, more variable, and have a higher maximum than those from stationary measurements (Figure 2, top and Table S8). After adjusting the mobile measurements to lessen the contribution of on-road sources, the PNC predictions from the on-road plume-adjusted mobile data show better agreement with the predictions derived from the stationary measurements (Figure 2, bottom). The intercept is not statistically-significantly different from zero (-89 pt/cm3 (95% CI: -160, 19)), and the slope is larger than 1 (1.059 (95% CI: 1.049, 1.068)), but closer to 1 than the unadjusted predictions (1.12 (95%CI: 1.11, 1.13)).

Model evaluation

We observed the same regression-based $R^2$ for the unadjusted (Figure 2, top) and adjusted mobile predictions, and a lower MSE-based $R^2$ and larger RMSE for the unadjusted compared to the adjusted model (Table 2). Both have comparable, though slightly lower, regression-based $R^2$ compared to the cross-validated stationary predictions. We also reported prediction performance for the other plume algorithms considered and found similar prediction performance to the adjusted model below (Note S5; Table N5S1). One of the models we considered adjusts measurements based on the best-fit line between the unadjusted mobile and stationary predictions. While prediction performance for this model was comparable with the other adjusted models, we show this approach may over-adjust locations less contaminated by on-road sources and may give inadequate attention to what should be adjusted (Note S5, Table N5S1).

**Table 2.** Model performance statistics evaluated at stationary locations

| Model | Regression-based $R^2$ | MSE-based $R^2$ | RMSE (pt/cm$^3$) | Location-visits (N)[†] | Hours sampled |
|---|---|---|---|---|---|
| CV-Stationary (comparison)[‡] | 0.77 | 0.77 | 1,180 | 8,652 | 288 |
| PNC predictions from on-road plume-adjusted data | 0.73 | 0.71 | 1,510 | 165,733 | 908 |
| PNC predictions from unadjusted data | 0.73 | 0.60 | 1,800 | 165,733 | 908 |

[†]Location-visits are the sum of number of visits across all locations included in the model.
[‡]Cross-validated stationary
Note: Model performance statistics are all evaluated at 309 stationary locations. These are 10-fold cross-validated predictions for stationary data and external predictions for mobile data. Mobile data are either unadjusted or on-road plume-adjusted (99% scaled; 75% truncated, discussed in Note S3). We note that the stationary data are inherently less noisy than the mobile data because the longer sampling (median duration per visit: 2 minutes (stationary) versus 10 seconds (mobile)). Thus, we expect the RMSE from the cross-validated stationary model to be lower than the external validation with the mobile data.

Characterizing the PNC predictions from on-road plume-adjusted data

We further explored what we learned from using on-road plume-adjusted mobile measurements to predict PNC compared to predictions based on stationary PNC measurements. PNC predictions from on-road plume-adjusted mobile measurements (Figure 3, left panel) show greater variation, with particularly high levels near the airport, downtown, and

south of downtown. The difference in predictions from the on-road plume-adjusted mobile data minus stationary data (Figure 3, right panel) show the mobile measurements pick up trends in PNC exposure that are different from those captured by the stationary measurements (Figure 3, center), in part because we have information at many more locations when using the mobile measurements. Areas that are bright green or yellow (Figure 3, right) reflect larger PNC exposure predictions from on-road plume-adjusted mobile measurements compared to stationary measurements. This indicates that the mobile measurements are identifying spatial structure not captured by the stationary data. For example, the map shows that the adjusted mobile data identify additional areas with elevated PNCs, including one area that may capture elevated PNCs to the north of a regional airport in Auburn (Figure S6), in addition to likely industrial sources south of downtown.

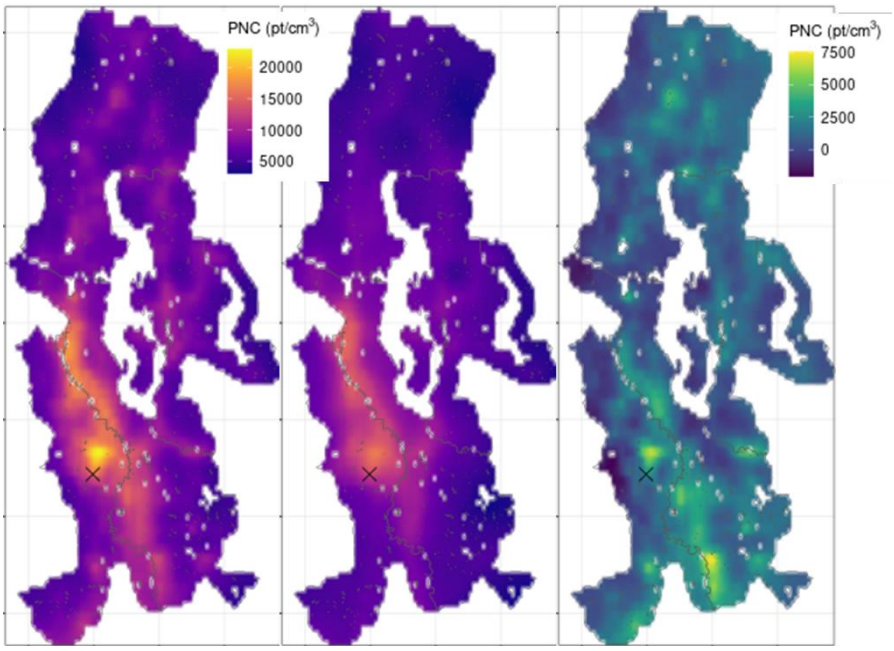


**Figure 3.** PNC predictions from on-road plume-adjusted mobile data (left) and from stationary data (center); and the difference between those two (right). The black 'x' represents Seattle-Tacoma International Airport.

To further understand the PNC predictions from the on-road plume-adjusted data, we examined cohort locations where predictions from the adjusted mobile data are much larger than the stationary predictions, as shown in Figure S7, and could be considered outliers relative to the bulk of the data on the scatterplot. When mapped, the locations with the highest adjusted mobile PNC predictions relative to the stationary predictions ("high tail" locations) are all clustered around Seattle-Tacoma International Airport. Locations with relatively high mobile

predictions compared to stationary predictions that are in the middle of the distribution of PNC concentrations ("middle tail" locations) show a less clear pattern. Examining the distribution of geo-covariates at these locations compared to the other cohort locations, we observe the "high tail" locations are closer to SeaTac airport and, on average, further from the nearest stationary location (Figure S8). The "middle tail" locations are nearer, on average, to any airport and further from the nearest stationary location compared to the other cohort locations (Figure S9). Thus, these cohort locations are more likely to have higher PNC, given their proximity to airports, and may be better predicted by mobile data, given their larger distance from stationary locations.

We additionally explored whether our on-road plume-adjustment approach over adjusts PNC exposure at cohort locations that are near major roadways (Figure S10). We examined predictions at cohort locations within 250m of A1 roads and separately within 250m of a truck route, and compared predictions derived from on-road plume-adjusted mobile data to those derived from stationary data. We used the predictions from the stationary data as a reference, as the stationary sampling occurred off the side of the road at locations representative of cohort residences and thus is more representative of near-roadway residential exposure. We found that predictions derived from on-road plume-adjusted mobile data at these cohort locations were generally higher than those from the stationary measurements, consistent with the pattern in the predictions overall. Further, the few predictions that deviated from this pattern were higher for the on-road plume-adjusted data. This indicates that we are likely not over adjusting PNC exposure at locations near major roadways.

The results from Table 2 tell us about the model accuracy of the mobile predictions compared to the stationary predictions. However, given the inherent differences between mobile and stationary data, these metrics do not capture all the important aspects of how predictions from mobile data might be useful in epidemiologic studies. A key advantage of mobile measurements is the large increase in spatial information compared to stationary data, which

translates tangibly into increased spatial variation in the predictions. To understand how each model captures spatial variation, we examined the variability of the predictions at cohort locations. However, insights from this comparison are inherently clouded by measurement uncertainty, namely that with an average of 2 minutes per visit to a stationary location for the stationary data, compared to a median of 10 seconds over a 100m road segment for the mobile data, predictions from mobile data are noisier, which is captured in quantification of their spatial variation. We attempted to understand this by comparing the variance of predictions across models, with a range of comparisons to address the differences in sampling duration and number of locations between the mobile and stationary data (Note S6). Overall, we found that the mobile PNC predictions were more variable than those from the stationary locations. This analysis (Table N6S1) suggests that while noisier, predictions at cohort locations from the mobile measurements are picking up additional spatial variability, particularly at locations that are not well represented by the stationary data. This implies that the increase in spatial locations from using the mobile data, particularly in areas with little stationary sampling, allows the predictions to reflect new information. This result is not well-captured in the model performance metrics as they use reference data at the stationary locations.

In summary, we observed, as others have (20,21,23), that mobile PNC data yield higher annual averages and predictions at residential locations compared to PNC data from stationary locations (Table S9). Based on these findings and prior research showing the rapid decrease of UFP concentration as distance from roadway increases, we hypothesized that mobile data contain localized on-road plumes and therefore likely overestimate exposures at residential locations. We proposed a novel approach that employs multi-pollutant mobile and stationary measurements to adjust the mobile data to better reflect conditions that are near but not on-road. We built on previous work on plume identification (69,70) and separating mobile data into component sources (53,62) to lessen the contribution of localized on-road plumes that are not representative of residential exposure. Adjusting PNC mobile data to lessen the contribution of

on-road plumes, optimized to mobile locations that are representative of stationary locations, yielded less biased predictions at residential locations compared to a model without adjustment (Table 2). Further, our predictions at cohort locations from on-road plume-adjusted mobile data identified new spatial information (Figure 3), particularly near known UFP sources and where stationary locations were sparse. A naïve approach of rescaling the mobile measurements based on the best-fit line between the unadjusted mobile and stationary predictions did not show the same advantages. These findings suggest mobile measurements can generate unbiased estimates of annual exposure for application to epidemiologic cohort studies after suitable adjustments to diminish the influence of on-road sources that are not representative of residential exposures. We add to the literature by characterizing features that mobile data added to our prediction models that were not picked up by our stationary measurements.

Previous studies comparing measurements from mobile to short-term stationary locations sometimes remove extreme outliers in the mobile data that they state represent localized exhaust plumes (20,23). However, these studies do not verify that these outliers are from localized on-road plumes, nor do they check for the presence of on-road plumes in other measurements. In epidemiologic studies, both the scaling of the exposure as well as its uncertainty can affect inference (71,72). Thus, while our work remains to be evaluated in an epidemiologic cohort, the results we present here suggest lessening the contribution of localized on-road plumes may minimize the impact of exposure measurement error on the inferential analysis. Rescaling the exposure will impact scaling of the health effect parameter estimate, allowing it to be more consistent with the effect of residential exposure. The noise in the mobile data is a form of classical measurement error; future work is needed to determine whether the plume adjustment reduces or enlarges this source of error. Future research will need to directly evaluate the implications of using mobile versus stationary data for inference in epidemiological cohort studies.

Validation of the on-road plume-adjustment approach required thoughtful consideration of the approach. We utilized several tools to check the adjusted mobile PNC predictions. We used external validation, where we predicted PNC at our reference (stationary) locations using the on-road plume-adjusted mobile measurements, to understand the model performance using traditional performance metrics. The RMSE, MSE-based and regression-based $R^2$ address performance at the locations with data (stationary locations); however, they do not tell us how well our model performed at other locations. Selecting the best plume-adjustment method purely based on these metrics would be incomplete since our goal is to identify additional spatial information that is not captured by the stationary measurements. We addressed this by characterizing residential locations where the mobile predictions are the most discrepant from stationary predictions. These locations tend to be close to airports and further from stationary locations, suggesting that we are learning new information that the stationary measurements are not picking up.

Our study has several strengths. Our mobile monitoring campaign was designed to target an epidemiologic cohort and to provide unbiased estimates of long-term exposure. This included collecting multi-pollutant data over a year-long period in all seasons, all days of the week, and most times of the day. Our campaign included reference (stationary) sample collection, intended to capture residential exposures. This balanced design allowed for a sensible validation set from the same campaign as our mobile data, using data that were less noisy than the mobile data, and were measured at locations that represent cohort locations. We showed that on-road mobile measurements were, on average, higher than stationary measurements and this was magnified at higher exposure locations. This discrepancy is particularly important in an epidemiologic context that focuses on residential exposures and requires careful consideration of how to lessen the contribution of on-road plumes. Because our mobile platform collected multiple pollutants simultaneously, including several instruments that measured PNC, we were able to leverage multi-pollutant data to lessen the contribution of on-

road pollutant sources. By using mobile data to produce predictions at cohort locations, we had data from nineteen times the number of locations available from the stationary data alone, thus yielding potentially finer spatial resolution in our predictions. Our maps show that mobile measurements improve our ability to capture the PNC impact from aircraft traffic and industrial sources, which appear to be underestimated by using stationary data alone (Figure 3). This helps us better understand exposure differences in communities impacted by aircraft traffic and other localized sources of PNC, and thus aids in illuminating potential health disparities. Additionally, we aimed to describe model performance beyond traditional prediction model performance metrics (i.e., MSE-based and regression-based $R^2$, and RMSE) to increase our understanding of the added value of mobile monitoring data for epidemiology. Further, given our study allowed direct comparison of predictions based on stationary and mobile data, we were able to quantify the additional spatial variation captured by the mobile data, and investigate the amount of variation due to noise, by comparing to the time-sampled stationary and location-sampled mobile predictions. Finally, we focused this analysis on UFP, as UFP varies widely over space. This or a similar method could be applied to other pollutants in future studies.

Our study has a few limitations. We were not able to identify definitively on-road plumes with the available data. Thus, we proposed a method to lessen the contribution of on-road sources that are not representative of residential locations. However, it is not clear how to optimally lessen the contribution of on-road sources. While the APCS method we employed has been used previously in similar contexts (53,61–63), various other muting procedures could be used to lessen the contribution of on-road plumes (69,70). Additionally, we are assuming the long-term averages estimated from the repeat drive passes represent the true underlying long-term average at each location. While we are inherently unable to address this assumption in these data, our separate investigation of study design using a full year of measurements from regulatory monitoring sites in California showed that a balanced sampling design like the one used in this study does result in unbiased estimates of the target long-term average at each

location (54). Further, due to privacy considerations, we did not sample at participant residences. However, the stationary locations in this study were selected to be near participant residences and we have shown that they are representative of those locations (56).

Additionally, for our approach to be reproducible and generalizable, campaigns need a multipollutant mobile platform with many reference locations. This type of mobile monitoring campaign requires substantial time and resources. Further, based on findings from this work, we suggest balanced sampling across the entire campaign, including at near-road stationary locations that are representative of the locations of interest. These features are critical to appropriately use mobile measurements. Thus, campaigns should prioritize sampling at stationary locations and use good study design principles to target the study's primary quantity of interest (e.g., unbiased long-term average estimate).

Moreover, we used an unscreened P-Trak in our analyses, which is known to measure particles less accurately in lower size ranges compared to condensation particle counters (CPC) (73), and may have bearing on our estimation of PNC exposure from aircraft in particular. Furthermore, in this study we are predicting outdoor ambient exposure to PNCs at residential locations. While this is a common strategy in air pollution epidemiology, this approach does not capture all important exposure sources, since most adults spend the majority of their time indoors. Additionally, our approach characterizes the overall spatial structure in the data by using universal kriging but does not acknowledge the fine-scale spatio-temporal correlation that is inherent in mobile monitoring data. While this concern is diminished given that we aggregated the temporally balanced visits at each location and drove all routes forwards and backwards, unavoidably there will be some remaining spatio-temporal correlation since we collected data along roads on fixed routes. Finally, future studies could explore alternate approaches to lessening the contribution of on-road plumes in mobile data, such as leveraging the geographic covariates, further investigation into the contribution of the principal components to the pollutant of interest, or excluding non-residential roads in the mobile exposure model.

In conclusion, we developed an approach to lessen the contribution of localized on-road plumes in mobile monitoring data for application to exposure prediction modeling for epidemiology. Future research is needed to improve our proposed method, and to determine if this type of approach is required. Our findings suggest that leveraging both multipollutant mobile data and stationary near-road sampling at locations that are representative of the locations of interest are useful to correct mobile measurements so they are more representative of near-roadway residential exposure. Mobile sampling alone will likely lead to overestimation of near-road residential exposure, which may bias epidemiologic inference. We recommend correction of mobile measurements when the goal is residential exposure prediction for epidemiology.

**Supplemental Material**

**Note S1**. Description of mobile measurements and exclusions (unscreened P-Trak)

Prior to any exclusions, we had 3,769,325 1-second measurements, representing 8,080 100-meter segments. Prior to analysis, we excluded several 1-second measurements from the mobile data. To ensure we were not capturing measurements while the vehicle was at one of the 309 stationary sites, we excluded measurements where the distance between the measurement and the site was 0 m, or where the recorded speed was 0 km/hr and the distance between the measurement and the site was ≤ 10m. Additionally, we excluded 1,779 road segments on major interstates and highways (A1 roads), as we did not consider them to be representative of residential exposure. To reduce noise, we further excluded 108 segments with fewer than a median of 5 1-second measurements per visit, and 306 segments with fewer than 23 repeat visits ($5^{th}$ percentile of segments). Previous research has shown that at least 20-25 repeat visits are needed to yield robust estimates (17,54). Excluding A1 roads, remaining exclusions resulted in removing 2.8% of the total 1-second mobile measurements taken during the campaign (Table S1).

After summarizing at the 10-second level, there were 346,174 measurements. Prior to conducting the multi-pollutant PCA, we removed all observations with any missing measurements among those included in the PCA (Table S3), a requirement for PCA. This resulted in removing 40,931 10-second measurements.

**Table S1**. Summary of 1-second mobile measurements and exclusions

| Segment type | Segments (%) | 1-second measurements (%) |
|---|---|---|
| All, before exclusions | 8,080 (100) | 3,769,325 (100) |
| Total excluded | 2,193 (27) | 501,008 (13) |
| A1 roads | 1,779 (22) | 395,889 (10.5) |
| <5 1-s measurements/visit | 108 (1) | 28,006 (0.7) |
| <23 repeat visits | 306 (4) | 77,113 (2) |

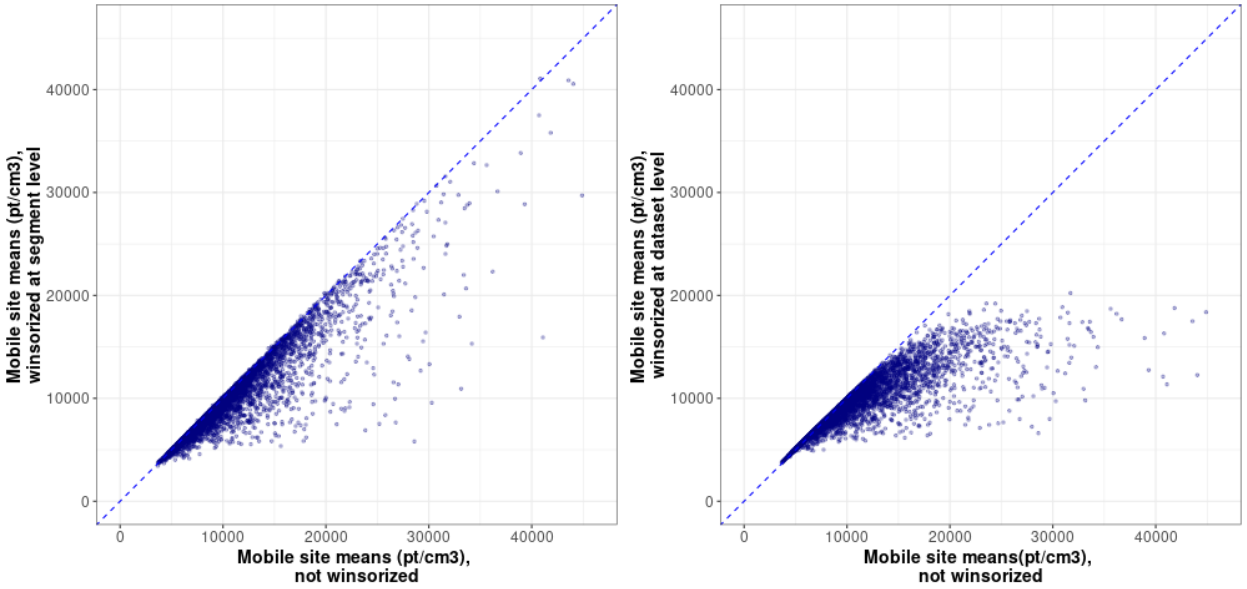


**Figure S1**. Comparison of segment means (annual averages) contrasting winsorizing at the segment vs dataset level. Mobile PNC segment means winsorized at the segment level (left) or the dataset level (right) versus mobile PNC means that are not winsorized. Winsorizing at the dataset level (right) may over-adjust spatial outliers. Additionally, winsorizing at the dataset level changes the spatial distribution of the data.

**Note S2**. Reduction of Geographic covariates

Our original dataset contained 350 geographic covariates at all cohort, mobile and stationary locations (see Table S6 in Blanco et al. 2022 for complete list (56)). Following previous methods, we reduced the geographic covariates to 191 (56,58). Detailed documentation of covariates can be found in the Data Organization and Operating Procedures (DOOP) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) (74). In brief, we first excluded variables with low variability (defined as less than 40% of the data for that variable were different from the mode). We then excluded all land use proportion variables where the maximum proportion was 20%. We also dropped variables with many outliers (defined as ≥ 2% of data had extreme values). Finally, we used log-transformed proximity variables.

**Table S2**. Description of geographic covariates in UK-PLS model

| Group | Definition | Buffers | Number |
|---|---|---|---|
| Location | Latitude, longitude<br>Lambert | NA | 4 |
| Elevation | Nearby locations above, below, at same elevation and the standard deviation | 1km<br>5km | 8 |
| Impervious surface | Average percent of area in a buffer that is covered with an impervious surface | 50m, 100m, 150m, 300m, 400m, 500m, 750m, 1km, 3km, 5km | 10 |
| Road intersections | Intersection of road types (A1, A2, A3) within buffer | 500m, 1km, 3km | 7 |
| Road lengths | Sum of A1, A2, or A3 road length within buffer | 50m, 100m, 150m, 300m, 400m, 500m, 750m, 1km, 1.5km, 3km, 5km | 16 |
| Normalized Difference Vegetation Index (NDVI) | NDVI quartiles or season averages within buffer | 250m, 500m, 1km, 2.5km, 5km, 7.5km, 10km | 35 |
| Population density | Population density within buffer | 500m, 1km, 1.5km, 2km, 2.5km, 3km, 5km, 10km, 15km | 9 |
| Land use type | Mixed forest, evergreen, water, herbaceous wetland, woody wetland, deciduous forest, barren, pasture, developed | 50m, 100m, 150m, 300m, 400m, 500m, 750m, 1km, 3km, 5km | 71 |
| Truck route lengths | Sum of truck route length within buffer | 300m, 400m, 500m, 750m, 1km, 1.5km, 3km, 5km, 10km, 15km | 10 |
| Distance to feature | Road type, airport, coast, port, rail yard, railroad, truck route, waterway, commercial land | NA | 13 |
| Distance to road intersection | Intersections between A1s, A2s, A3s, and combinations between | NA | 6 |

**Table S3**. Instruments included in the multi-pollutant PCA in the first step of the APCS model

| Instrument | Pollutant | Measurement frequency[1] |
|---|---|---|
| P-Trak 8525 (unscreened) | UFP (20-1,000 nm) | 1s |
| P-Trak 8525, with diffusion screen (screened) | UFP (36-1,000 nm) | 1s |
| P-Trak difference[2] | UFP (20-36nm) | 1s |
| DiSCmini | UFP (10-700nm) | 1s |
| CAPS $NO_2$ | $NO_2$ | 1s |
| microAeth MA200 | BC | 10s |
| LI-850 | $CO_2$ | 1s |

[1]All measurements were winsorized at the segment level and averaged to the 10s level for inclusion in the PCA.

[2]We included the paired difference between the unscreened and screened P-Trak (capturing the number of particles between 20-36 nm) to provide additional information about the particle size distribution

**Note S3**. Plume removal algorithm development

To create an algorithm to remove localized on-road plumes from the mobile measurements, we optimized our approach to a subset of 2,033 cohort locations that are well predicted by stationary locations. We identified these cohort locations based on their proximity to stationary monitoring locations: either geographic (<250m) proximity or geo-covariate proximity. Geo-covariate proximity was determined by Euclidean distance between the first two scores from a PCA of all 191 geo-covariates in a dataset with both the stationary and cohort locations. Table N3S1 shows the median and interquartile range (IQR) by location type for three representative covariates.

As noted in the methods section, the algorithm has two components: a scaling factor that multiplies the PCs that are dominated by $NO_2$, $CO_2$, and BC, and a truncation of these same components by an upper percentile threshold. In developing the algorithm, we considered two different scaling factors to the relevant PC coefficients (0.99 and 0.95), and truncation of the high end of the distribution of these components at the following percentile thresholds: 0.75, 0.8, 0.85, 0.9. Scaling the coefficients by a factor of 0.99, and truncating the components at the $75^{th}$ percentile yielded the best slope (closest to 1) and intercept (closest to 0) combination when compared to the predictions from the stationary measurements (Figure N3S1) at the 2,033 selected locations. We focused on optimizing the slope first as a slope different from one may lead to biased inference in an epidemiologic analysis, whereas an intercept different from zero will not lead to biased inference (72,75). However, we did not choose the algorithm with the slope closest to 1 (0.75 truncating, 0.95 scaling), as the very negative intercept may lead to overfitting. Thus, we chose the algorithm that optimizes the slope, while keeping the intercept close to 0. We then applied this algorithm to the full set of cohort locations (Figure 2, bottom).

We examined the PNC predictions from on-road PNC data compared to stationary predictions, both adjusted using all 8 combinations of scaling and truncating, and unadjusted (Table N3S2, Figure N3S2).

To check our plume algorithm, we evaluated the impact on predictions of truncating the top 25% of the PCs that are dominated by $NO_2$, $CO_2$, and BC to create the adjusted PNC. To do so, we identified the locations with a large difference in predictions based on only the 0.99 scaling factor for these PCs compared to the full adjustment (scaling the coefficients and truncating the scores). These locations are generally closer to A1 roads, truck routes and railroads, but further from airports (Figure N3S3). This is consistent with our algorithm, as we chose to focus on truncating the PC scores, which characterized by roadway and truck features, and we did not adjust the PCs with a strong signal near the airport.

**Table N3S1.** Description of groups of cohort locations compared to stationary locations

| | N | UFP prediction (pt/cm$^3$) (Median (IQR)) | Distance to A1 Road (m) (Median (IQR)) | Distance to truck route (m) (Median (IQR)) | Distance to nearest stop location (m) |
|---|---|---|---|---|---|
| Stationary locations | 309 | N/A | 1,513 (655; 2,594) | 900 (330; 1,827) | N/A |
| All cohort locations | 11,904 | 7,280 (5,960; 8,700) | 1,885 (904; 3,066) | 1,002 (445; 2,186) | 603 (373, 962) |
| Select cohort locations[1] | 2,033 | 7,650 (6,460; 9,100) | 1,281 (629; 2,322) | 657 (314; 1,155) | 330 (146, 700) |
| Other cohort locations | 9,871 | 7,190 (5,870; 8,630) | 2,023 (999; 3,289) | 1,126 (296; 2,417) | 641 (426, 1,043) |

[1]Locations chosen based on geographic and/or geo-covariate proximity to stationary locations, and used for developing the plume identification algorithm.

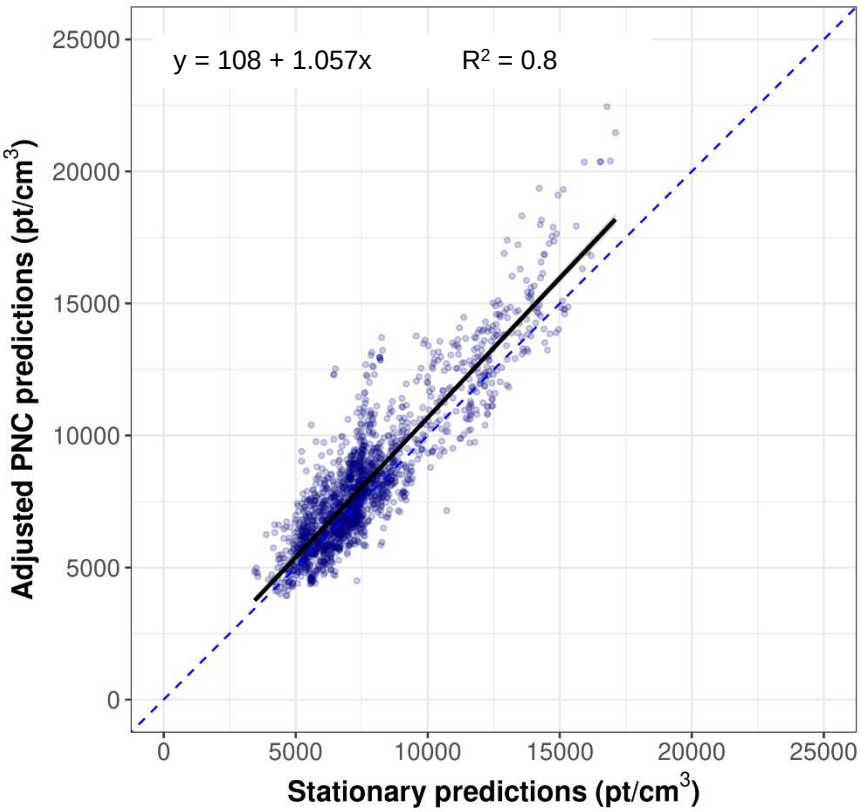


**Figure N3S1**. PNC predictions from on-road plume-adjusted data (0.99 scaled, 0.75 truncated) at 2,033 selected cohort locations (y-axis) versus predictions from stationary data (x-axis)

**Table N3S2.** Intercept and slope from mobile PNC predictions versus stationary predictions at 2,033 selected cohort locations for all combinations of scaling factors and truncation of components.

| PNC mobile prediction model | Intercept | Slope |
|---|---|---|
| Unadjusted | 584 | 1.099 |
| Scaled: 0.95, truncated: 0.75 | -648 | 1.050 |
| **Scaled: 0.99, truncated: 0.75[1]** | **108** | **1.057** |
| Scaled: 0.95, truncated: 0.80 | -482 | 1.085 |
| Scaled: 0.99, truncated: 0.80 | 218 | 1.062 |
| Scaled: 0.95, truncated: 0.85 | -435 | 1.058 |
| Scaled: 0.99, truncated: 0.85 | 332 | 1.065 |
| Scaled: 0.95, truncated: 0.90 | -333 | 1.061 |
| Scaled: 0.99, truncated: 0.90 | 438 | 1.069 |

[1]Selected PNC prediction model from on-road plume-adjusted data

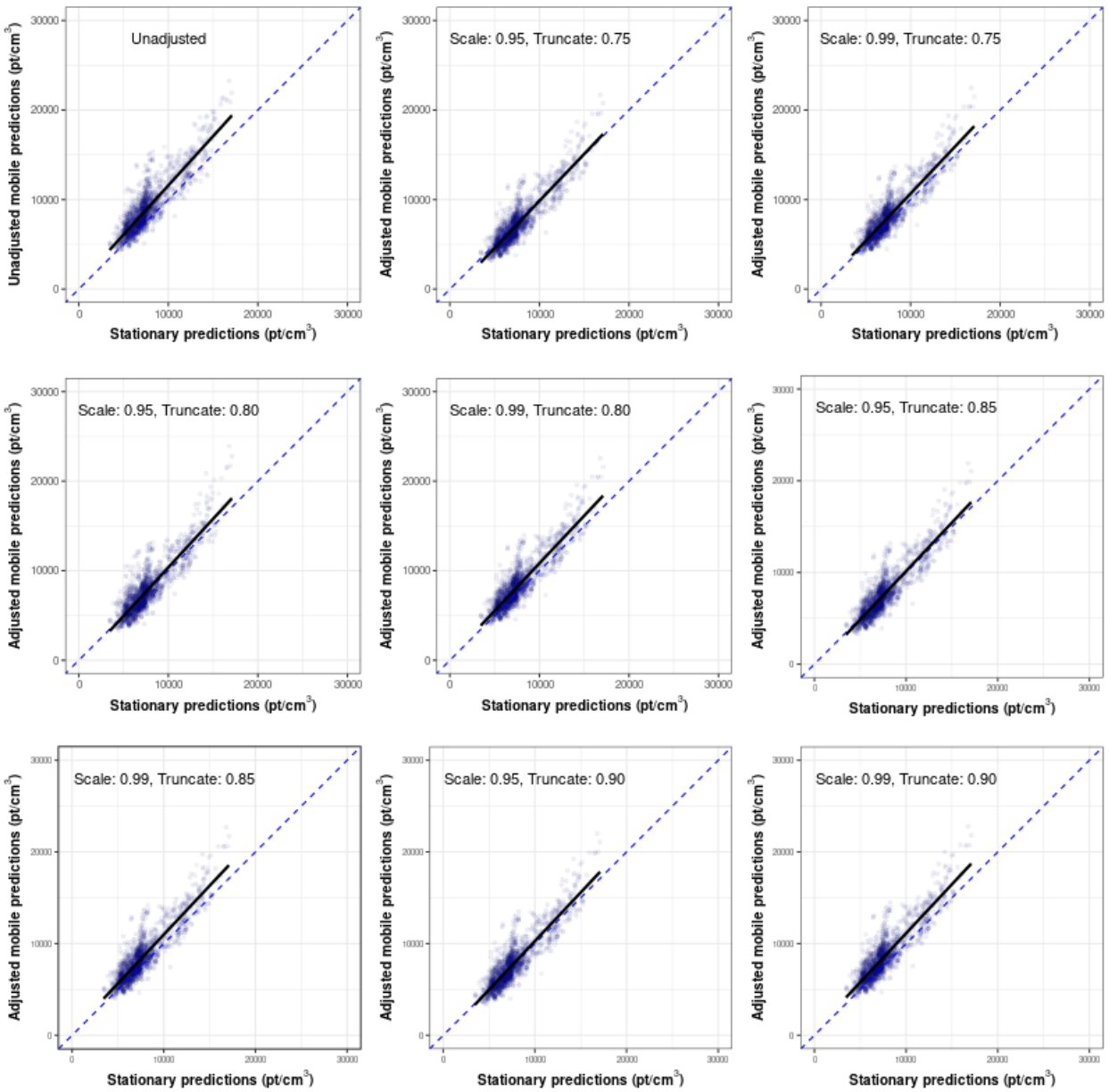


**Figure N3S2**. All combinations of PNC predictions at the 2,033 selected cohort locations from on-road data vs. stationary data. The upper left plot shows unadjusted on-road predictions; the remaining plots show plume-adjusted on-road data. The top-right plot is the algorithm we chose to adjust the mobile data for this analysis.

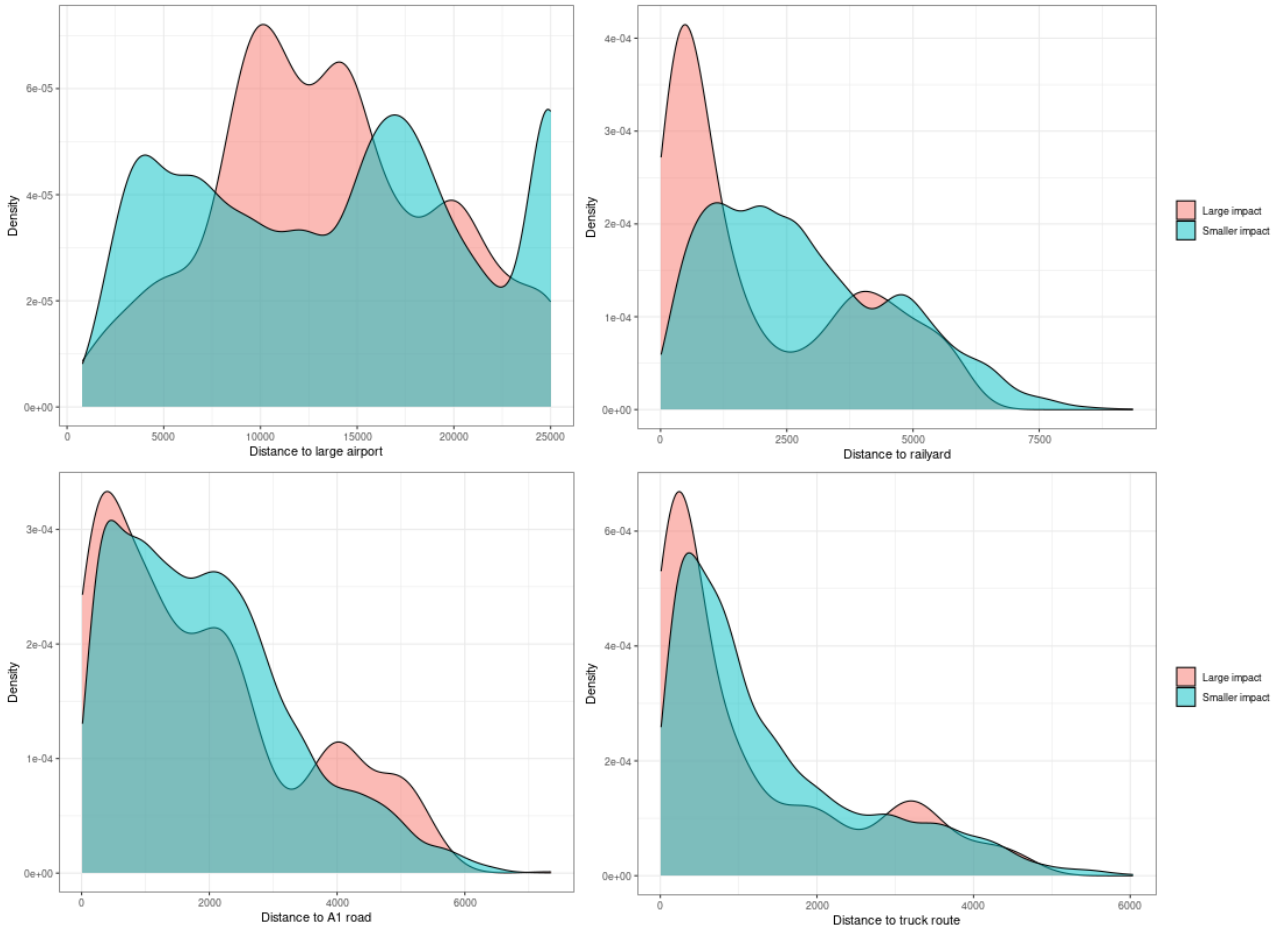


**Figure N3S3**. Distribution of distance between cohort locations and PNC sources, classified by cohort locations that are most impacted by the truncation part of the plume algorithm

**Table S4**. Summary of unadjusted mobile PNC measurement data (pt/cm$^3$)

| Category | Min | 25% | Median | 75% | Max |
|---|---|---|---|---|---|
| 1s | 0 | 4,400 | 7,130 | 11,400 | 500,000 |
| 10s | 1 | 4,420 | 7,190 | 11,480 | 500,000 |
| 10s on-road plume-adjusted | 1 | 3,900 | 6,380 | 10,460 | 453,240 |

**Table S5.** Summary of selected geo-covariate values for the 309 stationary and 5,887 mobile locations

| Location | Road class (N (%))[1] | | | Population density (per 1km$^2$) (Median (IQR)) | Distance to large airport[2] (km) (Median (IQR)) | Distance to A1 road[3] (m) (Median (IQR)) |
|---|---|---|---|---|---|---|
| | A2 | A3 | A4 | | | |
| Stationary | 0 | 27 (9) | 282 (91) | 5,179 (3,682; 7,488) | 13.8 (7.1, 20.6) | 1,528 (663, 2,595) |
| Mobile | 280 (5) | 3,060 (52) | 2,538 (43) | 4,808 (3,398; 6,515) | 15.0 (8.2, 21.8) | 1,694 (744; 2,779) |

[1]Road class refers to the classification of road segments based on TeleAtlas road data. A1: primary road with limited access or interstate (these were removed from the mobile data); A2: primary roads without limited access; A3: secondary and connecting roads; A4: local, neighborhood, and rural roads.
[2]Large airport refers to Seattle-Tacoma International Airport
[3]Represents distance to nearest A1 road, defined as interstates and major highways

**Note S4.** Comparisons between stationary and mobile locations

In order to address the hypothesis that on-road data are generally higher than stationary near-road data due to presence of on-road sources, we examined annual averages based on geographic proximity (arriving or departing each site) versus stationary annual averages (Figure N4S1). Because these locations are highly correlated with the stationary measurements, both spatially and temporally, their generalizability to the full set of mobile measurements is limited. Nonetheless, they are generally higher than the annual averages at adjacent stationary locations. It is notable that the departure averages are more similar to the stationary averages than the arrival averages. This makes sense since a vehicle is more likely to wait for traffic to clear when departing from a stationary location, whereas it can still safely pull to the side of the road when arriving at this location.

Annual averages at mobile locations characterized as near stationary locations based on proximity in geo-covariate space (Figure N4S2) show these 309 mobile averages are generally higher than those at the stationary locations. To rule out the alternative hypothesis that this pattern is an artifact of the PCA matching process, we repeated our analyses of locations matched based on geo-covariate proximity for the mobile data alone as well as for the stationary data alone. For each dataset, we split the data randomly in half, and matched this half of the data with values in the other half based on geo-covariate proximity. Figure N4S3 shows each of these comparisons, with the 2,943 pairs of mobile averages depicted in the left plot, and the 154 pairs of stationary averages depicted in the right plot. Both comparisons show scatter around the 1:1 line, with scatter both above and below the 1:1 line. Results from sign tests indicate that close to half of points were above the 1:1 line for each comparison, where the probability of being above the 1:1 line was not statistically different from 0.5 (stationary vs stationary: 0.53 (95% CI: 0.45, 0.61); mobile vs mobile: 0.49 (95% CI:0.47, 0.51).

The comparisons in Figures N4S1 and N4S2 show most of the scatter above the 1:1 line. To confirm this, we conducted additional sign tests. For the arrival versus stationary means (Figure N4S1, left), 271 of 309 points were above the 1:1 line (0.88 probability; 95% CI: 0.84, 0.91). We report similar results for departure versus stationary means (0.70 probability above 1:1 line, 95% CI: 065, 0.75), and mobile locations close to stationary locations based on geo-covariate proximity vs stationary means (Figure N4S2; 0.77 probability above 1:1 line; 95% CI: 0.72, 0.82). Thus, we can conclude that the mobile data are generally higher than the stationary data, likely due to the presence of on-road sources.

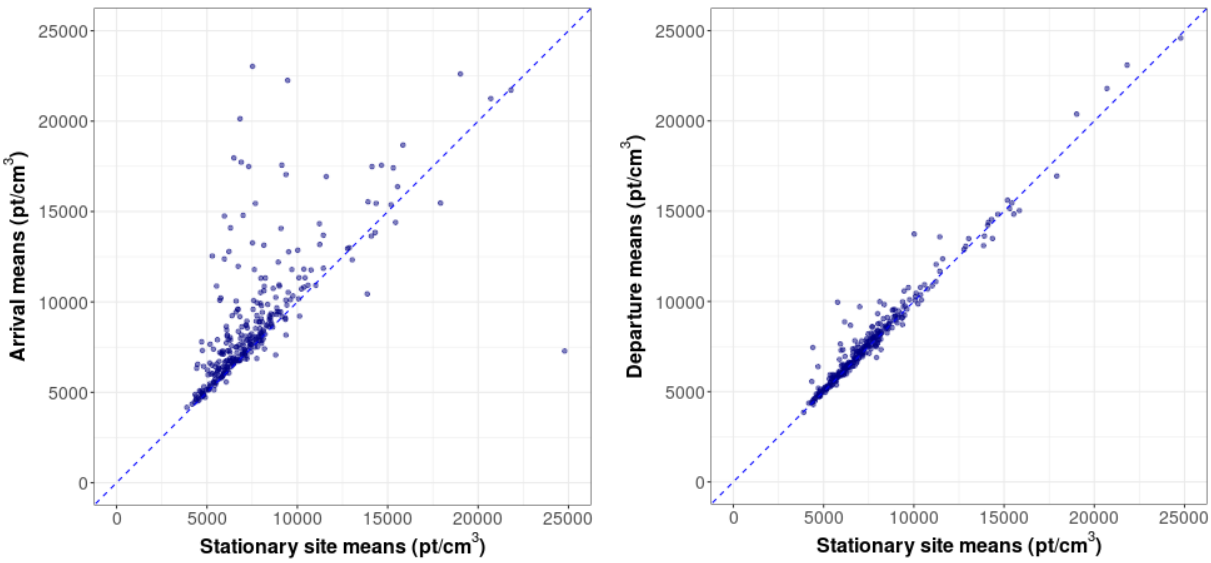


**Figure N4S1**. Annual averages from mobile versus stationary data. Annual averages from 1-second mobile PNC measurements arriving (left) at or departing (right) from each stationary location versus the stationary data annual averages.

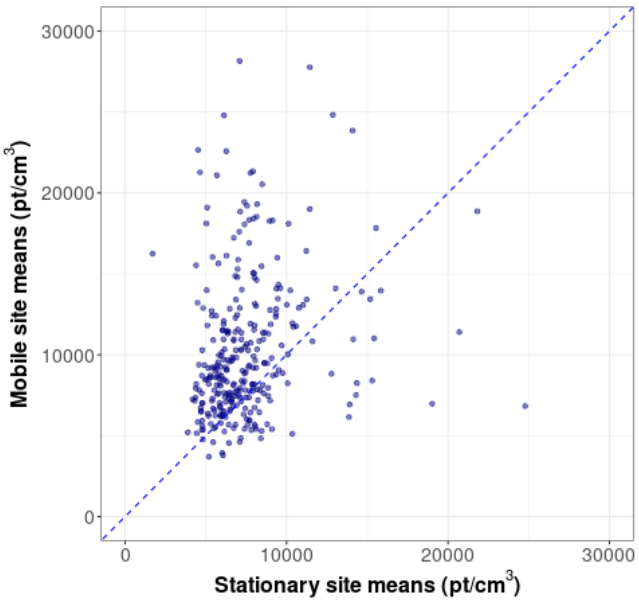


**Figure N4S2**. Annual averages from 1-second mobile versus stationary PNC data with pairing based on geo-covariate proximity

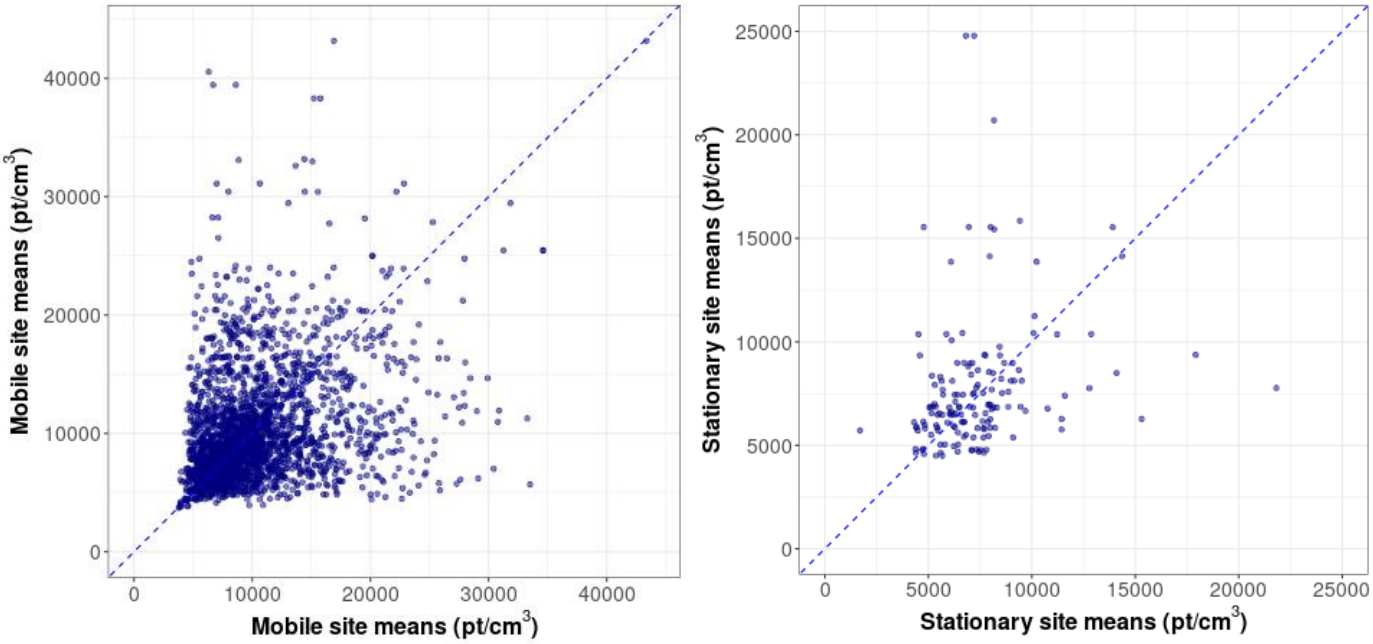


**Figure N4S3**. PNC annual averages from 1-second measurements at mobile versus mobile locations (left); and at stationary versus stationary locations (right), paired based on geo-covariate proximity. Datasets were divided in half before pairing.

**Table S6.** PCA factor loadings for the first three components from the multi-pollutant 10s winsorized unadjusted data

| Variable | PC1 | PC2 | PC3 |
|---|---|---|---|
| UFP – unscreened | **0.935** | 0.224 | 0.244 |
| UFP – screened | **0.799** | 0.3 | 0.398 |
| UFP (unscreened-screened) | **0.945** | 0.162 | 0.124 |
| UFP (DiSCmini) | **0.928** | 0.139 | 0.157 |
| $NO_2$ | 0.289 | **0.871** | 0.133 |
| BC | 0.286 | 0.225 | **0.905** |
| $CO_2$ | 0.156 | **0.728** | 0.450 |
| Proportion Variance explained | 0.494 | 0.218 | 0.185 |

Note: Bolded numbers reflect large contributions with loadings greater than 0.7. PNC measures include all listed UFP variables.

**Table S7.** Summary of first three PCA components from the multi-pollutant 10s winsorized unadjusted data

| Component | Variance explained (%) | Main contributing pollutants | Potential sources |
|---|---|---|---|
| 1 | 49 | UFP, UFP difference[1] | Aircraft |
| 2 | 22 | $NO_2$, $CO_2$ | Car exhaust |
| 3 | 19 | BC | Trucks, diesel exhaust |

[1]UFP difference is the difference between the unscreened and screened P-Trak, which measures the number of particles in the size range 20-36 nm.

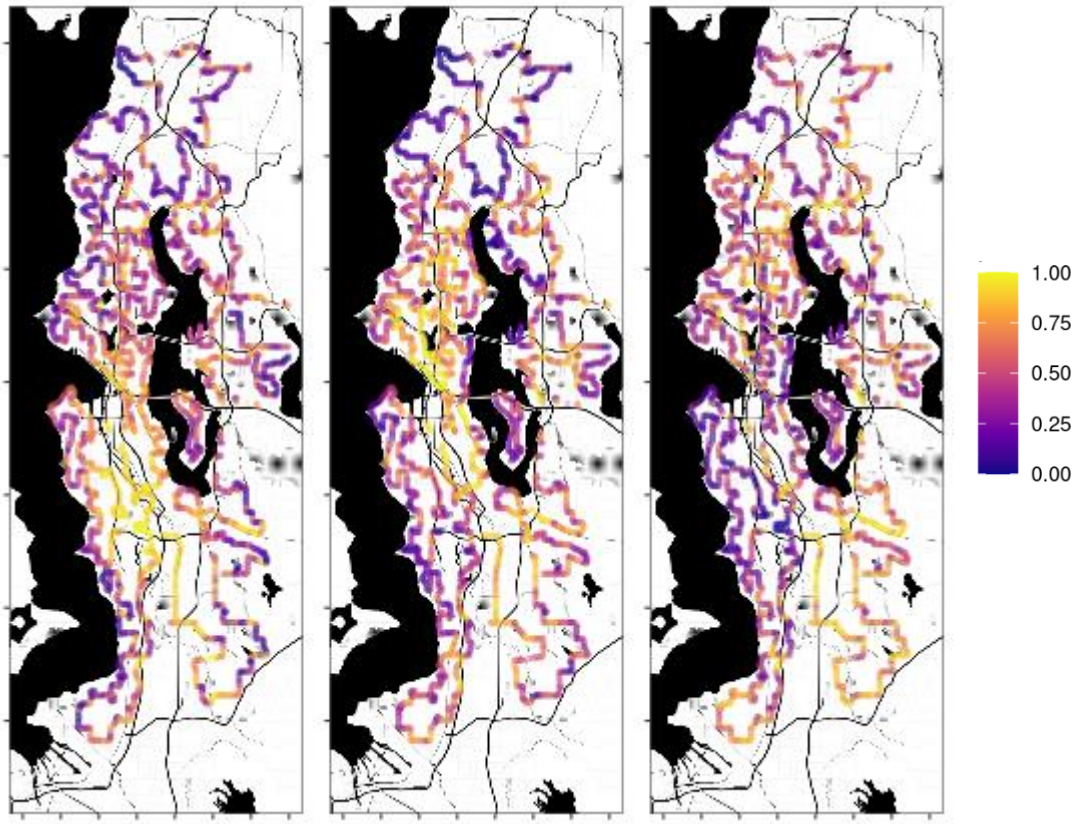


**Figure S2**. Principal component score rank (0-1). (Left to right: PC1, PC2, PC3). PC1 (left) defined as an aircraft feature, PC2 (center) defined as a roadway feature, PC3 (right) defined as a truck/diesel feature.

**Table S8**: Summary of PNC annual average predictions from unadjusted and adjusted mobile measurements (pt/cm$^3$) at 11,904 cohort locations

| Category | Min | 25% | Median | 75% | Max |
|---|---|---|---|---|---|
| Unadjusted predictions[1] | 970 | 6,520 | 7,890 | 9,460 | 28,670 |
| On-road plume-adjusted predictions[2] | 780 | 5,830 | 7,140 | 8,540 | 28,260 |
| Stationary predictions | 950 | 5,930 | 6,990 | 7,950 | 18,740 |
| Difference, unadjusted–adjusted mobile predictions | 190 | 690 | 750 | 920 | 410 |
| Difference, adjusted mobile–stationary predictions | -170 | -100 | 150 | 590 | 9,520 |

[1]Mobile PNC measurements were summarized to the 10 second level and winsorized at the segment level before estimating predictions at cohort locations
[2]0.99 scaled and 0.75 truncated

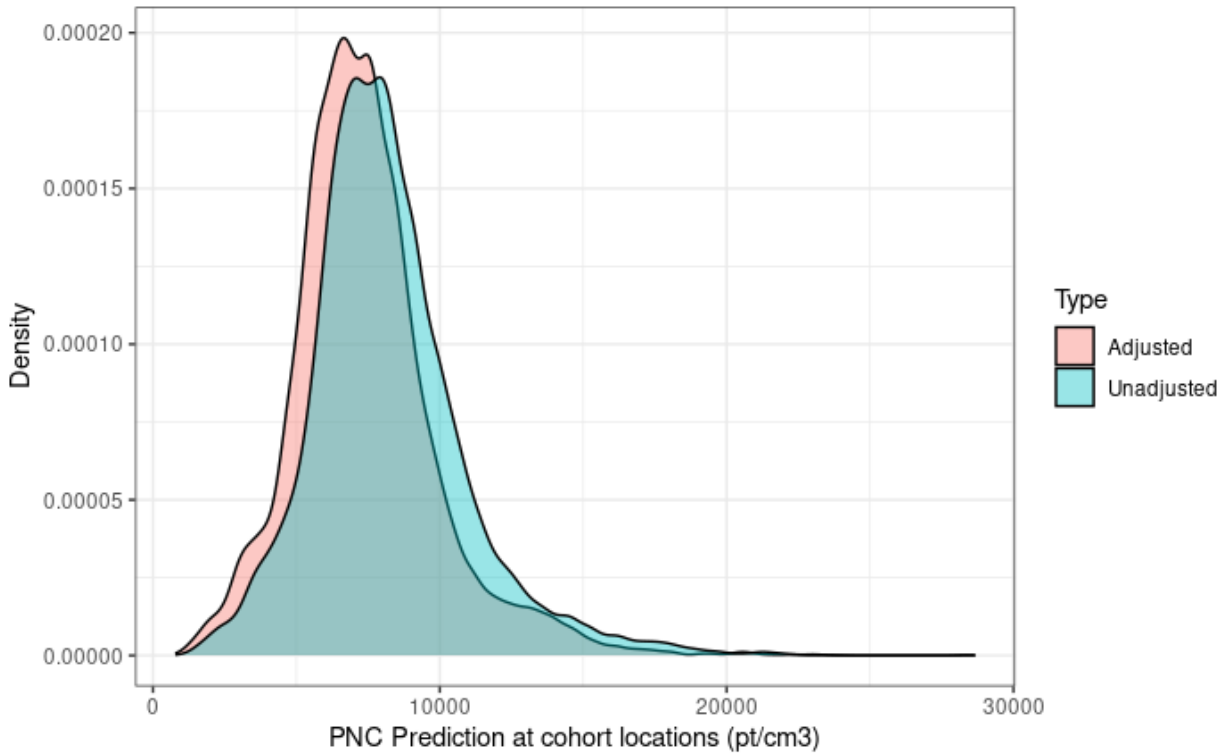


**Figure S3**. Distribution of PNC exposure predictions at cohort locations from unadjusted and on-road plume-adjusted mobile data. 8.4% of cohort locations shifted quartiles of PNC exposure after plume adjustment.

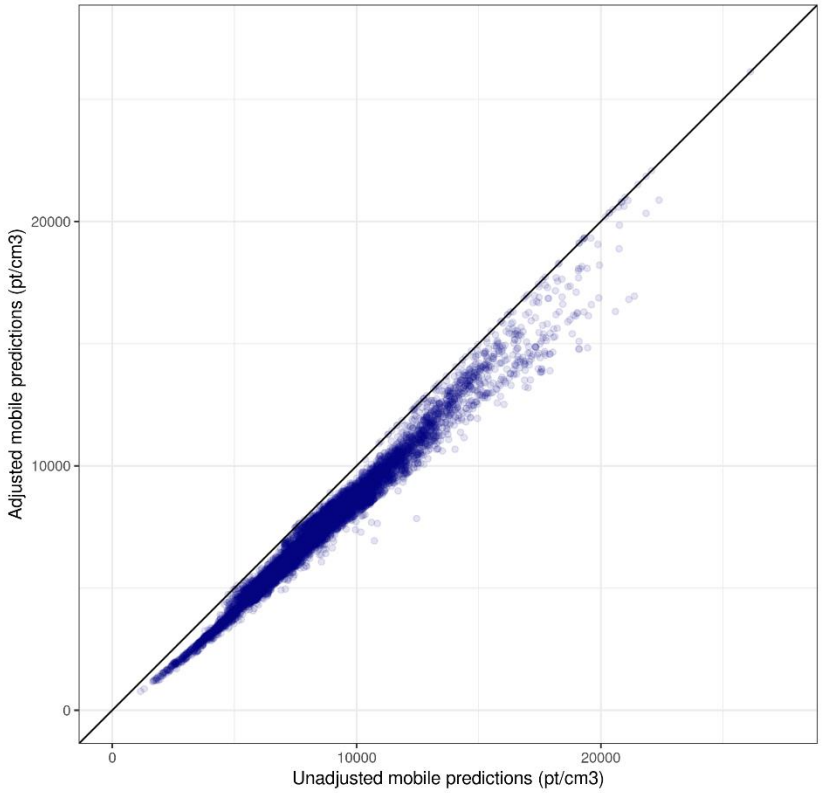


**Figure S4.** PNC exposure predictions at cohort locations comparing predictions from on-road plume-adjusted mobile measurements (y-axis) to those from unadjusted mobile measurements (x-axis).

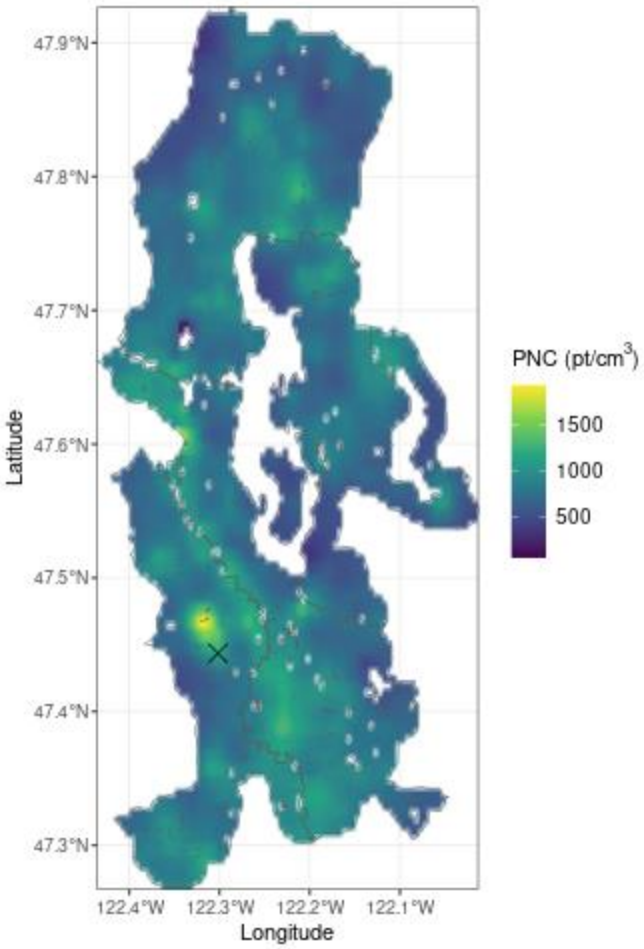


**Figure S5**. Difference between predictions derived from unadjusted PNC mobile measurements and on-road plume-adjusted PNC mobile measurements. The magnitude of this difference is smaller across most the study area compared to the difference between the predictions derived from the on-road plume-adjusted PNC mobile measurements and those from the stationary PNC measurements (Figure 3).

**Note S5.** Additional external model validation

To further examine the model performance, we also considered adjusting the mobile measurements by the best fit line estimated by comparing the unadjusted predictions to the stationary predictions (Figure 2, top). We then used these scaled measurements, called the "best fit adjusted" PNC measurements, to predict PNC at the stationary sites using external validation. We found performance metrics that are nearly the same and slightly better than the mobile model based on the on-road plume-adjusted PNC measurements (Table N5S1, last row). We note that this assessment is not a purely external assessment as it uses the stationary data twice, once to rescale the mobile PNC data to optimally align with the stationary data predictions, and then to "externally" validate them to the stationary data.

To determine whether these apparently better model performance statistics were providing evidence that the best fit adjusted PNC measurements are preferable to the on-road plume-adjusted PNC measurements, we further examined results from this model by plotting predictions at cohort locations. In Figure N5S1 we show the comparison of predictions from the best fit adjusted PNC model (y axis) versus those from the model using the on-road plume-adjusted data, and from a model based on the stationary data. We found that merely rescaling PNC measurements to calibrate them to the stationary data yields predictions that are slightly lower than the predictions from the stationary data, and lower than those from the on-road plume-adjusted mobile data (Figure N5S1, left and right plots, respectively). We also mapped the best-fit adjusted PNC predictions, and the difference between the on-road plume-adjusted and best-fit adjusted predictions (Figure N5S2). We find that this scaling method does not pick up the high PNC concentrations around the airport and may over-adjust predictions across much of the map. However, we do not have measurements collected at cohort locations, and thus cannot fully confirm whether this model misses high PNC concentrations, or if other scaling methods overestimate these high PNC concentrations.

We additionally included external validation for the 7 other plume algorithm adjustments we examined (discussed in Supplemental Note S3).

**Table N5S1**. Model performance statistics at stationary locations (N=309) for cross-validated predictions using stationary data and external predictions using mobile data. Mobile data are either unadjusted or adjusted

| Model | Regression-based $R^2$ | MSE-based $R^2$ | RMSE | Location-visits (N)† | Hours sampled |
|---|---|---|---|---|---|
| CV-Stationary (comparison)‡ | 0.77 | 0.77 | 1,180 | 8,652 | 288 |
| Mobile locations, unadjusted | 0.73 | 0.60 | 1,800 | 165,733 | 908 |
| Mobile locations, adjusted | | | | | |
| **0.99 scaled, 0.75 truncated** | **0.73** | **0.71** | **1,510** | **165,733** | **908** |
| 0.95 scaled, 0.75 truncated | 0.73 | 0.67 | 1,610 | 165,733 | 908 |
| 0.99 scaled, 0.8 truncated | 0.73 | 0.70 | 1,540 | 165,733 | 908 |
| 0.95 scaled, 0.8 truncated | 0.73 | 0.69 | 1,570 | 165,733 | 908 |
| 0.99 scaled, 0.85 truncated | 0.73 | 0.69 | 1,580 | 165,733 | 908 |
| 0.95 scaled, 0.85 truncated | 0.73 | 0.70 | 1,540 | 165,733 | 908 |
| 0.99 scaled, 0.9 truncated | 0.73 | 0.67 | 1,620 | 165,733 | 908 |
| 0.95 scaled, 0.9 truncated | 0.73 | 0.71 | 1,530 | 165,733 | 908 |
| Best-fit adjusted PNC predictions* | 0.73 | 0.71 | 1,520 | 165,733 | 908 |
| Mobile locations (adjusted) by geo-covariate distance** | 0.52 | 0.43 | 2,140 | 8,652 | 45 |

†Location-visits is the sum of number of visits across all locations included in the model.
‡Stationary data are inherently less noisy than the mobile data because their sampling duration was 2 minutes versus a median of 10 seconds (IQR: 8,17), and were based on 1-s versus 10-s measurements. Thus, we expect the RMSE from the CV-Stationary model to be lower than the mobile model.
*Mobile measurements (at the 10-second level and winsorized at the site level) were scaled by the best fit line between unadjusted mobile and stationary predictions (Figure 2, top).
**A set of 309 mobile locations that are like the stationary locations based on geo-covariates

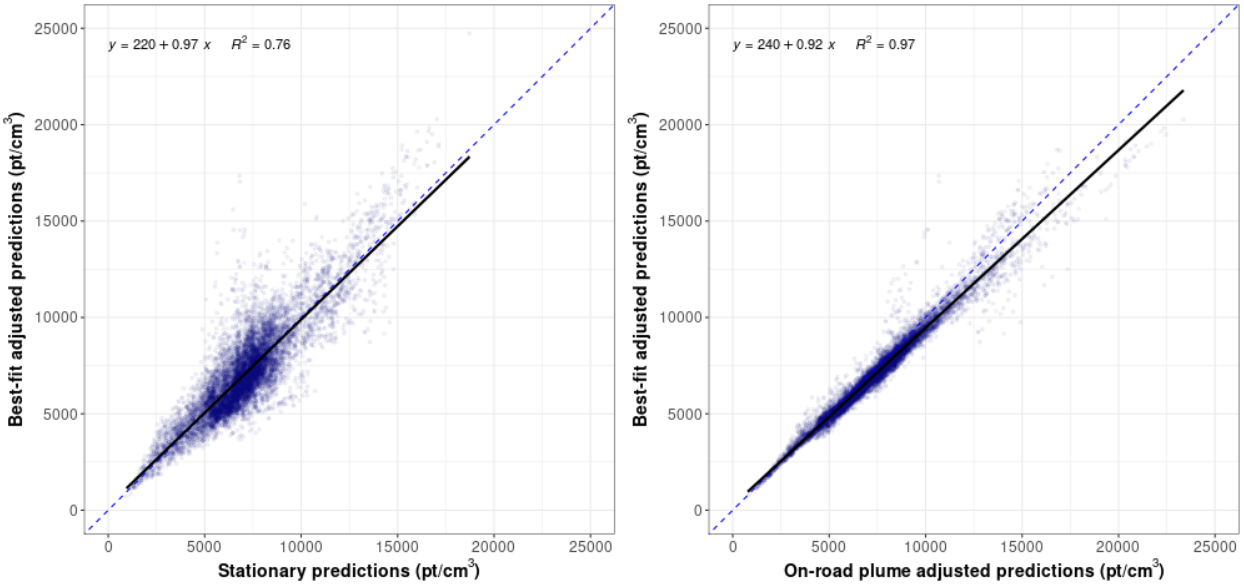


**Figure N5S1**. Predictions from best-fit adjusted predictions (y axis) versus stationary predictions (left) and on-road plume-adjusted predictions (0.99 scaled, 0.75 truncated) (right)

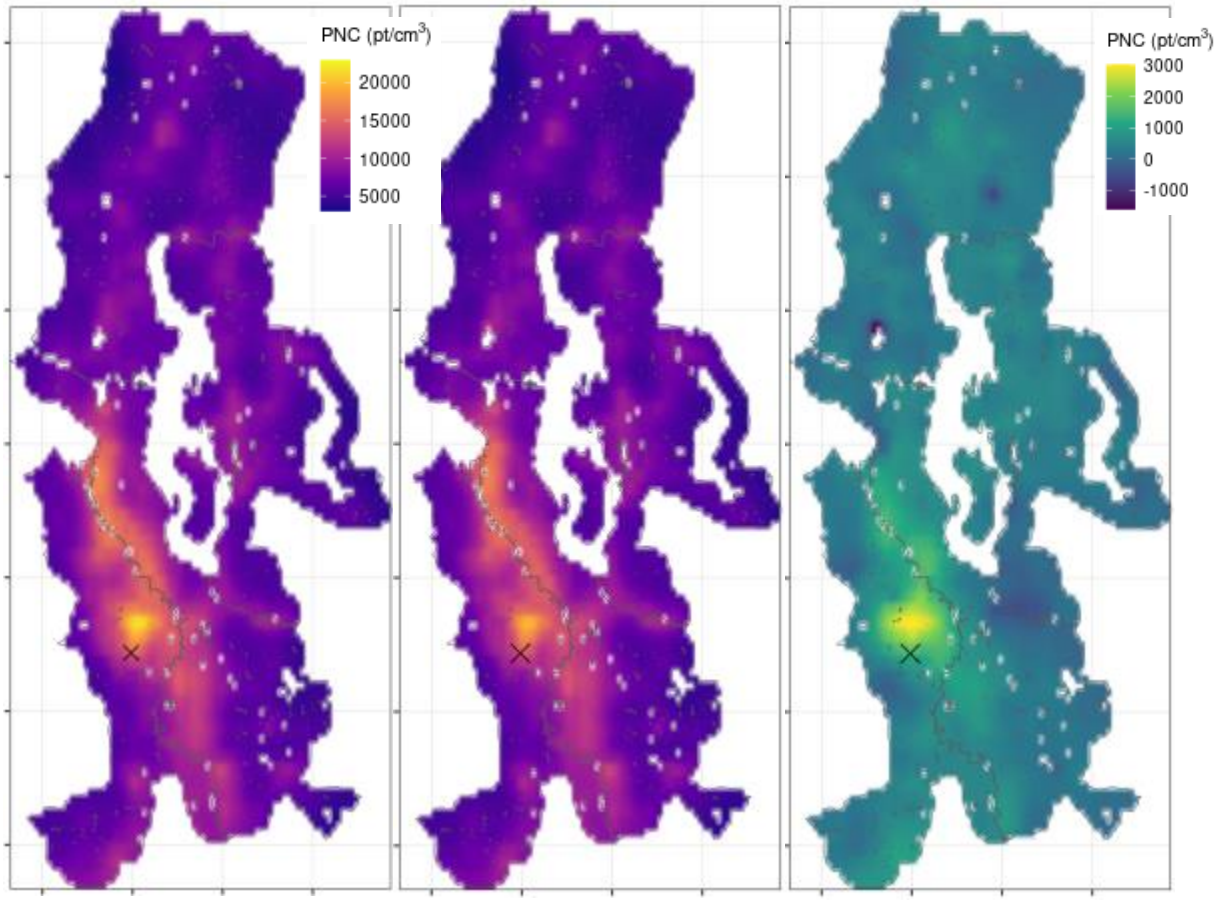


**Figure N5S2**. PNC predictions from on-road plume-adjusted mobile data (0.99 scaled, 0.75 truncated) (left); PNC predictions from the best-fit adjusted mobile data (center); Difference in adjusted PNC predictions: on-road plume-adjusted, minus best-fit adjusted predictions (right). The black ‘x’ represents Seattle-Tacoma International Airport.

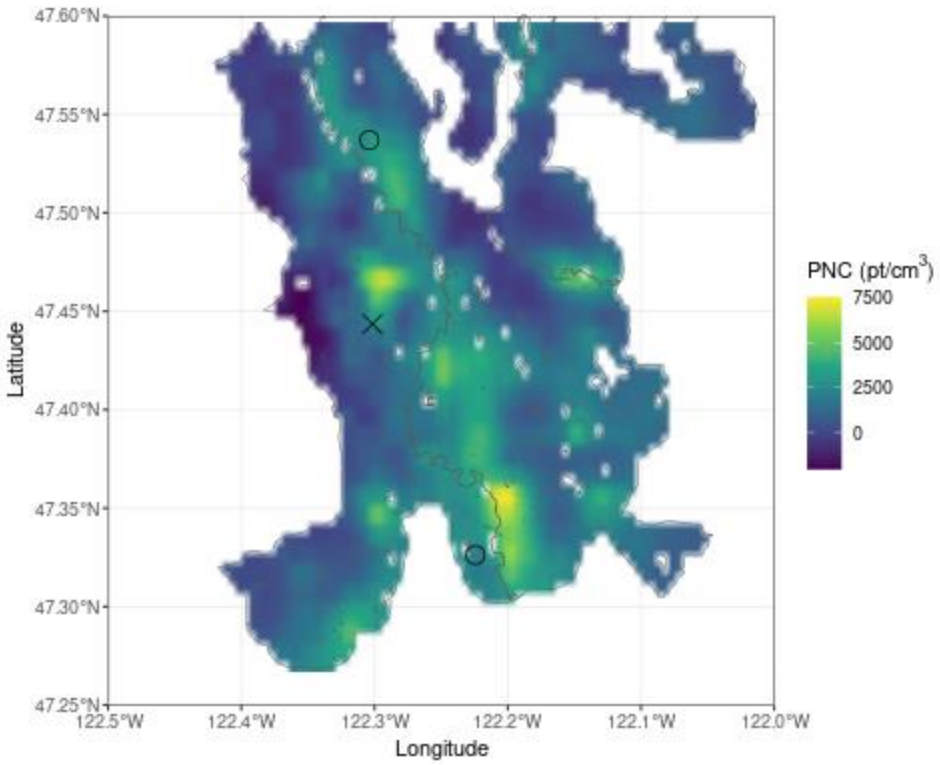


**Figure S6.** Difference in predictions near the airport. Zoomed in PNC predictions, showing on-road plume-adjusted mobile predictions minus stationary predictions. The black ‘x’ indicates Seattle-Tacoma international airport, and the black circles indicate regional airports.

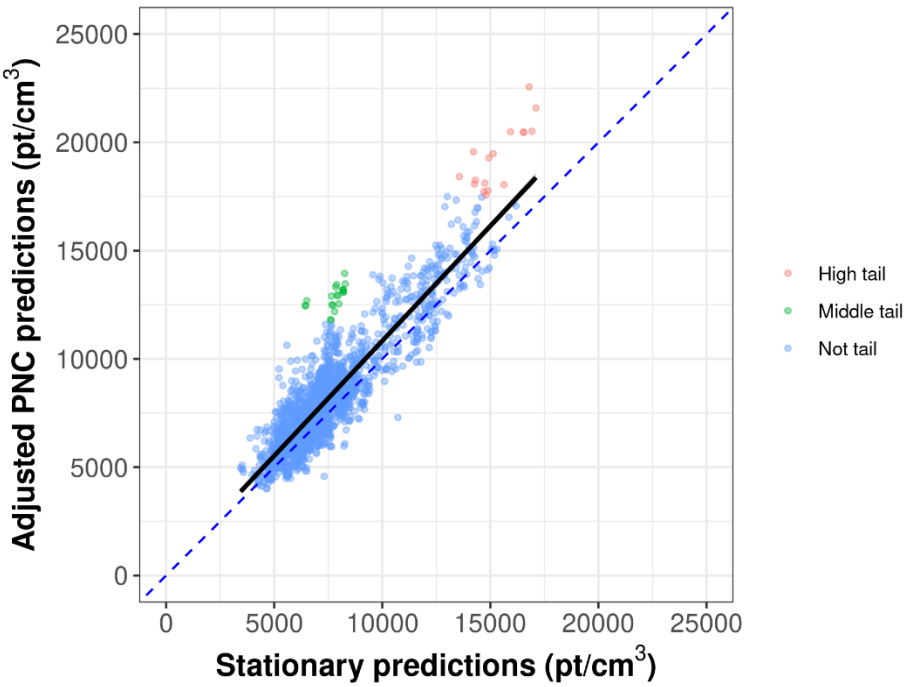


**Figure S7.** Predictions different from bulk of data (best fit line). On-road plume-adjusted mobile PNC predictions at select cohort locations (y-axis) versus predictions from stationary data (x-axis). Colors indicate predictions that deviate from the bulk of the data at the high end of the prediction range (red), or in the middle (green)

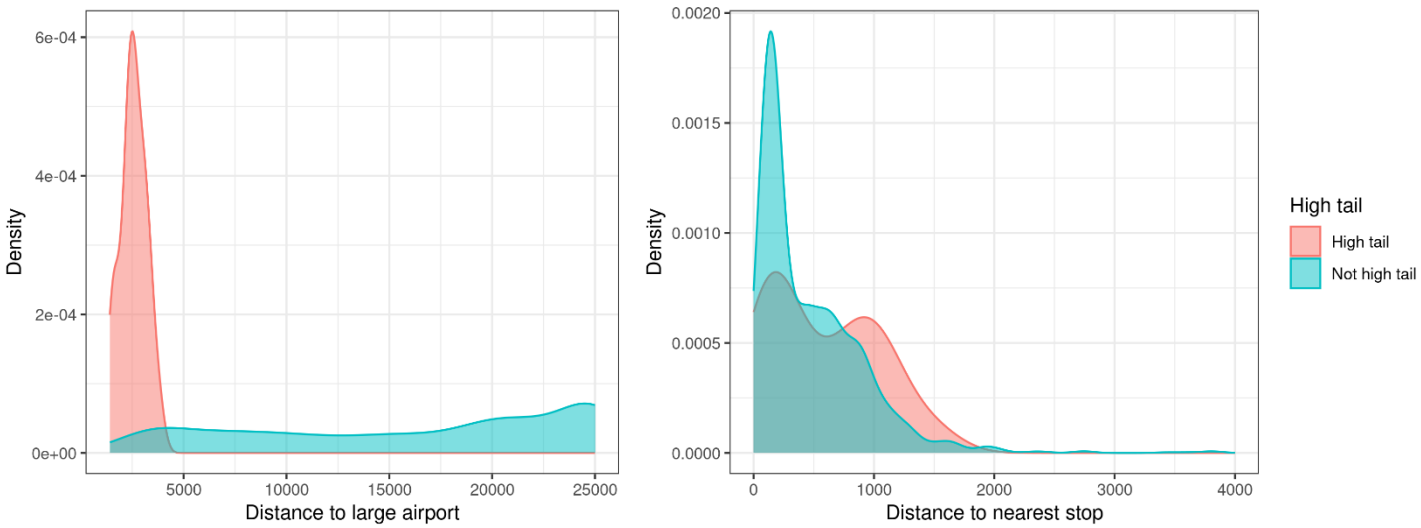


**Figure S8**. Distribution of distances between cohort locations and Seattle-Tacoma airport (left) and nearest stationary site (right), classifying by “high tail” locations versus all other locations

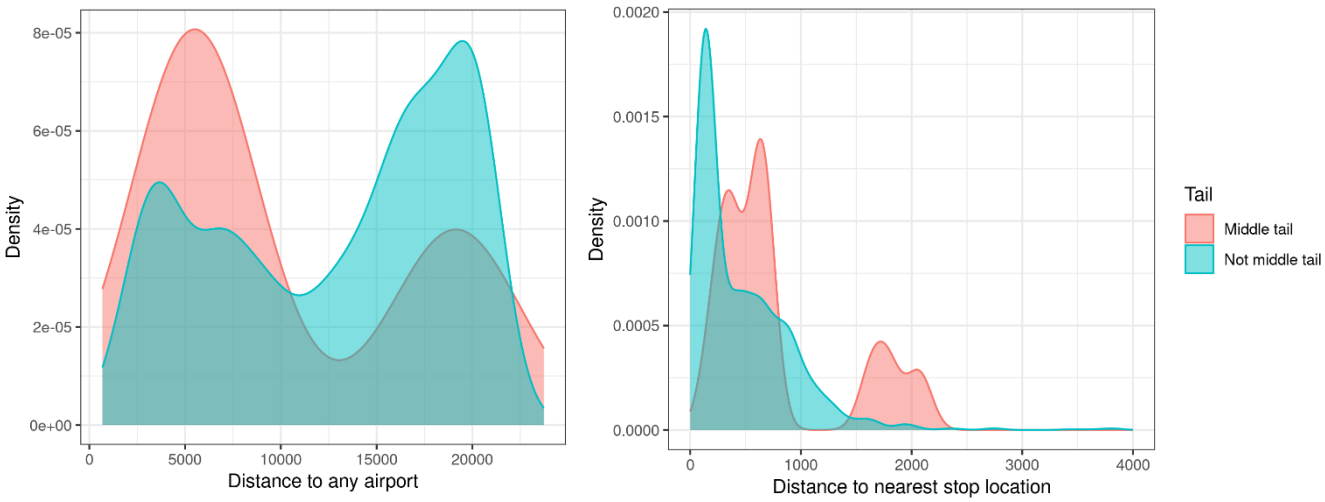


**Figure S9.** Distribution of distances between cohort locations and nearest airport (left) and nearest stationary site (right), classifying by “middle tail” locations versus all other locations

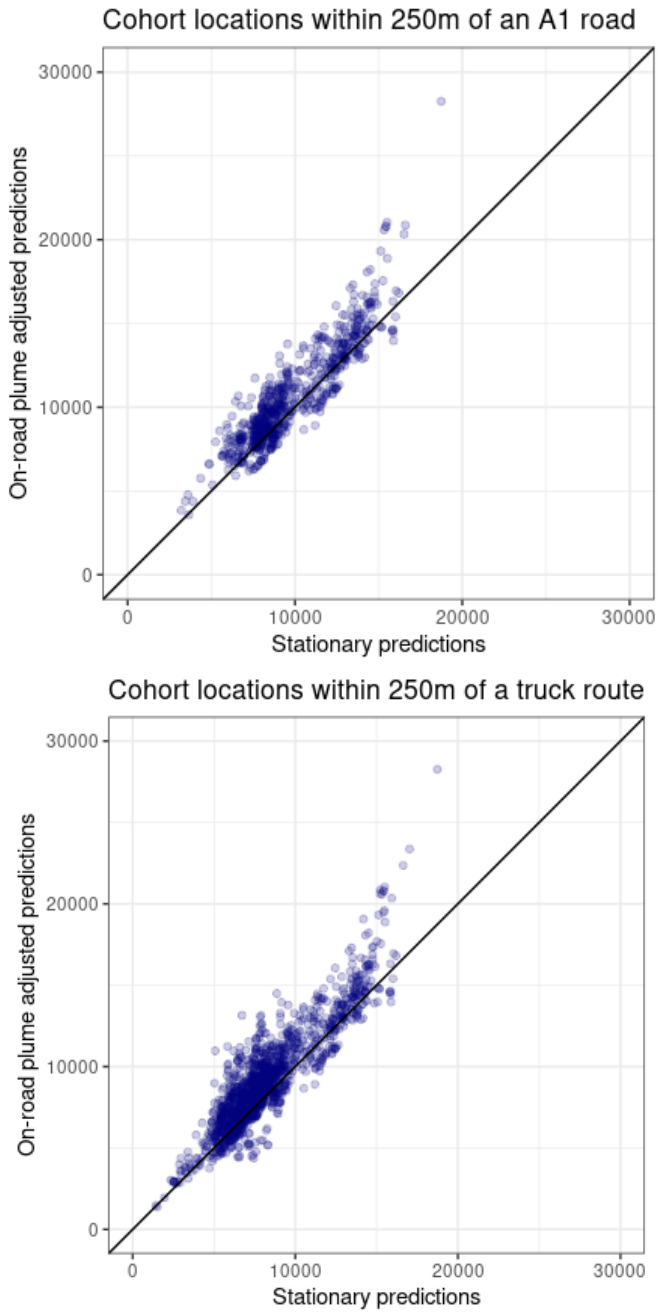


**Figure S10:** Predictions at cohort locations within 250m of an A1-road (top) or within 250m of a truck route (bottom) derived from stationary measurements (x-axis) versus from mobile plume-adjusted measurements (y-axis)

**Note S6.** Variability of predictions at cohort locations

The results from Table 2 tell us about the accuracy and precision of the mobile predictions when compared to the stationary observations (pure external validation) and they can be compared to these same statistics for the stationary predictions, which must be cross-validated to provide more comparable out-of-sample prediction validation. However, these performance metrics do not capture all the important features of the mobile data predictions, particularly since the uncertainty of the mobile data is inherently larger than for the stationary data, and this uncertainty is incorporated into the performance statistics. Thus, for instance, the RMSE for the adjusted mobile data is larger than the cross-validated RMSE from the stationary data. At least some of this increase is most likely driven by the difference in duration of sampling at the mobile vs. stationary locations. Because the mobile data represent 5,887 rather than 309 locations, we hypothesize that the mobile data model will capture more underlying spatial variation in the exposure surface than the stationary data. Quantifying this clearly is challenging, however, given there are differences in mobile and stationary data based on location, location accuracy, duration of sampling, and inclusion of adjustment for on-road plumes. The following analysis provides a preliminary investigation. Future work is needed to promote better understanding of the added value of on-road data for exposure predictions.

To gain insight into how each model captures spatial variation, we examined the variability of the predictions at cohort locations. Table N6S1 shows the variance of the predictions at cohort locations derived from mobile and stationary data separately, overall and by distance to Seattle-Tacoma International Airport, and distance to the nearest stationary site. We also show ratios of variances to facilitate comparisons. Overall, predictions at cohort locations from on-road plume-adjusted mobile data are more variable than those derived from stationary measurements (Table N6S1, columns 2-4) with an overall variance ratio of 1.40. We hypothesized that the gains in variability due to an increase in spatial information would be

greatest in areas where there was inadequate sampling from stationary measurements, such as near the airport and at cohort locations further away from stationary locations. We found that at cohort locations less than 3 km from the airport, the variance ratio is much higher than the overall ratio (2.19 and 2.15, relative to 1.40), whereas the ratios for cohort locations 5 km or greater from the airport was comparable to the overall ratio (1.38 and 1.47, relative to 1.40). Similarly, the second and third quartiles of cohort locations categorized by distance from stationary locations had progressively higher variability than the first quartile and the overall variance ratio; however, the variance ratio in the fourth quartile was like the first and the overall variance ratio.

To anchor these observations, this table also shows 1) predictions and ratios in columns 5 and 6 based on time-sampled stationary data (so the duration of measurements per visit approximately matches the mobile measurement duration), 2) predictions and ratios in columns 7 and 8 based on location-sampled mobile data (so the number of locations in the mobile data is the same as the stationary data), and 3) the ratio of the location-sampled mobile data to the time-sampled stationary data. These comparisons allow deeper understanding of the greater uncertainty and lower precision due to shorter duration measurements and fewer spatial locations, respectively. Overall variability of predictions within stationary data (unsampled vs. time-sampled) and within mobile data (unsampled vs. location-sampled) are relatively similar. Thus, their ratios are fairly close to 1, in contrast to the ratio for the mobile to the stationary data. The patterns of the ratios by distance to large airport are fairly comparable when considering the two sampling approaches, and distinctly different from the mobile vs. stationary prediction variance ratio. Finally, the overall variance ratio for the location-sampled mobile data to the time-sampled stationary data is similar to but smaller than the mobile to stationary data comparison (1.30 versus 1.40). These additional analyses suggest that the mobile data, while

clearly more uncertain than the stationary data, are capturing additional information not available from the stationary data.

**Table N6S1**. Variance of predictions at cohort locations, based on stationary measurements and on-road plume-adjusted mobile measurements (multiplied by 1000)

| Category | Mobile[1] | Ratio: Mobile to stationary | Stationary | Time-sampled stationary[2] | Ratio: Time-sampled stationary to stationary | Location-sampled mobile[3] | Ratio: Location-sampled mobile to mobile | Ratio: Location-sampled mobile to time-sampled stationary |
|---|---|---|---|---|---|---|---|---|
| All | 6299 | 1.40 | 4493 | 4808 | 1.07 | 6231 | 0.99 | 1.30 |
| **Distance to large airport[4]** | | | | | | | | |
| 0-1km | 538 | 2.19 | 245 | 234 | 0.95 | 958 | 1.78 | 4.10 |
| 1-3km | 11431 | 2.15 | 5320 | 4270 | 0.80 | 9543 | 0.83 | 2.23 |
| 3-5km | 8532 | 1.94 | 4402 | 4185 | 0.95 | 7165 | 0.84 | 1.71 |
| 5-10km | 5359 | 1.38 | 3875 | 5504 | 1.42 | 5985 | 1.12 | 1.09 |
| >10km | 3989 | 1.47 | 2710 | 2875 | 1.06 | 4266 | 1.07 | 1.48 |
| **Distance to nearest stationary site[5]** | | | | | | | | |
| <373m | 5317 | 1.45 | 3670 | 4374 | 1.19 | 5463 | 1.03 | 1.25 |
| >373-603m | 4587 | 1.59 | 2887 | 3285 | 1.14 | 4313 | 0.94 | 1.31 |
| >603-962m | 5734 | 1.70 | 3368 | 3454 | 1.03 | 4969 | 0.87 | 1.44 |
| >962m | 7317 | 1.47 | 4992 | 5314 | 1.06 | 7470 | 1.02 | 1.41 |

[1]Refers to the PNC predictions from the on-road plume-adjusted mobile data. This is true for the 'sampled mobile' predictions as well.
[2]Predictions derived from a random sample of consecutive 10 seconds at each stationary location (rather than the full 120 seconds).
[3]A set of 309 mobile locations that are like the stationary locations based on geo-covariates.
[4]Distance to Seattle-Tacoma International Airport
[5]Distance from each cohort location to the nearest stationary site is stratified by quartiles: 1st quartile (<373 m); 2nd quartile (373-603m); 3rd quartile (603-962m); 4th quartile (>962m)

**Table S9**. Summary of literature comparing mobile and stationary measurements and/or predictions for UFP and/or BC

| Authors | Location and year | Pollutants | Mobile sampling | Stationary sampling | Key Findings |
|---|---|---|---|---|---|
| Kerckhoffs et al. 2016 (20) | Netherlands, winter/spring 2013 | UFP, BC | 2,964 road segments; segments average 130m in length; average 12s of UFP measurements collected per segment, and an average of 3 visits (once per season) | 161 short-term stationary sites; total of 30 min sampled at each | Mobile UFP predictions at residential locations 1.41x higher than stationary; for BC, this difference was 1.91x. (Figure 2) |
| Minet et al. 2018 (21) | Toronto, May-August 2016; 7am-7pm | UFP, BC | 3,895 road segments sampled on bicycle (over 270 km); visited at least 5 times each (mean visits: 6, averaging a total of 121s sampled per segment) | 92 sidewalk locations. 20 minutes of sampling per drive pass with an average of 102 minutes sampled at each location over the campaign.. | Pearson correlation coefficients between mobile and stationary surfaces were 0.23 for UFP and 0.49 for BC. UFP medians from mobile surface were 31,201 versus 19,057 pt/cm$^3$ from the stationary surface, about 1.63x greater. For BC, the medians were 1,799 and 1,469 ng/m$^3$ for mobile and stationary, respectively, about 1.22x larger. (Figure 1, 2) |
| Chambliss et al. 2020 (22) | Oakland, CA; summer 2017 | BC | 2 mobile platforms with >300 hours of total sampling time. 30m road segments (~21,000 total segments) visited an average of 31 times | 97 near-road and 3 background fixed sites. Total of 304 sampling hours | Spatial $R^2$ between mobile and fixed-site measurements is 0.51. They reported overlap between mobile and fixed-site medians for BC. (Figure 1) |
| Kerckhoffs et al. 2017 (23) | Netherlands, 2014-15 | UFP, BC | 5,236 road segments, average 110m long. Each segment had 25s of sampled UFP data total, and was visited about 3 times (once per season) | 240 short-term stationary sites, each sampled 3 times for 30 minutes each. | UFP predictions from mobile models were about 30% higher than those from short-term stationary (Figure 2) |

# Chapter 3: Ultrafine Particle Exposure and Physical Performance in the Adult Changes in Thought Cohort

**Abstract**

**Background:** Declining physical function is common among older adults and is associated with chronic diseases. Exposure to air pollution is associated with many conditions that lead to worsening physical function. Ultrafine particles (UFPs) and other traffic-related air pollutants (TRAPs) can enter the circulatory system and may lead to systemic inflammation. No studies have investigated the impact of UFP exposure on physical functioning.

**Objectives:** Estimate the relationship between UFPs and other TRAPs and longitudinal change in physical function.

**Methods:** We conducted a longitudinal analysis in the Adult Changes in Thought study, a cohort of older adults in the Puget Sound area enrolled over 1994-2020. We examined the association between UFPs and other TRAPs and physical performance assessed using a modified version of the Short Physical Performance Battery, a measure of physical function that combines three physical functioning assessments: chair-stand time, gait speed, and grip strength. We used linear mixed models to understand whether UFP or other TRAP (black carbon, nitrogen dioxide, fine particulate matter, UFP of several size bins) exposure was associated with physical performance, both cross-sectionally at baseline and longitudinally for a 5-year change in physical performance.

**Results:** The average physical performance score at baseline was 9.05, and the average within-person decline in physical performance score per year was 0.32 points. In models adjusted for baseline age, sex, race, education, and neighborhood deprivation index, higher UFP exposure was consistent with a wide range of rates of decline in physical performance, including no decline. An interquartile range increase in black carbon exposure resulted in a decline in physical performance equivalent to an additional 3 months of aging over 5 years.

**Discussion:** Our results suggest that exposure to UFPs is not associated with decline in physical functioning. Exposure to other pollutants, including black carbon, fine particulate matter, and certain UFP size bins may be associated with decline in physical functioning.

## Introduction

Toxicology studies suggest that ultrafine particles (UFPs), particles ≤ 100nm in diameter, may be more toxic than larger particles (7,25). Due to their small size, UFPs may enter the circulatory system through the lungs or the olfactory bulb (26). Through these mechanisms, UFPs are hypothesized to lead to systemic inflammation, thus contributing to the development of chronic diseases (6). Epidemiologic studies have reported associations between long-term UFP exposure and cardiovascular and inflammatory impacts, including associations with blood pressure, heart rate variability, and systemic inflammation (6). As individuals age, chronic conditions, including cardiovascular and respiratory diseases, often take a toll on physical functioning, an early indicator of later disability and other poor health outcomes (27,76). However, few studies have examined the impact of long-term exposure to air pollution on physical functioning (28,29,77), and none have examined the impact of UFP exposure.

Diminished physical function is common among older adults, and of growing concern due to the increasing population of older adults worldwide (34). Poor or worsening physical function may be detectable much earlier than the onset of other adverse outcomes, and thus may aid in earlier identification of individuals at high risk of hospitalization or mortality, for example (35). Physical function is measured by performance-based tests of lower and sometimes upper-extremity function and is an important measure of aging. The Short Physical Performance Battery (SPPB) is a common measure of physical function. The SPPB is a well-established test, has been used in several epidemiology studies, and reports a high degree of reliability and validity (76,78,79). We use a modified version of the SPPB to measure physical function.

Three studies report increases in traffic-related air pollutants (TRAPs) are associated with worsening physical function over time (28,29,77). A study in a cohort of older adults in the Netherlands assessed the longitudinal association between several air pollutants ($NO_2$, $NO_x$, $PM_{2.5}$, $PM_{10}$, $PM_{coarse}$) and a measure of physical function. They reported annual decreases in

physical function scores associated with an increase in $NO_2$ and $NO_X$, equivalent to an additional 9 months of aging for an interquartile range increase in $NO_2$ (29). A second study in a cohort of older adults in Chicago, US examined the longitudinal association between $NO_X$ and distance to major roadways and physical function. A 1-year decline in physical performance score among those in the highest quartile of NOx exposure was equivalent to over 7 additional months of aging compared to those in the lowest quartile of exposure (28). A third study in a cohort of middle-aged and older adults in China examined the longitudinal association between $PM_{2.5}$ and physical function. They reported a decrease in the physical function score equivalent to an additional nearly 2 months of aging for a 10 $\mu g/m^3$ increase in $PM_{2.5}$ (77).

This study is the first to examine the association between UFPs and changes in physical function, leveraging a well-characterized cohort of older adults in the Puget Sound region. These data will allow us to evaluate the impact of UFPs and other TRAPs on physical function, both cross-sectionally and longitudinally.

## Methods

### Study overview

The Adult Changes in Thought (ACT) Study is an on-going community-based cohort study examining aging and brain health in the Puget Sound area (80). Enrollment of cognitively intact older adults (≥65 years) began in 1994-1996, and the study maintains a cohort of 2000 participants per year. Participants undergo a thorough medical assessment every 2 years. In addition to extensive clinical and medical information collected from each visit, ACT collects a comprehensive set of data on potential confounders, including age, sex, race, ethnicity, education, family income, and socioeconomic status. ACT monitors residential address information which allowed us to estimate annual average residential UFP and other TRAP exposure.

Exposure assessment

We leveraged predictions from the 2019-2020 ACT-TRAP mobile monitoring campaign, described elsewhere (56). Briefly, we designed a mobile monitoring campaign to characterize several traffic-related air pollutants at ACT cohort locations. The campaign sampled at 309 short-term stationary roadside locations as well as on the road between roadside locations. We created predictions at ACT cohort locations from data collected at the short-term stationary locations using universal kriging with partial least squares regression (UK-PLS).

The mobile monitoring campaign used several instruments to measure particle number concentration (PNC) to capture UFPs, in addition to black carbon (BC), nitrogen dioxide ($NO_2$), fine particulate matter ($PM_{2.5}$), and carbon dioxide ($CO_2$) (56). Our primary exposure is PNC predictions obtained from an unscreened P-Trak, capturing particles with an aerodynamic diameter of 20-1000nm. We also examined PNC for 36-1000nm (screened P-Trak), 20-700nm (DiSCmini), 10-420nm (NanoScan), several PNC size bins from the NanoScan (10-13.3nm, 13.3-17.8nm, 17.8-23.7nm, 23.7-31.6nm, 31.6-42.2nm, 42.2-56.2nm, 56.2-75nm, 75-100nm, 100-133.4nm, 133.4-177.8nm, and 10-100nm), 20-36nm (difference between the unscreened and screened P-Trak), and PNC predictions from on-road locations using the unscreened P-Trak. The measurements from the on-road locations were adjusted for on-road plumes prior to creating predictions at ACT cohort locations. We additionally examined BC and $NO_2$ exposure from this campaign (16).

We created cumulative 5-year averages of air pollutant exposure based on the 2019 annual averages from the mobile monitoring campaign. Five-year averages were generated for each year of follow-up, and then averaged from baseline until the exam. Thus, within-person exposure only varied if the participant moved during the study. These exposure predictions from the mobile monitoring campaign are a proxy for long-term exposure (56). We assumed the annual average estimates of UFP from our mobile monitoring campaign represent long-term

exposure back to 1990 based on prior work that has shown limited temporal changes in the predicted spatial distribution over time (81–83).

In secondary analyses, we used annual average concentrations at participant addresses from 2-week average concentrations derived from hierarchical spatiotemporal prediction models for $PM_{2.5}$ (84) and for $NO_2$ (58). We leveraged a $PM_{2.5}$ spatiotemporal model, described in Shaffer et al. (84) and updated to extend to 2021, and a $NO_2$ spatiotemporal model, which was developed in the Puget Sound region following the approach described in Keller et al (58). Time-varying 5-year averages of the predictions were calculated following the same approach as the mobile monitoring data. To compare more directly with the exposure predictions derived from the mobile monitoring (MM) campaign, we additionally considered spatial model (SP) predictions derived from the spatiotemporal models (ST), leveraging measurements only from 2019. We did not use $PM_{2.5}$ from the mobile monitoring campaign as the measurements were deemed too uncertain (56).

Outcome assessment

ACT collects information about physical performance at each in-person biennial visit. The modified Performance-Based Physical Function (PPF-mod) score is made up of three assessments: chair-stand time (a measure of muscle strength), measured in seconds and repeated 5 times; grip strength, measured in kilograms in the dominant hand; and a 10-ft timed walk, measured in seconds (32). A measure of standing balance was previously recorded at biennial visits and was included in the original score (PPF), but was dropped from biennial visits; thus, we use the modified score composed of three assessments (PPF-mod). Each assessment is scored from 0-4, with a score of 0 indicating the participant did not complete the assessment, and scores of 1-4 assigned based on sex-specific quartiles for the study population. The total score ranges from 0 to 12, with higher scores indicating higher physical function (32). We modeled the outcome continuously, as done elsewhere (28,29,77). These three assessments

together measure both lower and upper-extremity strength. Many physical function assessments have been used and evaluated, and measurements with multiple assessments report higher validity and reliability compared to single assessments (78).

We included all ACT participants with at least 2 recorded PPF-mod scores from enrollment to March 5th, 2020. See Table S1 for participants excluded from inferential analyses.

Inferential analysis

In the primary analysis, we examined the association between long-term UFP exposure and changes in physical performance (PPF-mod) scores over time. We used linear mixed effect models to assess both the rate of change over time and the cross-sectional association at baseline. Linear mixed models account for repeated measures data, allow for individual-level random effects (intercept and slope), and accommodate the assumption that data are missing at random (MAR). Like the model used previously in the MESA Air study (85), our model for the PPF-mod score ($Y_{it}$) for the *i*th person at their *t*th follow-up visit ($v_{it}$) (measured as time since baseline) is:

$$Y_{it} = (\alpha_0 + a_i + X_{i0}\alpha_1 + Z_{i0}\alpha_2) + [\beta_0 + b_i + X_{it}\beta_1 + W_{it}\beta_2]v_{it} + e_{kit}\ ,$$

where *X* signifies exposure and *Z* and *W* are vectors of potential confounders or predictors of the outcome. The term $X_{i0}\alpha_1$ captures the cross-sectional UFP association at baseline, and $X_{it}v_{it}\beta_1$ characterizes the decline in physical performance scores associated with long-term UFP exposure. The rate of change parameter, $\beta_1$, is our primary parameter of interest, and the cross-sectional parameter, $\alpha_1$, is of secondary interest. For comparison to other studies and ease of interpretation, we report the rate of change parameter over a 5-year interval (of aging). The rate of change parameter can be interpreted as the difference in the 5-year PPF-mod score change per IQR increase in the pollutant (or from the lowest quartile when considering quartiles of exposure).

We adjusted for confounding in a series of staged models. All models include the air pollutant exposure of interest. In model 1 (M1), we adjusted for age at baseline, and the cross product of time since baseline (in years) with a categorical visit age variable, binned into 5-year age groups (65-69, 70-74, 75-79, 80-84, 85-89, ≥90). In model 2 (M2), we adjust for variables in M1 plus sex, race (White, non-White), education (high school degree or less, bachelors or more, other), and the cross product of time since baseline with all confounders. In our models, race is treated as a confounder, and is a proxy for structural racism that would otherwise be unmeasured. In model 3 (M3), our *a priori* model, we adjust for all variables in M2 plus a neighborhood disadvantage index (NDI), and the cross product of time since baseline with all confounders. NDI (range: -2.66 to 1.97) characterizes neighborhood socioeconomic status by using census variables included in the Multi-Ethnic Study of Atherosclerosis (MESA) Neighborhood Disadvantage Index, weighted by a principal components analysis (86). In model 4 (M4), our extended model, we adjust for all variables in M3 plus baseline smoking status (current, former, never), regular exercise at baseline (yes, no), alcohol use at baseline (current, former, never), and the cross product of time since baseline with time-varying regular exercise, time-varying alcohol use, and smoking status. We examined one additional model as a secondary analysis. In model 5 (M5), we adjusted for all variables in M3 plus several chronic disease indicators that may be on the causal pathway (i.e., potential mediators) between air pollution exposure and physical performance decline: hypertension, diabetes, heart disease, and respiratory disease.

In the spatio-temporal models with and without time-varying $NO_2$ or $PM_{2.5}$ exposure (ST and SP, respectively), we accounted for calendar time by adjusting for the calendar year of the visit in two-year bins. We included this in both the longitudinal and cross-sectional parts of the model. All other covariate adjustments were consistent with the models that used the mobile monitoring derived exposures.

We report results per interquartile range (IQR) increase in the pollutant of interest, where IQR is based on the distribution of the pollutant at baseline. We additionally report longitudinal decline results in the equivalent amount of aging for comparison across pollutants and studies by dividing the effect estimate by the time since entry estimate, which is the rate of decline due to aging alone for those ages 65-69. We also report the results for each quartile of exposure with the first quartile as the reference group. Exposure quartiles are defined based on exposure at baseline.

In sensitivity analyses, we re-ran all models with all ACT participants who enrolled on 1/1/2000 or later, due to concerns that the spatial structure of UFP may be different from 1990-1995 compared to after 1995. Restricting the cohort to all participants who enrolled in 2000 or later ensures all participants have 5-year exposure averages that begin in 1995 or later. We also assessed for bias due to selective attrition by implementing inverse probability of attrition weighting (IPAW), following methods of Weuve et al (87). This method upweights participants with a greater likelihood of loss to follow-up due to mortality or drop-out. Finally, we ran several additional sensitivity analyses: a model truncating our dataset at visit age 90 due to concerns about participants over age 90 being influential; a baseline linear model to compare with the cross-sectional effect estimated in our *a priori* longitudinal model for UFP; a model without five-year visit age bins in the longitudinal part of the model; a model with neighborhood deprivation index in the longitudinal part of the model; and a model dropping all visits with low air pollutant exposure coverage or with poor address history imputation (described in Table S2).

All analyses were conducted using R statistical software, version 4.2.2 (66). This study was approved by the University of Washington Institutional Review Board.

**Results**

There were 4,536 ACT cohort participants in our study (Table 1). The mean age at baseline was 74, 58% of participants were female, and over 90% were White. At the baseline

exam, 40% of participants had hypertension, 10% had diabetes, and 16% had heart disease. The mean PPF-mod score at baseline was 9.1, on a scale of 0-12, with a higher score corresponding to higher physical functioning. The distribution of participant demographics varied only slightly by quartile of baseline UFP exposure, though the PPF-mod score at baseline was slightly higher among those with lower UFP exposure compared to the highest quartile of exposure.

**Table 1**. Characteristics of ACT Participants at baseline by quartile of baseline UFP exposure

| **Characteristic** | **Quartile of baseline UFP exposure** | | | | **Overall** |
|---|---|---|---|---|---|
| | **Lowest** (2,505-6,264 pt/cm$^3$) (N=1,134) | **Second** (6,264-7,082 pt/cm$^3$) (N=1,134) | **Third** (7,082-8,021 pt/cm$^3$) (N=1,134) | **Highest** (8,021-15,884 pt/cm$^3$) (N=1,134) | (Mean PNC: 7,395 pt/cm$^3$ (SD: 1,752 pt/cm$^3$) (N=4,536) |
| Baseline Age (Mean (SD)) | 73.5 (5.8) | 73.7 (6.1) | 73.6 (6.2) | 74.0 (6.2) | 73.7 (6.1) |
| Female (%) | 663 (59%) | 658 (58%) | 656 (58%) | 670 (59%) | 2,647 (58%) |
| Race (%) | | | | | |
| American Indian or Alaskan Native | <10 | <10 | <10 | <10 | <10 |
| Asian | 24 (2%)[1] | 26 (2%) | 32 (3%) | 79 (7%) | 161 (4%) |
| Black | <10 | <10 | 46 (4%) | 97 (9%) | 158 (4%) |
| Native Hawaiian or Pacific Islander | <10 | <10 | <10 | <10 | <10 |
| Other, including mixed | 37 (3%) | 29 (3%) | 32 (3%) | 41 (4%) | 139 (3%) |
| White | 1,061 (94%) | 1,067 (94%) | 1,024 (90%) | 911 (80%) | 4,063 (90%) |
| Missing | <10 | <10 | <10 | <10 | <10 |
| Education | | | | | |
| GED/HS | 468 (41%) | 413 (36%) | 364 (32%) | 447 (39%) | 1,692 (37%) |
| Bachelor's | 261 (23%) | 249 (22%) | 302 (27%) | 243 (21%) | 1,055 (23%) |
| Masters | 164 (15%) | 195 (17%) | 211 (19%) | 154 (14%) | 724 (16%) |
| Doctorate | 44 (4%) | 90 (8%) | 86 (8%) | 59 (5%) | 279 (6%) |
| Other | 107 (9%) | 104 (9%) | 98 (9%) | 113 (10%) | 422 (9%) |
| Missing | 90 (8%) | 83 (7%) | 73 (6%) | 118 (10%) | 364 (8%) |
| Smoking Status | | | | | |
| Current Smoker | 55 (5%) | 45 (4%) | 48 (4%) | 66 (6%) | 214 (5%) |
| Never Smoker | 552 (49%) | 565 (50%) | 579 (51%) | 535 (47%) | 2,231 (49%) |
| Past Smoker | 527 (47%) | 521 (46%) | 503 (44%) | 530 (47%) | 2,081 (46%) |
| Missing | <10 | <10 | <10 | <10 | 10 (0%) |
| Exercise Regularly | 865 (71%) | 906 (75%) | 879 (72%) | 866 (71%) | 3,516 (72%) |
| NDI (Mean (SD)) | -0.93 (0.59) | -0.92 (0.59) | -0.77 (0.66) | -0.22 (0.77) | -0.71 (0.72) |
| Hypertension | 451 (40%) | 455 (40%) | 430 (38%) | 481 (42%) | 1,817 (40%) |
| Diabetes | 108 (10%) | 111 (10%) | 95 (8%) | 129 (11%) | 443 (10%) |
| Heart Disease | 177 (16%) | 172 (15%) | 173 (15%) | 192 (17%) | 714 (16%) |

| | | | | | |
|---|---|---|---|---|---|
| Respiratory disease | 173 (15%) | 185 (16%) | 182 (16%) | 188 (17%) | 728 (16%) |
| PPF-mod Score (Mean (SD) | 9.2 (2.2) | 9.1 (2.3) | 9.1 (2.2) | 8.9 (2.3) | 9.1 (2.2) |

[1]For categorical variables, percentages do not always add up to 100 due to rounding.

ACT participants had an average of 5.4 study visits, an average of 4.5 visits with a PPF-mod score recorded, and an average of 9 years of follow-up (Table S3). UFP (primary instrument) exposure at baseline ranged from 2,505 to 15,884 pt/cm$^3$, with a median of 7,075 pt/cm$^3$. There were also differences in the variation in the PNC instruments: the unscreened P-Trak (20-1000nm) had a standard deviation (SD) of 1,754 pt/cm$^3$, while the screened P-Trak (36-1000nm) had an SD of 623 pt/cm$^3$. Further, the distribution and variation of $NO_2$ and $PM_{2.5}$ varied by collection method and time period. The models with time-varying exposure had larger SDs compared to models with only 2019 data (Table 2).

In our *a priori* model for UFP, we observed results consistent with a wide range of rates of decline in physical performance, including no decline in the 5-year change in PPF-mod score for an IQR increase in UFP exposure and comparing quartiles of exposure (Table 3). We observed the same trend for all other models (M1-5) for UFP exposure.

Cross-sectionally, when we compared quartiles of UFP exposure, increasingly higher quartiles of UFP exposure generally had a larger negative impact on physical performance score at baseline compared to the lowest quartile, although the confidence intervals suggest the results are consistent with a wide range of effects including no decline (Table 4).

**Table 2.** Distribution of pollutant exposures at baseline

| Pollutant | Min | 25th% | Median | 75th% | Max | SD |
|---|---|---|---|---|---|---|
| **PNC (pt/cm$^3$)** | | | | | | |
| Primary instrument, Unscreened P-Trak: 20-1000nm | 2,505 | 6,264 | 7,075 | 8,018 | 15,884 | 1,754 |
| Screened P-Trak: 36-1000nm | 1,417 | 2,980 | 3,226 | 3,500 | 6,907 | 623 |
| P-Trak difference: 10-36nm | 1,797 | 3,212 | 3,691 | 4,592 | 11,832 | 1,207 |
| DiSCmini: 20-700nm | 2,161 | 7,761 | 8,979 | 10,169 | 31,339 | 2,940 |
| NanoScan: 10-420nm | 4,576 | 9,401 | 10,232 | 11,282 | 20,850 | 1,962 |
| 10-13.3 nm | 191 | 395 | 502 | 622 | 1,955 | 220 |
| 13.3-17.8 nm | 396 | 754 | 938 | 1,143 | 3,823 | 389 |
| 17.8-23.7 nm | 480 | 776 | 914 | 1,078 | 2,989 | 300 |
| 23.7-31.6 nm | 821 | 1,229 | 1,380 | 1,563 | 3,294 | 347 |
| 31.6-42.2 nm | 899 | 1,430 | 1,553 | 1,700 | 3,614 | 334 |
| 42.2-56.2 nm | 774 | 1,308 | 1,404 | 1,519 | 3,422 | 271 |
| 56.2-75 nm | 624 | 1,018 | 1,148 | 1,268 | 2,866 | 243 |
| 75-100 nm | 555 | 825 | 943 | 1,061 | 2,057 | 182 |
| 100-133.4 nm | 377 | 627 | 703 | 765 | 1,257 | 109 |
| 133.4-177.8 nm | 203 | 297 | 323 | 351 | 541 | 43 |
| 10-100 nm | 5,147 | 8,304 | 9,006 | 9,928 | 19,983 | 1,892 |
| Mobile data (Unscreened P-Trak): 20-1000nm | 3,522 | 6,371 | 7,493 | 8,818 | 22,201 | 2,200 |
| **Other pollutants** | | | | | | |
| BC (ppb) | 253 | 507 | 568 | 632 | 1,118 | 106 |
| $NO_2$: MM[1] (ppb) | 4.08 | 8.60 | 9.46 | 10.69 | 20.45 | 2.15 |
| $NO_2$: SP[2] (ppb) | 5.77 | 8.83 | 9.59 | 10.53 | 22.25 | 1.98 |
| $NO_2$: ST[3] (ppb) | 5.42 | 12.18 | 13.78 | 15.44 | 37.80 | 3.77 |
| $PM_{2.5}$: SP (µg/m$^3$) | 4.20 | 5.40 | 5.58 | 5.80 | 7.14 | 0.33 |
| $PM_{2.5}$: ST (µg/m$^3$) | 4.40 | 7.91 | 10.18 | 12.50 | 17.75 | 2.67 |

[1]MM: Predictions derived from the mobile monitoring campaign.

[2]SP: Predictions derived from the spatio-temporal model using data from 2019.

[3]ST: Predictions derived from the spatio-temporal model using data across all years.

We additionally examined the effect per IQR increase and by quartiles of exposure for BC, $NO_2$, and $PM_{2.5}$ on PPF-mod score. Generally, in the primary model (M3) exposure to higher levels of BC and spatially modeled (SP) $PM_{2.5}$ resulted in faster decline in PPF-mod scores compared to lower levels of exposure. Cross-sectionally, higher levels of BC and $NO_2$ were associated with lower PPF-mod scores.

For BC, the 5-year decline among those with an exposure one IQR (124 ppb) higher was 0.052 points greater (95% CI: -0.098, -0.006) (Table 3), equal to an additional 3 months of aging over 5 years for 65–69-year-olds. Similarly, we observed a decrease in PPF-mod score at baseline among those with higher versus lower BC exposure (Table 4). For $NO_2$, our results are consistent with a wide range of rates of decline, including no decline in PPF-mod score associated with exposure to higher levels (Table 3). Cross-sectionally for all $NO_2$ models (MM, SP, and ST), we observed a decrease in PPF-mod score at baseline among those with higher versus lower $NO_2$ exposures (Table 4). For the $PM_{2.5}$ spatial model (SP), the 5-year decline in PPF-mod score among those with an exposure one IQR (0.41 μg/$m^3$) higher was 0.08 (95% CI: -0.13, -0.03) points greater, equivalent to an additional 3 months of aging. We observed no cross-sectional effect at baseline for $PM_{2.5}$ modeled spatially, and no effect longitudinally or cross-sectionally for time-varying $PM_{2.5}$.

**Table 3.** Differences (95% CI) in 5-year change in PPF-mod score per IQR increase or by quartile of air pollutant exposure

| **Exposure measure and model** | **Per IQR** | **Second quartile[1]** | **Third quartile** | **Highest quartile** |
|---|---|---|---|---|
| **UFP (20-1000nm)** | (1750 pt/$cm^3$) | (6264-7082 pt/$cm^3$) | (7082-8021 pt/$cm^3$) | (8021-15884 pt/$cm^3$) |
| M1 | -0.002 (-0.043, 0.039) | -0.054 (-0.148, 0.039) | -0.057 (-0.156, 0.042) | 0.022 (-0.082, 0.125) |
| M2 | -0.007 (-0.048, 0.035) | -0.062 (-0.156, 0.031) | -0.073 (-0.173, 0.026) | 0.001 (-0.103, 0.106) |
| M3, *a priori* | -0.01 (-0.055, 0.034) | -0.064 (-0.159, 0.031) | -0.065 (-0.166, 0.037) | -0.003 (-0.113, 0.107) |
| M4 | -0.009 (-0.056, 0.037) | -0.075 (-0.174, 0.024) | -0.069 (-0.176, 0.038) | 0.005 (-0.110, 0.120) |
| M5 | -0.024 (-0.073, 0.026) | **-0.118 (-0.223, -0.013)[2]** | -0.083 (-0.196, 0.030) | -0.046 (-0.168, 0.077) |

| | | | | |
|---|---|---|---|---|
| **BC** | (124 ppb) | (507-568 ppb) | (569-632 ppb) | (633-1118 ppb) |
| M1 | **-0.048** | -0.024 | -0.046 | **-0.155** |
| | **(-0.094, -0.002)** | (-0.115, 0.066) | (-0.143, 0.052) | **(-0.258, -0.051)** |
| M2 | **-0.052** | -0.028 | -0.051 | **-0.162** |
| | **(-0.098, -0.006)** | (-0.118, 0.063) | (-0.149, 0.046) | **(-0.266, -0.058)** |
| M3 | **-0.062** | -0.033 | -0.064 | **-0.187** |
| | **(-0.11, -0.014)** | (-0.125, 0.059) | (-0.163, 0.036) | **(-0.294, -0.080)** |
| M4 | -0.05 | -0.053 | -0.070 | **-0.166** |
| | (-0.101, 0.000) | (-0.15, 0.044) | (-0.175, 0.035) | **(-0.279, -0.054)** |
| **$NO_2$** | (2.1 ppb)[3] | (8.6-9.4 ppb) | (9.5-10.6 ppb) | (10.7-20.4 ppb) |
| MM, M3 | -0.039 | -0.016 | -0.031 | -0.067 |
| | (-0.079, 0.001) | (-0.107, 0.076) | (-0.131, 0.069) | (-0.172, 0.039) |
| SP, M3 | -0.014 | -0.033 | 0.060 | 0.041 |
| | (-0.057, 0.030) | (-0.125, 0.060) | (-0.038, 0.158) | (-0.064, 0.145) |
| ST, M3 | -0.025 | -0.393 | -0.162 | -0.294 |
| | (-0.051, 0.000) | (-1.494, 0.708) | (-2.152, 1.828) | (-1.41, 0.822) |
| **$PM_{2.5}$** | (0.41 µg/$m^3$)[4] | (5.40-5.58 µg/$m^3$) | (5.59-5.80 µg/$m^3$) | (5.81-7.14 µg/$m^3$) |
| SP, M3 | **-0.080** | **-0.117** | **-0.135** | **-0.192** |
| | **(-0.130, -0.030)** | **(-0.209, -0.024)** | **(-0.233, -0.037)** | **(-0.295, -0.089)** |
| ST, M3 | -0.013 | -0.061 | -0.043 | -0.037 |
| | (-0.036, 0.010) | (-0.14, 0.018) | (-0.146, 0.060) | (-0.192, 0.117) |

[1]The first quartile is the reference group for the analysis comparing quartiles of exposure, and thus not shown.

[2]Results in bold are significant at the p <0.05 level

[3]IQR and quartile exposure ranges are reported for the mobile monitoring data. All $NO_2$ models use 2.1 ppb for the IQR increase, but each model uses their own baseline defined quartiles for the comparison across quartiles of exposure.

[4]IQR and quartile exposure ranges are reported for the spatial $PM_{2.5}$ model. All $PM_{2.5}$ models use 0.41 µg/$m^3$ for the IQR increase, but each model uses their own baseline defined quartiles for the comparison across quartiles of exposure.

**Table 4.** Baseline change (95% CI) in PPF-mod score per IQR increase or by quartile of air pollutant exposure

| Exposure measure and model | Per IQR | Second quartile[1] | Third quartile | Highest quartile |
|---|---|---|---|---|
| **UFP (20-1000nm)** | (1750 pt/$cm^3$) | (6264-7075 pt/$cm^3$) | (7076-8018 pt/$cm^3$) | (8019-15884 pt/$cm^3$) |
| M1 | **-0.090** | 0.037 | -0.102 | **-0.233** |
| | **(-0.148, -0.033)[2]** | (-0.124, 0.198) | (-0.264, 0.060) | **(-0.396, -0.071)** |
| M2 | **-0.076** | 0.013 | -0.136 | **-0.187** |
| | **(-0.133, -0.018)** | (-0.146, 0.172) | (-0.296, 0.024) | **(-0.349, -0.025)** |
| M3, *a priori* | -0.044 | 0.025 | -0.118 | -0.113 |
| | (-0.108, 0.021) | (-0.137, 0.186) | (-0.281, 0.046) | (-0.289, 0.062) |
| M4 | -0.049 | 0.010 | **-0.165** | -0.149 |
| | (-0.113, 0.015) | (-0.15, 0.17) | **(-0.327, -0.002)** | (-0.324, 0.026) |
| M5 | -0.049 | 0.064 | -0.128 | -0.119 |
| | (-0.115, 0.017) | (-0.099, 0.227) | (-0.294, 0.038) | (-0.297, 0.060) |
| **BC** | (124 ppb) | (507-568 ppb) | (569-632 ppb) | (633-1118 ppb) |
| M1 | **-0.070** | **-0.190** | -0.066 | **-0.183** |
| | **(-0.138, -0.002)** | **(-0.351, -0.029)** | (-0.228, 0.096) | **(-0.346, -0.020)** |
| M2 | **-0.074** | **-0.227** | -0.070 | **-0.186** |

| | | | | |
|---|---|---|---|---|
| | **(-0.141, -0.007)** | **(-0.385, -0.068)** | (-0.229, 0.090) | **(-0.347, -0.026)** |
| M3 | -0.048 | **-0.210** | -0.044 | -0.122 |
| | (-0.119, 0.024) | **(-0.372, -0.048)** | (-0.207, 0.120) | (-0.290, 0.046) |
| M4 | -0.069 | **-0.212** | -0.055 | -0.155 |
| | (-0.139, 0.002) | **(-0.372, -0.052)** | (-0.217, 0.108) | (-0.322, 0.012) |
| **$NO_2$** | (2.1 ppb)[3] | (8.6-9.4 ppb) | (9.5-10.6 ppb) | (10.7-20.4 ppb) |
| MM, M3 | **-0.077** | **-0.195** | -0.128 | **-0.273** |
| | **(-0.137, -0.017)** | **(-0.357, -0.034)** | (-0.292, 0.037) | **(-0.442, -0.105)** |
| SP, M3 | **-0.077** | -0.126 | **-0.165** | **-0.182** |
| | **(-0.143, -0.012)** | (-0.288, 0.036) | **(-0.328, -0.002)** | **(-0.351, -0.013)** |
| ST, M3 | **-0.064** | **-0.196** | **-0.241** | **-0.383** |
| | **(-0.100, -0.028)** | **(-0.364, -0.029)** | **(-0.415, -0.067)** | **(-0.563, -0.204)** |
| **$PM_{2.5}$** | (0.41 µg/m$^3$)[4] | (5.40-5.58 µg/m$^3$) | (5.59-5.80 µg/m$^3$) | (5.81-7.14 µg/m$^3$) |
| SP, M3 | 0.071 | 0.052 | 0.018 | 0.129 |
| | (-0.003, 0.145) | (-0.109, 0.214) | (-0.144, 0.181) | (-0.036, 0.294) |
| ST, M3 | 0.017 | -0.117 | -0.079 | -0.016 |
| | (-0.012, 0.045) | (-0.371, 0.138) | (-0.457, 0.298) | (-0.407, 0.374) |

[1]The first quartile is the reference group for the analysis comparing quartiles of exposure, and thus not shown.
[2]Results in bold are significant at the p <0.05 level
[3]IQR and quartile exposure ranges are reported for the mobile monitoring data. All $NO_2$ models use 2.1 ppb for the IQR increase, but each model uses their own baseline defined quartiles for the comparison across quartiles of exposure.
[4]IQR and quartile exposure ranges are reported for the spatial $PM_{2.5}$ model. All $PM_{2.5}$ models use 0.41 µg/m$^3$ for the IQR increase, but each model uses their own baseline defined quartiles for the comparison across quartiles of exposure.

In secondary analyses, we ran the primary model (M3) for the additional PNC measures captured in the mobile monitoring campaign (Table 2, Figures S1 and S2). Compared to the primary UFP longitudinal and cross-sectional estimates from the unscreened P-Trak (20-1000nm), we observed similar longitudinal estimates for the DiSCmini (20-700nm), the NanoScan (10-420nm), and the mobile data from the unscreened P-Trak. However, the longitudinal decline estimated for the screened P-Trak (36-1000nm) was over 4 times larger, though still very small, and with confidence intervals that included 0 (Figure S1). Exposure to one IQR higher levels of PNC in the range of 56-178nm (across 4 size bins) was associated with faster decline in the PPF-mod score (Figure S1). This was the equivalent of aging an additional 3-5 months over 5 years for 65–69-year-olds compared to those with an IQR lower exposure. The cross-sectional estimates at baseline were generally similar to the estimates from the primary UFP model (Figure S2).

In sensitivity analyses, in a model with IPAW to assess if our results may be biased due to differential attrition (87), we observed only slight changes in the estimates (Tables S4 and S5). In an analysis with the dataset truncated at visit age 90 to exclude all visits where the individual was 90 or older (Tables S4 and S5), the longitudinal decline estimate was over 3 times larger, though still small and with confidence intervals that included 0 (Table S4). In a linear regression model with only data at baseline to check our cross-sectional results, the estimates were in the same direction with wider confidence intervals (Table S5). Additionally, in an analysis with all ACT participants who enrolled on 1/1/2000 or later, as a check on the exposure surface, we found the same trends with wider confidence intervals due to the fewer participants included in the models (Tables S4 and S5). In a model without five-year age bins in the longitudinal part of the model (keeping a binary indicator for visit age equal to or greater than 75), we observed the longitudinal decline parameter was slightly greater than 0, and with a very similar confidence interval to the primary model. In an analysis with the neighborhood deprivation index in the longitudinal part of the model, we found the longitudinal decline parameter was over twice as negative, though still consistent with a wide range of effects, including no decline. Finally, in a model excluding visits with poor air pollution exposure coverage or poor address imputation, we observed larger longitudinal decline, but still consistent with the *a priori* estimate. There was almost no impact on the baseline estimate.

**Discussion**

We conducted the first study examining exposure to UFPs and physical functioning. We found evidence of a decrease in physical functioning due to higher levels of UFPs cross-sectionally at the baseline exam, and no association between UFP exposure and average within-person change in physical functioning over time. Of the other three pollutants we examined, exposure to higher levels of BC and $PM_{2.5}$ modeled spatially resulted in faster declines in physical performance scores compared to lower levels of exposure. We generally

found higher levels of UFP, BC, and $NO_2$ were associated with lower physical performance scores cross-sectionally at baseline; however, these effects were small.

Three previous studies examined traffic-related air pollution and physical performance (28,29). Weuve et al. examined the association between $NO_x$ and distance to roadway with a measure of physical performance (28). They used a 5-year average of $NO_x$ exposure prior to baseline and a slightly different measure of physical performance composed of three elements, two of which are the same as our measure (chair stand time and gait speed), and a third that was different (tandem stand, measuring balance). They found that participants with the highest quartile of NOx exposure had a decline in physical performance score that was equivalent to 3 additional years of aging compared to those in the lowest quartile over 5 years. While not a direct comparison, for $NO_2$, we reported a lower physical performance score due to an IQR increase in exposure to $NO_2$ at baseline, equivalent to a baseline age 6 months older, but no evidence of a decline in physical performance over time. De Zwart et al. also examined several TRAPs including $NO_2$, $NO_x$, $PM_{2.5}$, $PM_{10}$, and $PM_{Coarse}$ (29). Similarly, they did not use time-varying exposure, and used a slightly different outcome measure, composed of gait speed, chair stand time, balance, and putting on/off a sweater. They reported declines in physical performance score with an IQR increase in $NO_2$ and $NO_x$, equivalent to aging an additional 9 months, but no association for PM. We find no evidence of decline in physical performance over time with increased $NO_2$, and evidence of decline with increased $PM_{2.5}$ using the spatial model. Finally, Wang et al. reported a 0.19% decrease (0.03, 0.36) in the physical performance score for each 10 $\mu g/m^3$ in $PM_{2.5}$. The difference in results across studies may be due to several factors; all studies have outcomes composed of different sets of assessments which measure slightly different aspects of physical functioning. Thus, these assessments may be capturing somewhat different impacts due to air pollution exposure. Further, each study reports effect estimates for a different increase in the pollutant and for different outcome metrics, and thus the

effect sizes are not directly comparable as reported. Additionally, each study reports estimates based on different cohorts, with different demographics and risk factors. We report estimates as an equivalent amount of aging for ease of comparison across studies and pollutants.

Our results indicate differences in the effect on PPF-mod score by pollutant. Exposure to higher levels of BC showed the clearest evidence of a faster decline in PPF-mod score among the pollutants we examined, equivalent to an additional 3 months of aging over 5 years. Exposure to higher levels of other pollutants was consistent with a range of effects, except for the spatial $PM_{2.5}$ exposure and PNC between 56-178nm, which were associated with faster decline. For the spatial $PM_{2.5}$ exposure, a one IQR increase resulted in a 5-year decline in the PPF-mod score equivalent to 3 additional months of aging. For the PNC size bins between 56-178nm, a one IQR increase in exposure resulted in a 5-year decline in the PPF-mod score equivalent to 3-5 additional months of aging.

Our results examining PNC size bins indicated a longitudinal effect of UFPs in the range of 56-178nm on physical functioning, but no effect for smaller particles or wider size ranges. We did not examine the effect of larger particles alone; the NanoScan had larger size bins, but we did not use them as their counts were too low. Prior toxicological evidence and epidemiologic research indicates that smaller particles may be more toxic (6,7,25). However, we found that PNC in the range of 56-178nm, but not PNC < 56nm, appear to have the greatest impact on decline in physical function. Further, other researchers have reported that PNC larger than 50nm (88), or in the range of 60-80nm (89), corresponds to combustion from diesel engines, which is also a major source of BC. We see longitudinal decline effects due to BC exposure and the four PNC size bins ranging between 56 and 178nm, suggesting that emissions from diesel engines may be driving some of this effect. However, there is limited epidemiologic research examining the impact of different particle size ranges on health end points and none for physical

functioning. Additional research is needed to fully understand the impact of particle size on physical functioning and aging-related outcomes.

We controlled for confounding by robust inclusion of potential confounders. In our *a priori* (M3) model, we controlled for neighborhood deprivation index (NDI) in the cross-sectional part of the model, which is an indicator of neighborhood socioeconomic status (SES). The difference in longitudinal effect estimates between the model with (M3) and without (M2) adjustment for NDI suggests confounding by neighborhood SES (90,91). Further, in a sensitivity analysis, we controlled for NDI in the longitudinal part of the model and observed a larger decline in PPF-mod score, though it was still consistent with a range of effects including no decline (Table S4). While the overall conclusion does not change, this result suggests it may be important to account for confounding due to neighborhood SES over time, not just at baseline. Prior research indicates that even with these adjustments, residual confounding likely remains due to the complexity of the air pollution-socioeconomic status relationship, both at the individual and neighborhood level (91). It is worth noting that these models with (M3) and without (M2) adjustment for NDI ask different scientific questions, and we learn something different about the effect of air pollution exposure from each.

Most of our exposure metrics varied spatially but not temporally; thus, we were not overly concerned about confounding due to calendar time. In the models with exposure that varied over space and time (ST models), we adjusted for calendar time in two-year bins. However, in a longitudinal setting with exposure time-averaged since baseline and with substantial decline in levels of $PM_{2.5}$ and $NO_2$ over the study period, this adjustment may not be adequate. For $PM_{2.5}$, the spatial model indicated higher exposure to $PM_{2.5}$ (based on 2019 data alone) was associated with a decline in PPF-mod score; however, the spatiotemporal model, based on time-varying data and controlling for calendar time, indicated higher levels of $PM_{2.5}$

was consistent with a range of effects, including no decline. Thus, there may be residual confounding due to calendar time present in this model.

Based on other studies examining traffic-related air pollution effects on measures associated with aging, we expected the effect of air pollution to be small, if detectable at all (28,29,77,84). Thus, it is critical to adequately capture the acceleration of decline due to aging in the longitudinal part of the model. Before considering the effect of exposure, we examined several ways of modeling the decline due to aging and found creating five-year visit age categories allowed for acceleration in the slope as participants aged in the study. Furthermore, in studies of older adults, there is concern about differential attrition; in other words, those who drop out are different from those who remain in the study. Participants who remain in the study at older ages may be healthier than the full cohort. The ACT cohort has a Completeness of Follow-up Index (CFI) (92) of 95.6%, indicating that nearly all planned visits have data available. This lowered our concerns about selection bias due to differential attrition. However, as an additional check, we ran a sensitivity analysis to further assess for differential attrition. We conducted inverse probability of attrition weighting (87), and found slight impact on our results, though they were still consistent with the primary model results. This indicated minor to minimal impact of differential attrition in our study. In another sensitivity analysis, we truncated the visit age at age 90 to further reduce the influence of survivor bias. The longitudinal effect estimate became more negative, indicating older ages may be driving some attenuation of the longitudinal association.

In additional sensitivity analyses, we excluded visits with poor exposure quality or poor address history. This impacted the results minimally, and the longitudinal effect was still consistent with positive and negative associations. Further, we examined how dropping the five-year visit age bins from the longitudinal part of the model impacted our results. The direction of the longitudinal effect estimate changed, though the confidence interval overlapped almost

completely with the primary model CI. The change in direction may indicate that the inclusion of five-year visit age bins better captures the overall aging effect, allowing us to pick up the subtle air pollution effect.

Our study has several strengths. To our knowledge, ours is the first study to examine the association between UFP and BC with physical performance. We utilized the ACT cohort, a well-characterized community-based cohort of older adults in the Puget Sound region with a rich set of covariates and high rates of follow-up (80). We also leveraged robust exposure metrics, including our recent and highly innovative mobile monitoring campaign (56), which includes several PNC instruments and size bins, and spatiotemporal models for $PM_{2.5}$ (84) and $NO_2$ (58) developed for the region. Further, the PPF-mod score is an objective and validated measure of physical functioning. To increase the strength of our and others' findings, additional research could examine other measures of physical decline. One such measure is frailty, a similar and correlated measure, though it adds additional metrics including weight loss, fatigue, and physical activity (93,94), which may shed light on additional aspects of physical decline not captured in the physical performance measure.

Our study also has limitations. Our primary exposure metrics were derived from our 2019 mobile monitoring campaign. We used predictions derived from these measurements as a proxy for historical exposure. We assumed the spatial distribution of UFPs is constant back in time (56,81–83). We were less confident that the predictions from the mobile campaign are representative of the spatial distribution of pollutants prior to 1995 due to changes in policies and emissions in the late 1980s and early 1990s (95). Thus, we conducted a sensitivity analysis excluding ACT cohort participants who enrolled prior to 2000 (who would not have a 5-year exposure average in the relevant time period). We find similar trends with wider confidence intervals, due to substantially less person-time contributing to our estimates, suggesting that our spatial surface is adequate for our analysis in all time periods. Furthermore, the exposure

estimates are based on participant residential address. Thus, we assumed participants spend most of their time at home, and that outdoor exposure is a proxy for true participant exposure, a common assumption in air pollution epidemiology studies. Additionally, the ACT cohort is about 89% White; thus, further disaggregation of race beyond the non-White and White groups in a regression model is not possible due to sample size constraints. Further studies leveraging cohorts with greater racial and ethnic diversity are needed to understand the true population impact of traffic-related air pollution exposure on physical functioning. Additionally, we treated the physical performance score as a continuous measure in our models, as other studies have done (28,29,77). However, the physical performance score is a discrete outcome on a scale of 0-12 and modeling it continuously does not acknowledge that participants who start with lower physical performance scores at baseline have less to decline compared to those who start with higher scores at baseline. Finally, though physical performances measures are highly predictive of poor health outcomes, they are not important outcomes on their own in a clinical setting. Rather, physical performance scores serve as an intermediate measure for outcomes that are important clinically.

Overall, in this prospective cohort study with rich covariate data and novel high quality exposure estimates, we observed higher levels of several traffic-related air pollutants resulted in faster declines in physical functioning, though the effects were small. Further research is needed to understand which pollutant(s) have the most substantial impact, and to understand the impact in more racially and ethnically diverse populations.

## Supplemental Material

**Table S1**. ACT Cohort exclusions for inferential analyses

| **Exclusions** | **Participants dropped (N)[1]** |
|---|---|
| Participants with 1 visit | 755 |
| Participants with <2 PPF-mod scores | 408 |
| Participants with no UFP exposure[2] | 64 |
| **Final data set** | 4,536 |

[1]Full cohort before exclusions is 5,763.
[2]No UFP exposure as these participants lived outside the mobile monitoring region.

**Table S2.** Exposure and address history coverage

| Measure | UFP (all instruments) | BC | $NO_2$ (MM) | $NO_2$ (ST) | $PM_{2.5}$ (ST) |
|---|---|---|---|---|---|
| Average weeks of coverage (proportion)[1] | 0.99 | 0.98 | 0.98 | 0.98 | 0.99 |
| Average exact coverage (proportion)[2] | 0.98 | 0.96 | 0.96 | 0.96 | 0.97 |
| Average imputed coverage (proportion)[3] | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 |

[1]Average weeks of coverage: Average proportion of weeks in the averaging period covered with an exposure prediction
[2]Average exact coverage: Proportion of the averaging time covered by an exact geocode
[3]Average imputed coverage: Proportion of averaging time covered by an imputed address

**Table S3.** ACT cohort person-level follow-up information for the analysis dataset (N=4,536)

| | ACT participants with 1+ follow-up (Mean (SD)) |
|---|---|
| Visits (N) | 5.4 (2.8) |
| Visits with PPF-mod score (N) | 4.5 (2.6) |
| Number of follow-up years | 9 (5.7) |

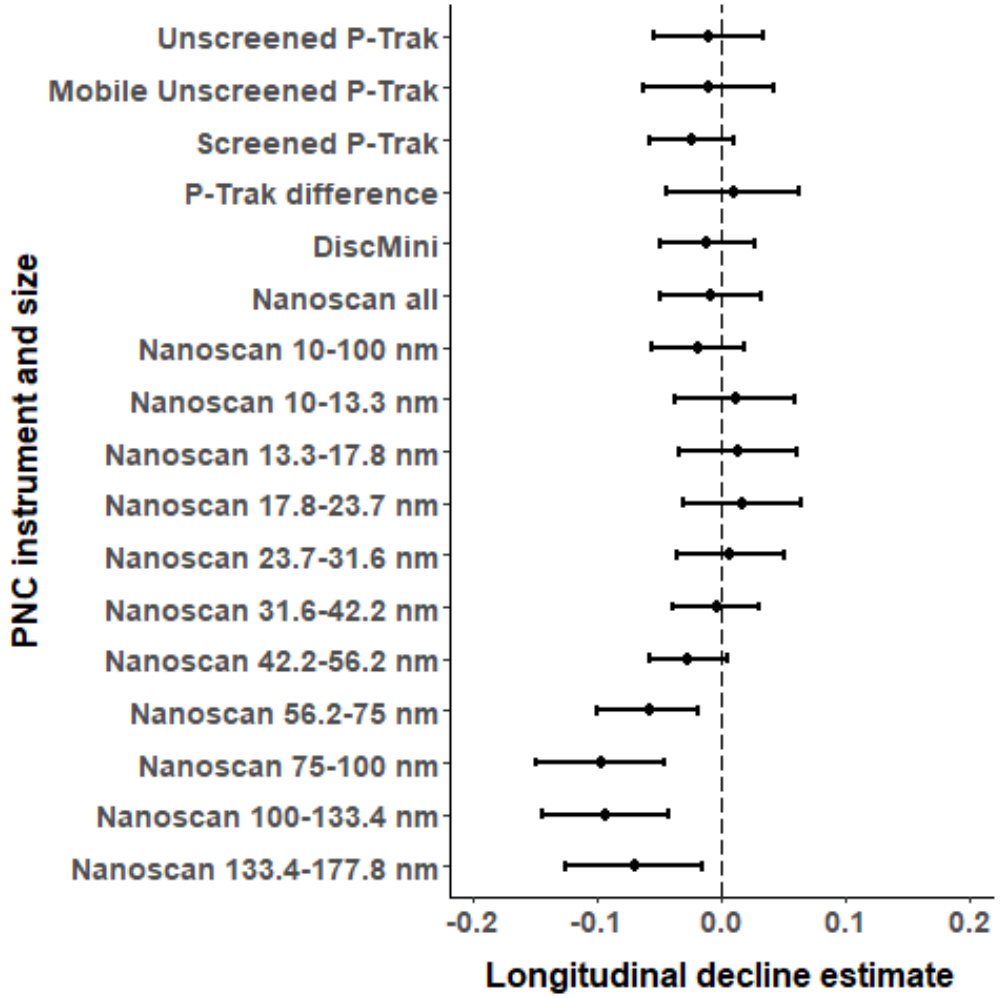


**Figure S1.** Differences and 95% CI in 5-year change in PPF-mod score per IQR increase in exposure for the *a priori* model (M3) for additional PNC instruments and size bins

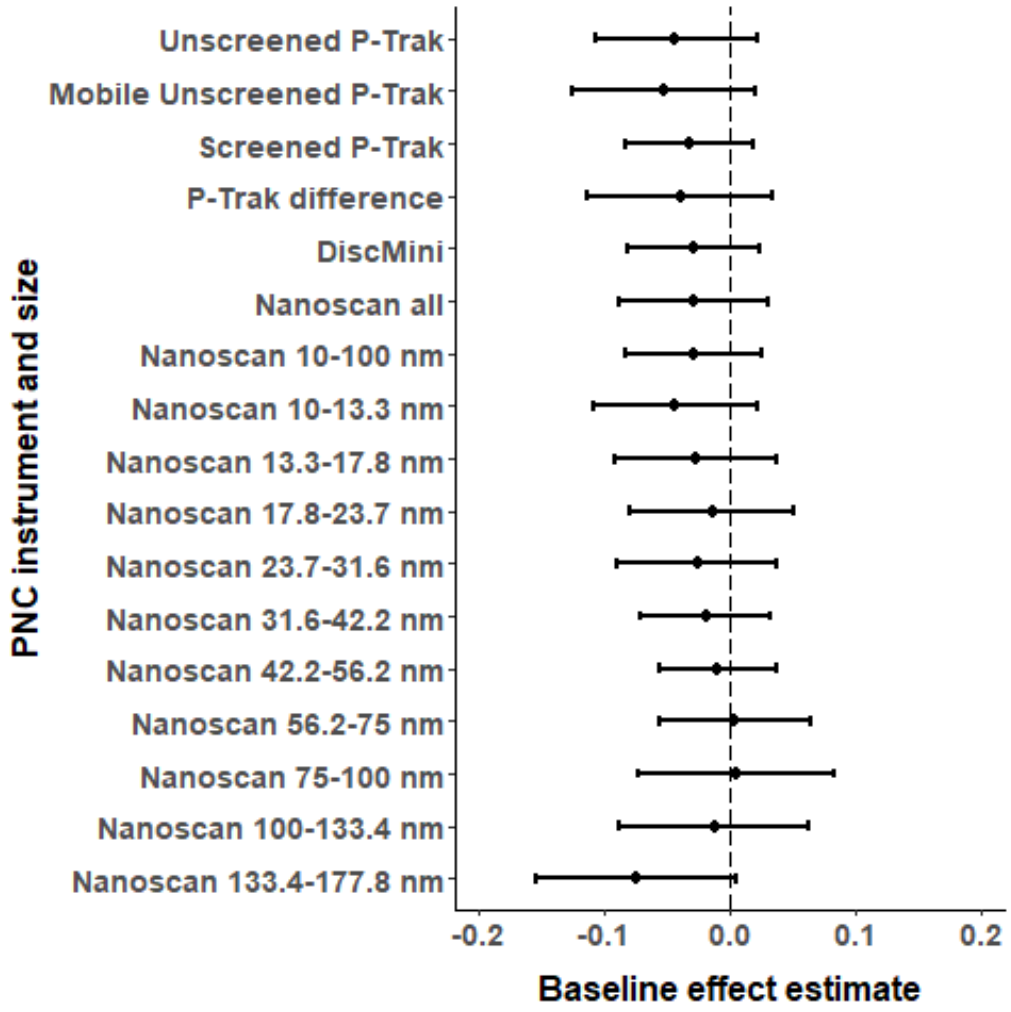


**Figure S2.** Baseline change and 95% CI in PPF-mod score per IQR increase in exposure for the *a priori* model (M3) for additional PNC measures and size bins

**Table S4.** Differences (95% CI) in 5-year change in PPF-mod score per IQR increase in exposure for the *a priori* model for UFP 20-1000nm for sensitivity analyses

| Sensitivity Analysis | Per IQR | 95% CI |
|---|---|---|
| Primary UFP model | -0.010 | -0.055, 0.034 |
| IPAW weights | 0.006 | -0.048, 0.059 |
| Truncated at age 90 | -0.018 | -0.065, 0.030 |
| Excluding visits with poor exposure quality or address history | -0.015 | -0.062, 0.033 |
| Cohort entry on 1/1/2000 or later | 0.001 | -0.068, 0.070 |
| Dropping 5-year age bin in longitudinal part of model[1] | -0.005 | -0.049, 0.039 |

[1]Model contains cross-product of a binary indicator of visit age 75 or greater with time since entry

**Table S5.** Baseline change (95% CI) in PPF-mod score per IQR increase in exposure for the *a priori* model for UFP 20-1000nm for sensitivity analyses

| Sensitivity Analysis | Per IQR | 95% CI |
|---|---|---|
| Primary UFP model | -0.044 | -0.109, 0.021 |
| IPAW weights | -0.050 | -0.117, 0.018 |
| Truncated at age 90 | -0.048 | -0.112, 0.017 |
| Excluding visits with poor exposure quality or address history | -0.044 | -0.111, 0.022 |
| Cohort entry on 1/1/2000 or later | -0.050 | -0.142, 0.042 |
| Dropping 5-year age bin in longitudinal part of model[1] | -0.049 | -0.114, 0.016 |
| Baseline linear model | -0.033 | -0.100, 0.034 |

[1]Model contains cross-product of a binary indicator of visit age 75 or greater with time since entry

# Chapter 4: Wildfire smoke exposure and Emergency Department Visits in Washington State[2]

[2]Reprinted from Environmental Research: Health. Annie Doubleday, Lianne Sheppard, Elena Austin, Tania Busch Isaksen. Wildfire smoke exposure and emergency department visits in Washington state. Copyright (2023), with permission from IOP Science under a CC BY 4.0 license. The publication is available at https://iopscience.iop.org/article/10.1088/2752-5309/acd3a1.

**Abstract**

Wildfires are increasing in prevalence in western North America due to changing climate conditions. A growing number of studies examine the impact of wildfire smoke on morbidity; however, few evaluate these impacts using syndromic surveillance data that cover many emergency departments. We used syndromic surveillance data to explore the overall and cumulative effect of wildfire smoke exposure on all-cause respiratory and cardiovascular emergency department visits in Washington state. Using a time-stratified case crossover design, we observed an increased odds of asthma visits immediately after and in all five days following initial exposure (lag 0 OR: 1.13; 95% CI: 1.10, 1.17; lag 1-5 ORs all 1.05 or greater with a lower CI of 1.02 or higher), and an increased odds of respiratory visits in all five days following initial exposure (lag 1 OR: 1.02; 95% CI: 1.00, 1.03; lag 2-5 ORs and lower CIs were all at least as large) comparing wildfire smoke to non-wildfire smoke days. We observed mixed results for cardiovascular visits, with evidence of increased odds emerging only several days following initial exposure. We also found increased odds across all visit categories for a 10 $\mu g/m^3$ increase in smoke-impacted $PM_{2.5}$. In stratified analyses, we observed elevated odds for respiratory visits among ages 19-64, for asthma visits among ages 5-64, and mixed risk estimates for cardiovascular visits by age group. This study provides evidence of an increased risk of respiratory ED visits immediately following initial wildfire smoke exposure, and increased risk of cardiovascular ED visits several days following initial exposure. These increased risks are seen particularly among children and younger to middle-aged adults.

**Background**

Wildfires are increasing in prevalence in western North America due to changing climate conditions, including prolonged droughts and heat waves (96); changing land-use resulting in an increase in the wildland-urban interface (97); and decades of poor forest management (98,99). Accordingly, smoke from wildfires continues to impact large populations over prolonged time periods in western North America, resulting in worsening air quality over the past decade, reversing improvements in air quality overall (100–102). Recent toxicology and epidemiology studies suggest increased toxicity comparing wildfire smoke-influenced $PM_{2.5}$ to $PM_{2.5}$ without wildfire smoke (5,103,104). Thus, characterizing the health impacts of wildfire smoke is critical in a region's efforts to mitigate adverse health outcomes from smoke exposure.

Studies examining the impact of wildfire smoke exposure on morbidity are increasingly abundant (36–38,41,42,45,46,48,105–107), with several hospitalization and emergency department (ED) visit studies. These studies consistently estimate increased risk of respiratory hospitalizations and ED visits associated with wildfire smoke exposure (36–38,41,42,45,105,106). However, the results for cardiovascular hospitalizations and ED visits are inconsistent across studies, including a decreased risk, no change in risk, and an increase in risk (36–38,42,45,46,105,106).

To our knowledge, there are only a few studies utilizing syndromic surveillance data in a wildfire smoke context (42,48,107,108). Syndromic surveillance is a near real-time monitoring system that requires all Emergency Departments in Washington to report healthcare encounter visit data to help rapidly identify outbreaks and other public health emergencies (49,109). Utilization of syndromic surveillance data increases our power to detect an effect due to the high frequency of ED visits compared to more severe outcomes such as hospitalization or mortality. Further, to date, there are only two statewide wildfire smoke epidemiologic studies in Washington, and only one examining non-mortality outcomes (37,50). Gan et al. examined hospitalization outcomes during summer 2012 and found an increased risk of respiratory

hospitalizations associated with wildfire smoke, as characterized by three different exposure metrics. However, due to lack of comprehensive data sources and difficulty accessing health data, few studies have examined the effect of wildfire smoke exposure across several years of morbidity data in Washington. We examined both the overall impact and the intensity of wildfire smoke exposure on emergency department visits in Washington during the summers of 2017-2020 using a wide-ranging dataset of emergency departments covering all of Washington. Our results will further elucidate our understanding of the overall impact of smoke presence in addition to how the magnitude of smoke exposure impacts ED visits. Our results will help inform Washington state and local agency efforts to craft more targeted risk communication and interventions.

## Methods

### Study location and time period

The study period spanned June-September, 2017-2020, and included all reporting emergency departments in Washington state. We chose June-September as all major wildfire smoke events across this time period occurred during these months.

### Exposure assessment

We used two primary exposure metrics: 1) a binary indicator of wildfire smoke-impacted days, and 2) wildfire smoke impacted $PM_{2.5}$ concentration at the zip code level. The binary indicator of wildfire smoke gives results that are easier to communicate to the public and to practitioners, and the smoke concentration increases our understanding of how the intensity of exposure impacts health.

To create the binary metric, we assigned wildfire smoke days based on a combination of modeled and monitored $PM_{2.5}$ (50). Briefly, 4x4km grid cells from the *Air Indicator Report for*

*Public Awareness and Community Tracking (AIRPACT-4)* model domain in Washington were assigned the daily average monitored concentration from the nearest regulatory air quality monitor, based on agreement between the summertime average at the grid cell (based on the AIRPACT modeled concentration, scaled by the monitored concentration) and the summertime average at the nearest monitor. See Note S1 for details. We then computed the daily population-weighted average zip code-level $PM_{2.5}$ concentration to match the spatial resolution of the ED visit data.

We identified wildfire smoke days through an iterative process, described elsewhere (50). Briefly, we tried a threshold of 15 $\mu g/m^3$, corresponding to the 99$^{th}$ percentile of measured $PM_{2.5}$ concentrations across two years with minimal smoke impacts. However, most days above this threshold in urban areas were not wildfire smoke impacted, while some days below this threshold in rural areas were wildfire smoke impacted. We thus applied a more nuanced approach to better capture wildfire smoke days in both urban and rural areas. For the binary analysis, smoke days are defined as:

1. All zip code-days greater than 20.4 $\mu g/m^3$.
2. Zip code-days with a 24-hour average $PM_{2.5}$ concentration between 9 and 20.4 $\mu g/m^3$ must meet the following set of criteria to be considered a smoke day (50):
    1. The day must be part of an event in which at least 2 of 3 consecutive days are greater than 9 $\mu g/m^3$;
    2. One of the days in the 3-day window must be greater than 15 $\mu g/m^3$; and
    3. For urban areas (Seattle, Tacoma, Spokane), at least 50% of the air monitors in those areas must be greater than 9 $\mu g/m^3$.

We chose the 20.4 $\mu g/m^3$ based on the cut point between Moderate and Unhealthy for Sensitive Groups in the previously used Washington Air Quality Advisory. The lower cut point of 9 $\mu g/m^3$ was used after sensitivity analyses that found wildfire smoke impacts at low levels, corresponding to 9 $\mu g/m^3$.

We used this exposure grid to calculate wildfire smoke $PM_{2.5}$ concentration for each zip code-day. Wildfire smoke $PM_{2.5}$ was defined as the total $PM_{2.5}$ on smoke-impacted days (using the definition above) minus typical $PM_{2.5}$, where typical $PM_{2.5}$ was defined as the median $PM_{2.5}$ concentration for each zip code and month on non-smoke impacted days. All non-smoke days were assigned a smoke $PM_{2.5}$ concentration of 0 $\mu g/m^3$. By subtracting out typical $PM_{2.5}$ levels by season (month) and location (zip code), this focuses our analysis on smoke-impacted $PM_{2.5}$ using an approach similar to that used in other wildfire smoke-health studies (110,111). We used this wildfire smoke-impacted $PM_{2.5}$ for our primary analysis examining the impact of wildfire smoke intensity on ED visits.

In a sensitivity analysis, we re-ran our primary analysis using NOAA's daily Hazard Mapping System (HMS) smoke plumes to determine the binary exposure metric (112). Briefly, trained NOAA analysts created smoke polygons over North America based on satellite imagery from GOES-16 and GOES-17. The polygons were designated as Light, Medium, or Heavy smoke. We used the polygons categorized as 'Heavy' smoke, as these are more likely to be at ground level, reducing misclassification. If a smoke plume intersected with the zip code on a specific day, we categorized that zip code as smoke impacted(46). We considered using smoke plumes categorized as 'Medium', but found they introduce many false positives.

Outcome assessment

We used selected emergency department (ED) visits from Washington state's syndromic surveillance system (RHINO) for our study period. The system contains patient demographic information as well as data about the visit and clinical information pertinent to the visit. We examined the following groups of diagnoses as our outcomes of interest, using all listed diagnosis codes for each visit: all non-traumatic (ICD-10-CM codes: A01-R99), all-cause respiratory (ICD-10-CM codes: J01-J99), asthma (ICD-10-CM codes: J45), all-cause

cardiovascular (including cerebrovascular outcomes) (ICD-10-CM codes: I05-I52, I60-63, I65-69, G45), and myocardial infarction (MI) (ICD-CM-10: I21). We additionally queried the chief complaint text field to ensure we captured all ED visits in each outcome category (described in Table S1). We leveraged queries developed by the Washington state Department of Health (DOH) Rapid Health Information Network (RHINO) (49) team to maintain consistency and utilize queries that have been previously developed and tested.

Inferential analysis

We leveraged a time-stratified case-crossover design to evaluate the effect of wildfire smoke presence versus absence and wildfire smoke intensity on ED visits. This design uses conditional logistic regression to compare exposures within individuals during a time-stratified referent window. The exposure on the case or index date, the day of the ED visit or a prior day for lagged exposures, is compared to the exposure on referent days, defined as the same day of the week and month of the index day. We report the odds ratio (OR) on wildfire smoke days compared to non-wildfire smoke days using the daily patient zip-code level binary smoke indicator, described above. We report the odds ratio separately for lag days 0-5 for respiratory outcomes, or lag days 0-7 for cardiovascular outcomes. We chose *a priori* to extend the lag period to 7 days for the cardiovascular outcomes based on evidence suggesting an increase in cardiovascular-related mortality several days following initial exposure (107,111). Lag 0 represents initial exposure on the day of the ED visit, lag 1 represents initial exposure on the day prior to the visit, and so on. We adjusted for time- and zip code-specific humidex at lag 0 (113), a measure of temperature and humidity, modeled as a natural cubic spline with 3 degrees of freedom, as practiced elsewhere (114,115) (see Equation S1 for the model notation).

The syndromic surveillance system was actively onboarding emergency departments in Washington during the study period. Thus, we ran the binary conditional logistic regression

model separately for each year and combined the results using a fixed effect meta-analysis to yield an overall odds ratio for the study period (116). This method upweights years with smaller standard errors (more data). While we do not expect year-specific results to be homogenous, Rice et al. show that the overall effect estimated from a fixed effect meta-analysis is unbiased if the individual study effects are unbiased (116). We report results using a mixed effects meta-analysis as a sensitivity analysis.

In a secondary analysis, we reported results for all-cause respiratory, asthma, all-cause cardiovascular, and myocardial infarction visits stratified by age group (0-4, 5-18, 19-44, 45-64, 65-84, ≥85), using the same model form as the binary analysis.

To understand how the intensity of wildfire smoke impacts ED visits, we again used the time-stratified case-crossover design with conditional logistic regression. We used the continuous wildfire smoke $PM_{2.5}$ exposure metric over the pre-defined lag period using a distributed lag model. A distributed lag model uses an assumed lag structure model to distribute the effect of the exposure across the entire lag period rather than on one lag day, as done in a single lag model (117). The resulting effect estimates describe the lag-specific and cumulative impact of a 10 $\mu g/m^3$ increase in wildfire smoke $PM_{2.5}$ across the lag period, taking into account any smoke exposure during the lag period for the cumulative OR (108,111,114,118). Estimates from a distributed lag model cannot be directly compared to results from a single exposure day model. We used the `dlnm` R package (117), and modeled the exposure-outcome relationship linearly, with the lag modeled as a natural cubic spline with 3 degrees of freedom, as documented in similar studies (108,111).

We additionally sought to visualize the wildfire smoke $PM_{2.5}$-health effect relationship within the context of our existing public health messaging. We used the same model structure as in the intensity analysis, using a total $PM_{2.5}$. We modeled total $PM_{2.5}$ using a natural cubic spline with one knot at 55 $\mu g/m^3$, corresponding to the US Environmental Protection Agency's air quality index (AQI) cut point between 'Unhealthy for Sensitive Groups' and 'Unhealthy' for

$PM_{2.5}$. We estimated the dose response curve for a 10 μg/$m^3$ in total $PM_{2.5}$ for respiratory visits on lag day 1 plotting the curve over the AQI categories for visualization.

We ran several sensitivity analyses. First, we evaluated the binary model with and without COVID-19 diagnoses at time of treatment (Table S1). This only impacts data from 2020. We then re-ran the binary model excluding all cases from 2017, and then separately by year due to differences in coverage and patterns among reported cases by year (Table 1). We also re-evaluated the binary model results using a mixed effect meta-analysis rather than a fixed effect meta-analysis to allow the effects to differ by year. Next, we re-ran the binary model assigning smoke presence versus absence using NOAA's HMS smoke plumes, as described above. Finally, we ran all models using a quasi-Poisson time series analysis, adjusting for humidex, day of the week, month, year, and a year-specific log zip code-level population offset. Unlike the conditional logistic regression models used in our primary analyses, this model does not make assumptions about patient readmissions during the referent period.

**Results**

The Washington syndromic surveillance platform began collecting ED visit data in 2011 and was still onboarding emergency departments during the study period, resulting in fewer visits in 2017 compared to 2019 and 2020 (Table 1). All-cause respiratory visits make up 19% of all visits recorded in the dataset over the four summers, while all-cause cardiovascular visits make up nearly 29% of visits.

**Table 1.** Summary of Emergency Department Visit characteristics

| Characteristic (N (%)) | All years | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|
| Total study visits (all non-traumatic)[1] | 1,864,470 | 235,314 | 424,833 | 629,702 | 574,621 |
| EDs ever reporting | 94 (100) | 39 (41) | 66 (70) | 83 (88) | 94 (100) |

| | | | | | |
|---|---|---|---|---|---|
| Average study visits per day | 3,821 | 1,929 | 3,482 | 5,161 | 4,710 |
| Age group | | | | | |
| 0-4 | 82,040 (4) | 13,389 (6) | 21,028 (5) | 30,728 (5) | 16,895 (3) |
| 5-18 | 139,878 (8) | 20,559 (9) | 32,700 (8) | 49,459 (8) | 37,160 (7) |
| 19-44 | 692,257 (37) | 86,855 (37) | 155,699 (37) | 228,832 (36) | 220,871 (38) |
| 45-64 | 483,231 (26) | 61,352 (26) | 110,194 (26) | 160,654 (26) | 151,031 (26) |
| 65-84 | 371,043 (20) | 42,211 (18) | 83,644 (20) | 125,789 (20) | 119,399 (21) |
| ≥85 | 94,196 (5) | 10,892 (5) | 21,449 (5) | 32,673 (5) | 29,182 (5) |
| Missing | 1,825 (<1) | 56 (<1) | 119 (<1) | 1,567 (<1) | 83 (<1) |
| Race | | | | | |
| American Indian/ Alaska Native | 37,356 (2) | 4,263 (2) | 8,705 (2) | 11,700 (2) | 12,688 (2) |
| Asian | 61,084 (3) | 4,304 (2) | 12,854 (3) | 23,731 (4) | 20,195 (4) |
| Black | 138,226 (7) | 16,023 (7) | 30,145 (7) | 49,321 (8) | 42,737 (7) |
| Native Hawaiian/ Other Pacific Islander | 21,456 (1) | 637 (<1) | 3,593 (1) | 9,328 (2) | 7,898 (1) |
| White | 1,363,223 (73) | 172,581 (73) | 323,787 (76) | 451,353 (72) | 415,502 (72) |
| Other | 142,720 (8) | 3,570 (2) | 24,266 (6) | 60,276 (10) | 54,608 (10) |
| Not reported | 100,405 (5) | 33,936 (14) | 21,483 (5) | 23,993 (4) | 20,993 (4) |
| Ethnicity | | | | | |
| Hispanic | 182,277 (10) | 21,635 (9) | 38,544 (9) | 63,102 (10) | 58,996 (10) |
| Not Hispanic | 1,587,564 (85) | 176,907 (75) | 367,681 (87) | 545,404 (87) | 497,572 (87) |
| Not reported | 94,629 (5) | 36,772 (16) | 18,608 (4) | 21,196 (3) | 18,053 (3) |
| Sex | | | | | |
| Female | 1,016,769 (55) | 126,723 (54) | 232,670 (55) | 345,510 (55) | 311,866 (54) |
| Male | 847,539 (45) | 108,584 (46) | 192,145 (45) | 284,138 (45) | 262,672 (46) |
| Outcome | | | | | |
| All Non-traumatic[1] | 1,864,470 | 235,314 | 424,833 | 629,702 | 574,621 |
| All Respiratory | 353,717 (19) | 38,137 (16) | 73,777 (17) | 123,592 (20) | 118,211 (21) |
| Asthma | 68,293 (4) | 4,981 (2) | 10,246 (2) | 24,915 (4) | 28,151 (5) |
| All Cardio | 537,007 (29) | 46,565 (20) | 99,144 (23) | 195,298 (31) | 196,000 (34) |
| Myocardial infarction | 131,668 (7) | 13,645 (6) | 29,987 (7) | 45,679 (7) | 42,357 (7) |

Note: Percentages are rounded to the nearest whole number and may not add to 100%.
[1]All study visits are non-traumatic. All other visit categories are not mutually exclusive. A visit can be counted in multiple outcome categories, as all listed diagnosis codes are used to identify the outcome.

Table 2 shows average total $PM_{2.5}$ concentration and humidex at the zip code level averaged across the study region by year, and by wildfire smoke day. The average concentration on non-smoke days is consistent across the study period, however the total $PM_{2.5}$

concentration on smoke days varies by year. The 2020 wildfire smoke season had nearly 3 times higher average smoke day $PM_{2.5}$ concentration than the next highest year, 2018. The 2019 wildfire smoke season experienced the lowest average smoke day $PM_{2.5}$ concentration of the four years, and the fewest number of smoke days by zip code. Additionally, most (97%) smoke days were a part of a multi-day event, defined as at least two consecutive smoke days by zip code. The distribution of $PM_{2.5}$ on smoke versus non-smoke days, a summary of exposure at the referent window level (each individual set of referent and index days), and a comparison of the primary exposure model and the HMS smoke plume exposure model are summarized in Figure S1 and Tables S2 and S3.

**Table 2**. Summary of $PM_{2.5}$ and humidex across the study period and by year

| | All years | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|
| **$PM_{2.5}$ (μg/m$^3$), Mean (SD)** | | | | | |
| Overall[1] | 9.6 (3.6) | 10.4 (4.4) | 10.4 (5.2) | 4.2 (1.1) | 13.6 (4.4) |
| Non-smoke-days[2] | 4.3 (1.1) | 4.5 (1.3) | 4.4 (1.0) | 4.0 (1.0) | 4.1 (0.9) |
| Smoke days | 48.7 (9.8) | 33.4 (8.8) | 37.3 (9.2) | 17.7 (3.9) | 94.7 (32.0) |
| Number of smoke days | 14.4 (12.1) | 22.4 (11.2) | 21.2 (12.4) | 1.7 (3.5) | 12.6 (4.3) |
| Index days[3] | 9.4 (25.2) | 10.1 (18.6) | 9.7 (16.9) | 4.3 (2.7) | 14.4 (40.5) |
| Referent days | 9.3 (24.8) | 10.3 (18.5) | 9.6 (16.7) | 4.3 (2.7) | 14.2 (39.8) |
| **Humidex (degrees C), Mean (SD)** | | | | | |
| Overall humidex | 25.5 (2.3) | 26.0 (2.4) | 25.5 (2.2) | 25.5 (2.1) | 25.2 (2.4) |
| Non-smoke days | 25.1 (2.2) | 25.1 (2.2) | 24.7 (2.0) | 25.4 (2.0) | 25.2 (2.5) |
| Smoke days | 28.7 (2.3) | 30.8 (2.4) | 28.8 (2.6) | 33.0 (2.4) | 25.0 (1.8) |
| Index days | 25.5 (5.5) | 25.8 (6.4) | 25.5 (5.6) | 25.6 (5.1) | 25.3 (5.3) |
| Referent days | 25.5 (5.5) | 25.8 (6.6) | 25.4 (5.7) | 25.6 (5.1) | 25.3 (5.4) |

[1]All metrics are averages of zip code-level averages except summaries on index and referent days.
[2]Smoke and non-smoke days defined by primary binary exposure definition at the patient's zip code-day level.
[3]Summaries of index and referent days summarize $PM_{2.5}$ and humidex across the full case-crossover dataset which includes all referent windows.

In our presence versus absence analysis, we examined the impact of initial wildfire smoke exposure on ED visits using single-day binary exposure models (Figure 1; Table S4). We observed both protective and null effects for all non-traumatic ED visits comparing smoke days to non-smoke days (Figure 1). For all-cause respiratory ED visits, we observed no increase on lag 0, and an increase in odds across lags 1-5, with peak odds on lag 2 (OR: 1.03; 95% CI:

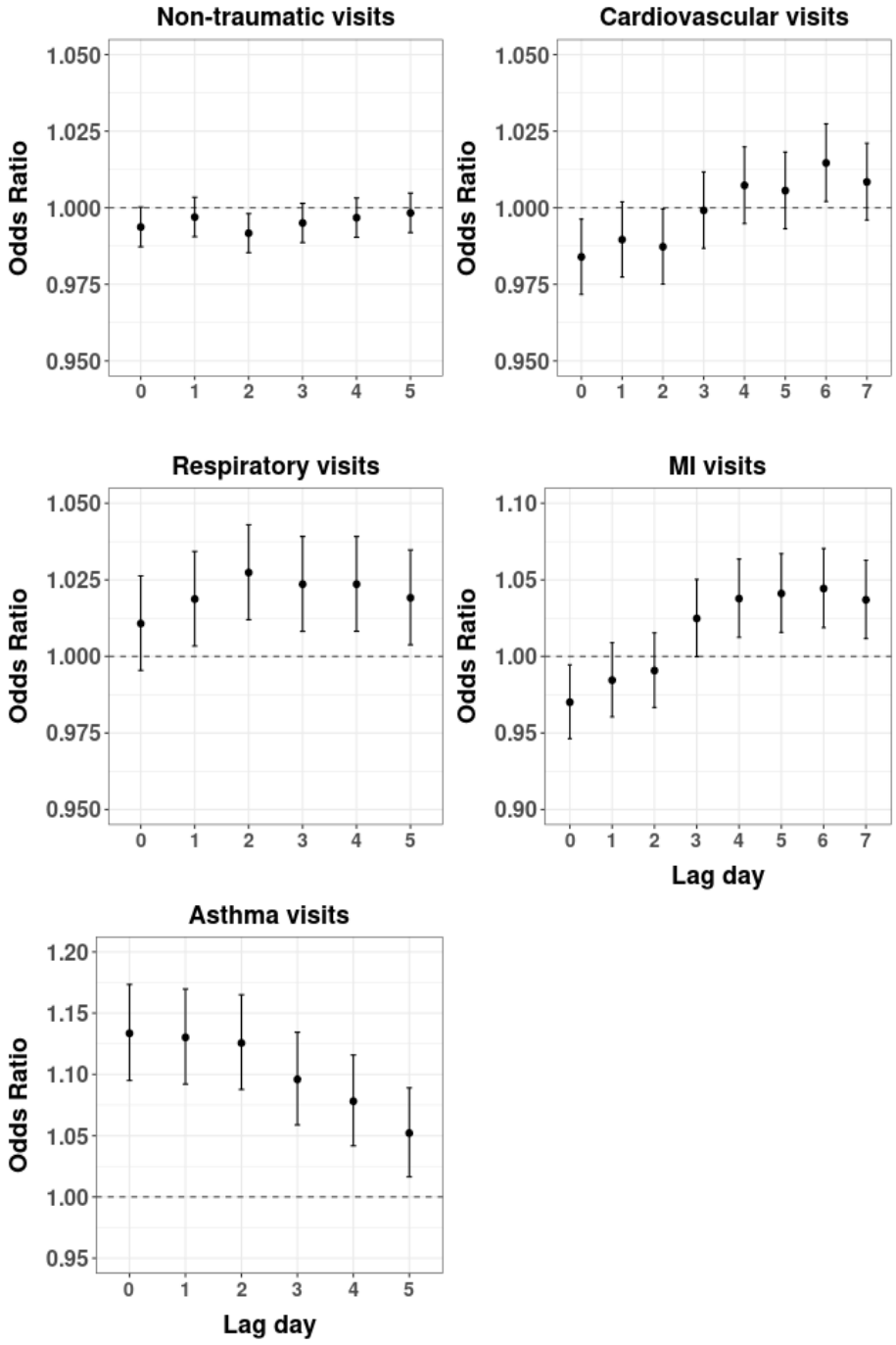


**Figure 1.** Lag-specific odds ratios of ED visits by outcome comparing wildfire smoke versus non-wildfire smoke days, controlling for humidex. Note y-axes vary by outcome to allow for visualization of results. All-cause respiratory includes asthma visits and all-cause cardiovascular includes MI visits. Visit categories are not mutually exclusive.

1.01, 1.04). For asthma-related ED visits, there was an increase in the odds of visits on smoke versus non-smoke days across all lags, with peak odds on the initial day of exposure (OR: 1.13, 95% CI: 1.10, 1.17). For all-cause cardiovascular visits, the single-day ORs were inconsistent across the lag period, including a decrease (lag 0 and lag 2), no increase (lag 1, 3-5, 7), and an increase (lag 6, OR: 1.02, 95% CI: 1.00, 1.03) in the odds of visits on smoke days compared to non-smoke days (Table S4). Similarly, for MI visits, we observed a range of effects consistent with a decrease (lag 0), no increase (lag 1-3), and an increase (lag 4-7) in the

odds, with peak odds on lag 6 (OR: 1.04, 95% CI: 1.02, 1.07). Results from a mixed effect meta-analysis show similar trends with slightly muted effects and wider confidence intervals (Table S5).

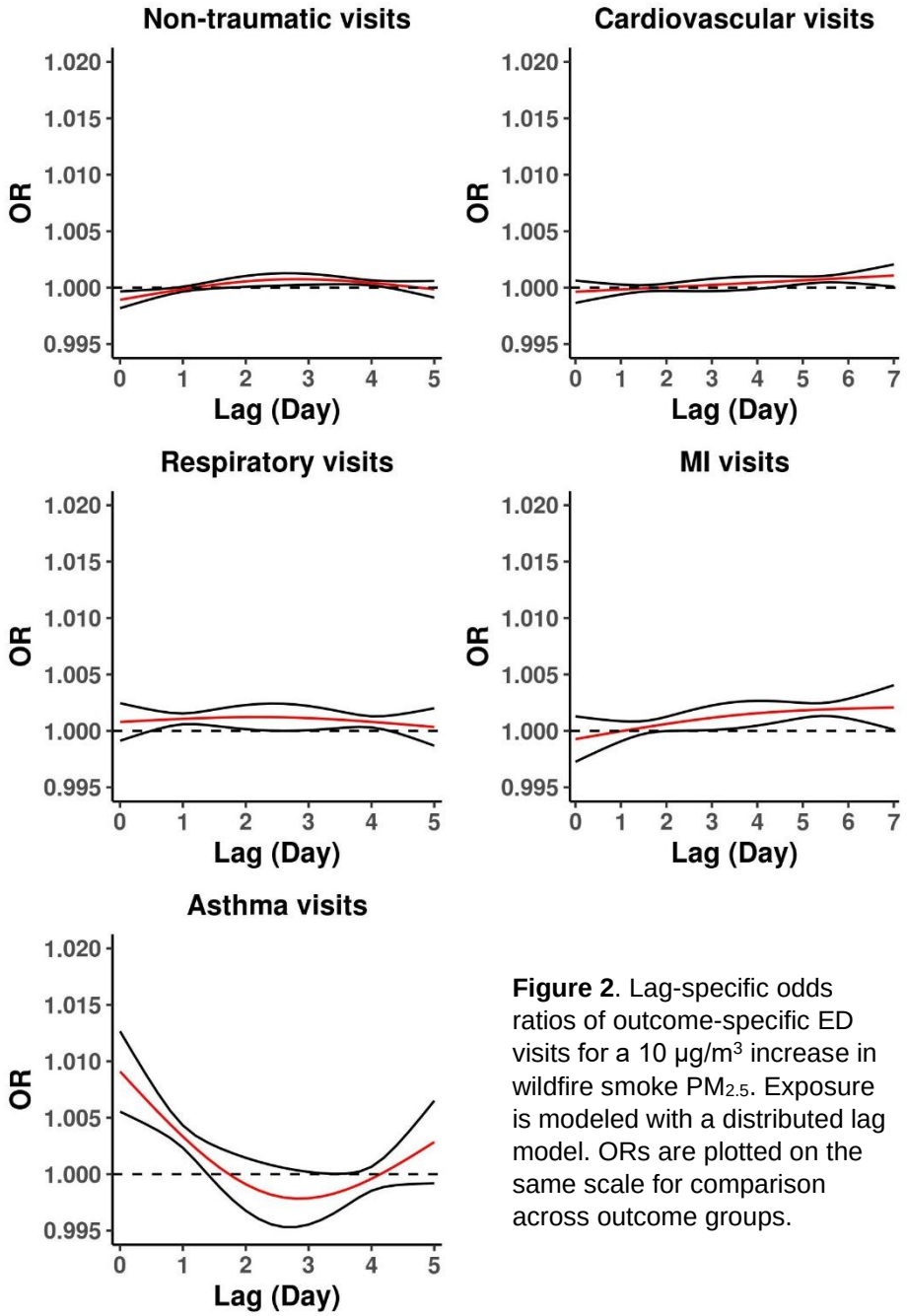


**Figure 2**. Lag-specific odds ratios of outcome-specific ED visits for a 10 μg/m$^3$ increase in wildfire smoke $PM_{2.5}$. Exposure is modeled with a distributed lag model. ORs are plotted on the same scale for comparison across outcome groups.

In our second primary analysis, we assessed the intensity of wildfire smoke exposure on ED visits. We separately examined the lag-specific and cumulative effects of exposure across 6 and 8 days for respiratory and cardiovascular visits, respectively. For all non-traumatic visits, we observed a lag-specific OR of about 1.00 on each lag day per 10 μg/m$^3$ increase in smoke-impacted $PM_{2.5}$. For all-cause respiratory visits, we observed a lag-specific OR of just greater than 1 across lag days 0-4, and for asthma visits, we observed an OR of 1.01 (CI: 1.01, 1.01) on lag day 0, and lower ORs through lag 3, with an increase in the OR on lags 4 and 5. For all-cause

cardiovascular and MI visits, we observed a lag-specific OR of 1 or slightly greater than 1 across most lag days (Figure 2, Table S12). Modeling the ORs cumulatively across the lag period shows an upward trajectory in cumulative risk for all outcomes (Figure S2, Table S13). Results from a quasi-Poisson model examining a lag-specific and cumulative effect of exposure were nearly unchanged (Tables S14 and S15).

In secondary exploratory analyses comparing the effect of wildfire smoke versus no wildfire smoke exposure, when stratified by age group (Table S6), results were consistent with a range of effects. For all-cause respiratory visits, results for ages 0-18 and ≥65 showed a decrease in risk or no increase in risk across all lags (Table S6). Results for ages 19-64 showed an increase in effect across all lags (ages 19-44) or lags 1-3 (ages 45-64). The results for asthma visits were similar, with generally higher ORs across all age groups. Children ages 5-18 were at increased risk of asthma visits on lags 0-1, with no further lagged effect. For all-cause cardiovascular ED visits, there was evidence of a range of effects across age groups, with a decreased risk for lags 0-2 among ages 65-84, and an increased risk among ages 19-44 for lags 4-6. For MI visits, we again observed a range of effects across age groups. Among ages ≥65, we observed a decreased risk on lags 0-1 (ages 65-84) or lags 0-3 (ages ≥85), and an increased risk among ages 19-44 on lags 2-7, and among ages 45-65 on lag 7.

We conducted a sensitivity analysis using a different binary indicator of exposure for our presence versus absence models. The results using the HMS smoke plumes were similar to the results with the primary exposure surface (Table S7). Like the primary analysis, the effect estimates for respiratory-related visits were elevated across the lag period, and for asthma-related visits was largest immediately following initial exposure (lag 0-1). Both cardiovascular and MI ED visits had the largest effect several days following initial smoke exposure (lag 4-5 for cardiovascular, and lag 3-7 for MI).

In an additional sensitivity analysis, we aimed to understand the impact of COVID-19 diagnoses on our results from the presence versus absence models. In 2020, we removed all

visits with a COVID-19 diagnosis (N=8,328), and found our results were nearly unchanged (Table S8). We also ran our main binary model excluding visits from 2017 due to differences in patterns of reporting in 2017 compared to 2018-2020 and due to the low number of EDs reporting visit data in 2017 (Table S9). Results were consistent with the main presence versus absence model (Figure 1; Table S4), although reported ORs were slightly larger across all outcomes and lags. We additional reported results stratified by year (Table S10). We observe the largest ORs and tightest confidence intervals in 2020, the year with all EDs included in the data. In a final sensitivity analysis, we ran all models using a quasi-Poisson time series analysis and found all results to be nearly unchanged (Table S11).

In a final secondary analysis, we sought to visualize the wildfire smoke-health effect relationship within the context of our existing public health messaging. Figure S3 models the exposure-response curve for all-cause respiratory visits on lag day 1 for a 10 $\mu g/m^3$ increase in total $PM_{2.5}$ across the Good, Moderate, Unhealthy for Sensitive Groups, Unhealthy, and Very Unhealthy categories of the US Environmental Protection Agency's air quality index (AQI). For all-cause respiratory visits, there was an approximately linear increase in odds as $PM_{2.5}$ increases. However, there is a large amount of uncertainty in the dose-response in the high end of the $PM_{2.5}$ range. Thus, we only modeled the exposure-response from 0-200 $\mu g/m^3$, as there was not enough exposure data above this threshold to yield meaningful information.

**Discussion**

We leveraged a comprehensive dataset to learn how wildfire smoke presence and intensity impact Washington, using syndromic surveillance data covering the entire state. We observed an increase in the odds of all-cause respiratory and asthma visits across the lag period, both comparing wildfire smoke vs non-wildfire smoke days, and for a 10 $\mu g/m^3$ increase in smoke $PM_{2.5}$. We observed an increase in the odds of all-cause cardiovascular and MI visits on lag day 6 and lag days 4-7, respectively, comparing wildfire smoke to non-wildfire smoke

days, while our models estimate a protective effect at lag day 0. Additionally, we observed a small increase in both cardiovascular and MI visits when exposure is modeled per 10 $\mu g/m^3$ increase in smoke $PM_{2.5}$.

Our respiratory results are consistent with the estimates reported in the literature, although our results are somewhat more muted. Other studies generally reported increased risk of respiratory impacts over the first 0-2 days following initial wildfire smoke exposure, particularly for asthma, although the magnitude of effect varies, and are larger than what we report in our continuous analysis (45,106,119). Alman et al. reported results across lag 0-2 that are consistent with an increased risk of asthma ED visits in Colorado using a case-crossover design. They reported an asthma lag 0 OR of 1.04 (95% CI: 1.02, 1.06) for a 5 $\mu g/m^3$ increase in $PM_{2.5}$ (45). Similarly, Hahn et al. reported results in Alaska using a case-crossover design across lag days 0-4 that are consistent with an increased risk of asthma ED visits, and with no increase in risk on lag day 5. Their reported same-day asthma OR was 1.13 (95% CI: 1.10, 1.16) for a 10 $\mu g/m^3$ increase in wildfire smoke $PM_{2.5}$ (106), similar to the OR reported by Alman et al. They also reported ORs across lag days 0-5 for all respiratory visits, which were increased on lag 0 (OR: 1.04; 95% CI: 1.02, 1.06), and consistent with no increase across lag days 1-5. Stowell et al. reported a 3-day average OR for asthma and respiratory visits (asthma OR: 1.08 (95% CI: 1.06, 1.11); respiratory OR: 1.02 (95% CI: 1.01, 1.03)) for a 1 $\mu g/m^3$ increase in wildfire smoke $PM_{2.5}$ in Colorado using a case crossover design (119). These results are much larger than those found by Alman et al. and Hahn et al., as they report per 1 $\mu g/m^3$ increase in wildfire smoke $PM_{2.5}$.

Effect estimates reported in the literature for cardiovascular ED visits are inconsistent. Like our results, Wettstein et al. reported no increase in the risk of all-cause cardiovascular visits on smoke versus non-smoke days across lag days 0-4 (46). Additionally, they found evidence for no increased risk of myocardial infarction visits across all lag days. However, to our

knowledge, no studies report a lagged effect of wildfire smoke on cardiovascular visits past 5 lag days. We observed increased risk of all-cause cardiovascular and MI visits several days after initial wildfire smoke exposure in both our presence versus absence and intensity models, suggesting a lagged effect of wildfire smoke on cardiovascular-related ED visits. It is possible that previous studies may have missed this lagged effect by not examining a long enough lag period. In the literature on $PM_{2.5}$ from non-wildfire smoke sources, several studies report a lag of 0-2 days for MI outcomes (120–122). In our study, we find an increase in MI-related ED visits for a lag of 3-7 days. An exposure to increased $PM_{2.5}$ may initiate a process of plaque rupture which can increase the risk of an MI (123). However, this process can take a couple of days, explaining the delayed effect in the general $PM_{2.5}$ exposure setting. In a wildfire smoke setting, based on increased wildfire smoke messaging and awareness, individuals may change their behavior by staying inside for the first couple days of an exposure event to reduce their exposure, which may delay MI events by additional day(s) compared to a setting of general $PM_{2.5}$ exposure. Additionally, we observed a strong delayed effect for MI outcomes in 2020 (Table S10), suggesting a possible confounding effect of the COVID-19 pandemic when individuals may have been delaying care more than usual (124,125). Furthermore, the high prevalence of multi-day smoke events during the study period may contribute to the delayed lag structure. We report lag-specific effects, however, the presence of multi-day smoke events builds cumulative exposure into this analysis. Thus, the effects we observe are likely not solely due to the initial smoke exposure on lag 0.

We aimed to understand the impact of wildfire smoke intensity by modeling the impact of a 10 $\mu g/m^3$ increase in smoke $PM_{2.5}$ on ED visits. The odds ratios from the lag-specific model are much smaller compared to those from the presence versus absence model, as we used different exposure metrics, and a distributed lag versus a single lag model. We found evidence of increased odds of respiratory and cardiovascular-related ED visits for each 10 $\mu g/m^3$ increase

in smoke $PM_{2.5}$, and our results follow similar trends to those reported in our presence versus absence analysis. In a secondary analysis, we examined the cumulative impact of exposure to a 10 $\mu g/m^3$ increase in smoke $PM_{2.5}$ over the full lag period. Few studies report the cumulative impact of wildfire smoke exposure over several days on health outcomes (42,108). Rappold et al. reported the cumulative impact of peat bog smoke versus no smoke across five days, and Yao et al. reported the cumulative OR of ambulance dispatches over 48 hours. Like our study, both studies found evidence of an increase in cumulative outcome-specific risk. These findings suggest a possible need for risk messaging in the several days following initial wildfire smoke exposure.

Our exposure-response curve shows an approximately linear increase in the odds of next day respiratory ED visits per 10 $\mu g/m^3$ increase in $PM_{2.5}$. This result indicates that an increase in $PM_{2.5}$ from 30 to 40 $\mu g/m^3$ may lead to a similar increase in the odds of next day respiratory ED visits as an increase from 100 to 110 $\mu g/m^3$, for example. This finding suggests a possible need for increased risk messaging when air quality is in lower ranges of the AQI (Moderate and Unhealthy for Sensitive Groups) in addition to higher ranges (Unhealthy and above). Above a concentration of 150-200 $\mu g/m^3$, there is a high degree of uncertainty in the true increase in odds due to the relatively few days above this threshold during our study period. Further work is needed to clarify how much the highest intensity of wildfire smoke impacts health outcomes.

Our analysis was bolstered by using syndromic surveillance data. Washington State requires all Emergency Departments to report health encounter information. Thus, by the last year of the study period, we were able to capture all visits to all emergency departments in Washington, allowing us to make inference across a large population and geographic area, rather than within one health care system or metropolitan area. Syndromic surveillance data also come with some limitations. The quality of the visit data varied by hospital and may have

varied across the reporting period; the data is only as good as what is entered by the hospital. To try to overcome this, we identified cases by using a combination of ICD-10 codes and queries in the chief complaint section. Furthermore, the syndromic surveillance platform in Washington was actively onboarding emergency departments during the study period, thus there was a different population captured by the data across time. This may have varied differentially by smoke versus non-smoke exposure areas, possibly yielding invalid inference. Thus, we chose to conduct a time-stratified case crossover analysis rather than a Poisson time series analysis, as the case crossover design does not require the population at risk to be known. Additionally, the time-stratified case crossover design compares exposure within each referent window, which is within a single month-year, decreasing the likelihood of bias. However, with the case crossover design, we assume patients do not have readmissions during their referent window. We cannot know this, since these data are visit-based rather than population based. While this is a limitation, we are confident in our results, as they are very similar to those from the Poisson time series analysis, which does not make this assumption. Overall, these data capture ED visits across a large population over four wildfire smoke seasons yielding effect estimates with high certainty relative to other studies.

Our study period included the initial period of the COVID-19 pandemic in June-September 2020. This was also a major wildfire smoke year for Washington. To ensure any effect we might see was not due to the increase in COVID-related ED visits, we ran a sensitivity analysis removing all ED visits with a COVID-19 diagnosis code. We found our effect estimates were nearly unchanged after removing these visits (Table S7). We additionally report results separately by year to better understand differences in EDs reporting by year (Table S10). We find that visits from 2020 dominate most of the positive effect we observe, and results from 2017 visits report protective effects, suggesting that the low coverage of EDs reporting visit data may have yielded biased estimates for 2017. We additionally re-ran our primary analysis excluding

visits from 2017 due to the low number of EDs reporting that year and to observed differences in outcomes reported (Table 1). We observed similar though slightly elevated effect estimates after removing 2017 compared to the primary model with 2017 visits (Table S8), suggesting that visits from 2017 are not impacting our overall estimates.

We utilized a previously developed exposure surface, population-weighted to the zip code level (50) (Supplemental Note S1). This surface leverages monitored and modeled $PM_{2.5}$ concentrations to better estimate daily local $PM_{2.5}$ concentrations compared to a model with only monitored $PM_{2.5}$. To assess the sensitivity of our results to the exposure model for the presence versus absence analysis, we used NOAA's HMS smoke plume daily data as a binary smoke exposure indicator, an exposure metric widely used in epidemiologic applications (46,106,126–129). We found the results from this analysis generally agree with our primary analysis leading to similar conclusions (Table S6). However, both our smoke exposure surface and the HMS smoke plumes likely result in some degree of misclassification. For the HMS smoke plume exposure, we only used smoke polygons categorized as 'Heavy', which tend to capture all major wildfire smoke plumes, but may miss some of the smoke-impacted days at lower $PM_{2.5}$ concentrations. On the other hand, our wildfire smoke exposure surface may be classifying some days as wildfire smoke-impacted that are not truly smoke-impacted. However, the results from these two methods yielded similar risk estimates (Table S6).

**Conclusion**

This study adds to the growing body of research supporting evidence of the adverse impacts of wildfire smoke exposure on population health. We reported an increase in all-cause respiratory ED visits, asthma ED visits, and a lagged increase in all-cause cardiovascular and MI-related ED visits using a large statewide dataset over four wildfire smoke seasons in Washington state. We find all-cause respiratory visits are most elevated among ages 19-64, and

asthma visits are most elevated among ages 5-64, suggesting that messaging and risk communication need to be targeted toward younger and middle-aged adults.

**Supplemental Material**

**Note S1.** Primary exposure grid methods

$PM_{2.5}$ in Washington State is monitored at 75 State and Local Air Monitoring Stations (SLAMS), and meteorological variables, measured at 18 stations, including temperature, dew point, and wind speed (130). The spatial coverage of the 75 regulatory air quality monitors in Washington is not adequate for estimating daily exposure to wildfire smoke due to the unique terrain and meteorology (131). Instead, we relied on the Air Indicator Report for Public Awareness and Community Tracking (AIRPACT-4) modeled $PM_{2.5}$, combined with the monitored daily $PM_{2.5}$ concentrations to produce a more accurate exposure grid (132). The AIRPACT-4 model domain is composed of 4x4 km grid cells. AIRPACT-4 models air quality at high resolution across the Pacific Northwest. However, the AIRPACT-4 model does not produce reliable daily $PM_{2.5}$ estimates. Rather, it performs better over longer time periods. Further, AIRPACT-4 accurately predicts relative concentrations, that is, whether one area has a higher $PM_{2.5}$ concentration than another area, but does a relatively poor job of predicting absolute concentrations. Thus, we decided to incorporate monitored $PM_{2.5}$ concentrations into our exposure assessment.

To create the $PM_{2.5}$ exposure grid, we assigned monitored data from the 75 regulatory monitors in Washington to corresponding grid cells of the AIRPACT model (132). This involved several steps:

1) To yield daily $PM_{2.5}$ concentrations at each grid cell, grid cells were assigned the $PM_{2.5}$ concentration of the nearest active and representative monitor on each day. A monitor was considered representative of a grid cell if the difference between the summertime mean of the grid cell (determined in steps 2a and b below) and the summertime mean of the monitor was less than 2 μg/m3. If a grid cell's nearest monitor was not representative

or not active on a given day, the grid cell was assigned its second- or third-nearest monitor if both active and representative.

2) To determine the summer time mean at the grid cell:
    a. Each monitoring site was assigned the ratio of the monitored summertime mean $PM_{2.5}$ concentration to the AIRPACT-modeled summertime mean $PM_{2.5}$ concentration. Ratios were interpolated across the grid cells using Empirical Bayesian Kriging (133).
    b. The interpolated ratio at each grid cell was multiplied by the (AIRPACT) modeled 2017-2020 summertime mean $PM_{2.5}$. We then had estimates of summertime means for each grid cell from the AIRPACT model, adjusted by monitored data to correct AIRPACT's inadequate predictions of absolute concentrations.
3) Each monitoring site was matched to its nearest National Weather Service meteorological site to obtain meteorological variables. We used daily average temperature and dew point to calculate humidex, a measure of apparent temperature calculated from air temperature and dew point (113), for each monitor. Grid cells were then assigned the humidex value associated with their assigned $PM_{2.5}$ monitor on each day.
4) The end result is a dataset with the following for each day and each grid cell: 24-hour average $PM_{2.5}$ concentration and humidex from a neighboring monitoring site with the population attributed to that grid cell.

Next, the exposure grid was overlaid with the Washington State Office of Financial Management's (OFM) yearly population estimates at the census block group level (134). In cases where census block group boundaries crossed grid cells, we determined the percent of the population's census block group attributed to each grid cell by the area that fell within the

grid cell. This method assumes populations are evenly distributed within each census block group and each grid cell. A small percentage of the grid cells and corresponding populations were not considered to be represented by any monitoring site, and were excluded from the analysis.

**Equation S1**. Conditional logistic regression equation notation

$$logit(p_{it}) = \beta_1 Smoke_{it} + \beta_2 Humidex_{it}(ns, 3\ df) + \alpha_{stratum(i)}$$

Conditional logistic regression equation, where $p_{it}$ is the probability of being a case (having a visit), and $logit(p_{it})$ is the log odds ($p_{it}/(1-p_{it})$) comparing individuals to themselves on the day of their visit compared to reference days. $Smoke_{it}$ is the binary smoke indicator for day *t* in the zip code where individual *i* lives. Similarly, $Humidex_{it}$ is the time- and zip code-specific humidex term with values modeled as a natural cubic spline with 3 degrees of freedom. $\alpha_{stratum(i)}$ is the referent window stratum for each individual, signifying that all comparisons are within individual.

**Table S1**. Outcome Assessment: Diagnosis (ICD-10) codes and Chief Complaint query text

| Outcome | ICD-10 codes | Chief Complaint query |
|---|---|---|
| All non-traumatic visits | A01-R99 | |
| Respiratory | J01-J99 | wheez\|shortness of breath\|difficulty breathing\|chest tightness\|cough\|breathing problem\|asthma\|bronchitis\|dyspnea |
| Asthma | J45 | asthma |
| Cardiovascular | I05-I52, I60-63, I65-69, G45 | chest pain\|myocardial infarction\|MI\|STEMI\|heart attack\|angina\|acute coronary syndrome\|ACS |
| Myocardial infarction | I21 | chest pain\|myocardial infarction\|MI\|STEMI\|heart attack |
| COVID | U07 | |

To identify cases in each outcome category, we examined all ICD-10 diagnosis code fields, and additionally queried the Chief Complaint field. The Chief Complaint field is a text field that typically lists the main complaint for the visit, as written by the care provider. The queries listed in the table are the text queries that were searched to identify additional cases.

**Figure S1**. Distribution of total PM2.5 on smoke versus non-smoke days

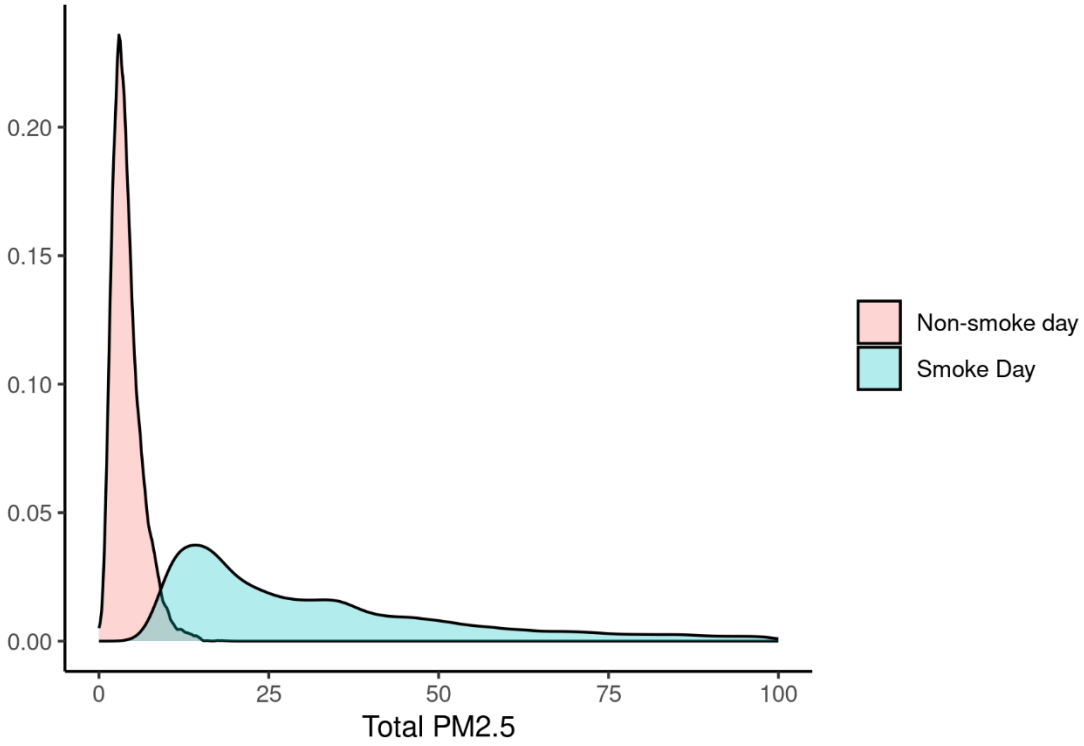

**Table S2.** Summary of smoke days at the referent window level[1]

| | All years | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|
| Referent windows (N)[1] | 1,864,470 | 235,314 | 424,833 | 629,702 | 574,621 |
| Mean (SD) days per referent window[2] | 4.4 (0.6) | 4.4 (0.6) | 4.4 (0.6) | 4.4 (0.6) | 4.4 (0.6) |
| Mean (SD) smoke days per referent window | 0.4 (0.8) | 0.8 (1.0) | 0.7 (1.1) | 0.1 (0.2) | 0.5 (0.8) |
| | **All years** | **2017** | **2018** | **2019** | **2020** |
| **Count of referent windows with 0, 1, ≥2 smoke days (N (%))** | | | | | |
| 0 smoke days | 1,409,315 (76)[3] | 113,352 (48) | 274,914 (65) | 597,906 (95) | 423,143 (74) |
| 1 smoke day | 174,063 (9) | 64,425 (27) | 47,428 (11) | 29,165 (5) | 33,045 (6) |
| ≥2 smoke days | 281,092 (15) | 57,537 (24) | 102,491 (24) | 2,631 (<1) | 118,433 (21) |
| **$PM_{2.5}$ within referent windows** | | | | | |
| Mean (SD) total $PM_{2.5}$[4] | 9.4 (24.5) | 10.2 (18.3) | 9.6 (14.6) | 4.3 (2.4) | 14.4 (40.5) |
| Mean (SD) wildfire smoke $PM_{2.5}$ | 5.0 (24.5) | 5.5 (18.2) | 5.0 (14.7) | 0.2 (1.8) | 10.0 (40.6) |

[1]A single referent window refers to the day of each individual visit (event) and corresponding set of referent days. The number of referent windows in the dataset is equivalent to the total number of visits, as each visit is assigned a set of referent days for comparison.

[2] The number of days per referent window corresponds to the index day plus the referent days.

[3] Percentages rounded to the nearest whole number. Percentages may not add up to 100%.

[4] Mean (SD) of within referent window mean $PM_{2.5}$, either total $PM_{2.5}$ or smoke-specific $PM_{2.5}$.

**Table S3.** Summary of smoke days and $PM_{2.5}$ comparing primary exposure surface to HMS smoke plumes

| | Primary exposure | HMS smoke plumes |
|---|---|---|
| **Total zip code smoke days (N (%))[1]** | | |
| **All years** | 36,411 (12.7) | 24,042 (8.4) |
| **2017** | 14,012 (19.5) | 9,066 (12.6) |
| **2018** | 13,374 (18.6) | 8,266 (11.5) |
| **2019** | 1,096 (1.5) | 1,055 (1.5) |
| **2020** | 7,929 (11.1) | 5,655 (7.9) |
| **Mean (SD) $PM_{2.5}$ ($\mu g/m^3$) on smoke days[2]** | | |
| **All years** | 48.7 (9.8) | 55.0 (15.4) |
| **2017** | 33.4 (8.8) | 36.2 (17.0) |
| **2018** | 37.3 (9.2) | 40.1 (15.1) |
| **2019** | 17.7 (3.9) | 13.1 (7.0) |
| **2020** | 94.7 (32.0) | 113.0 (35.2) |
| **Mean (SD) $PM_{2.5}$ ($\mu g/m^3$) on non-smoke days[2]** | | |
| **All years** | 4.3 (1.1) | 5.5 (1.3) |
| **2017** | 4.5 (1.3) | 6.2 (1.7) |
| **2018** | 4.4 (1.0) | 6.4 (1.9) |
| **2019** | 4.0 (1.0) | 4.1 (1.0) |
| **2020** | 4.1 (0.9) | 5.6 (1.5) |

**Number of zip code days categorized as smoke versus non-smoke comparing the main method to the HMS method**

| | | HMS Method | |
|---|---|---|---|
| **Main method** | | **Smoke** | **Non-smoke** |
| | **Smoke** | 19,880 | 16,402 |
| | **Non-smoke** | 4,162 | 243,605 |

[1]All 592 zip codes in Washington are included in the study.

[2]Averages of zip code-level averages

**Table S4**. Binary exposure results: ORs and 95% CIs for ED visit categories by lag day for wildfire smoke vs non-wildfire smoke days

| Outcome Category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 0.99 (0.99, 1.00) | 1.00 (0.99, 1.00) | **0.99 (0.99, 1.00)** | 1.00 (0.99, 1.00) | 1.00 (0.99, 1.00) | 1.00 (0.99, 1.01) | | |
| Respiratory | 1.01 (1.00, 1.03) | **1.02 (1.00, 1.03)** | **1.03 (1.01, 1.04)** | **1.02 (1.00, 1.04)** | **1.02 (1.00, 1.04)** | **1.02 (1.00, 1.04)** | | |
| Asthma | **1.13 (1.10, 1.17)** | **1.13 (1.09, 1.17)** | **1.13 (1.09, 1.17)** | **1.10 (1.06, 1.13)** | **1.08 (1.04, 1.12)** | **1.05 (1.02, 1.09)** | | |
| Cardiovascular | **0.98 (0.97, 1.00)** | 0.99 (0.98, 1.00) | **0.99 (0.98, 1.00)** | 1.00 (0.99, 1.01) | 1.01 (1.00, 1.02) | 1.01 (0.99, 1.02) | **1.02 (1.00, 1.03)** | 1.01 (1.00, 1.02) |
| Myocardial infarction | **0.97 (0.95, 0.99)** | 0.98 (0.96, 1.01) | 0.99 (0.97, 1.02) | 1.03 (1.00, 1.05) | **1.04 (1.01, 1.06)** | **1.04 (1.02, 1.07)** | **1.04 (1.02, 1.07)** | **1.04, 1.01, 1.06)** |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S5**. Binary results from mixed effects meta-analysis: ORs and 95% CIs for ED visit categories by lag day for wildfire smoke vs non-wildfire smoke days

| Outcome Category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 0.99 (0.97, 1.01) | 0.99 (0.97, 1.01) | 0.98 (0.96, 1.01) | 0.99 (0.96, 1.02) | 0.99 (0.96, 1.02) | 0.99 (0.96, 1.02) | | |
| Respiratory | 1.00 (0.95, 1.04) | 1.01 (0.96, 1.05) | 1.02 (0.97, 1.06) | 1.01 (0.96, 1.07) | 1.01 (0.95, 1.07) | 1.00 (0.94, 1.07) | | |
| Asthma | **1.12 (1.05, 1.21)** | **1.13 (1.05, 1.21)** | **1.12 (1.04, 1.21)** | **1.08 (1.00, 1.17)** | 1.07 (0.98, 1.17) | 1.03 (0.94, 1.13) | | |
| Cardiovascular | 0.98 (0.95, 1.00) | 0.98 (0.94, 1.01) | 0.97 (0.94, 1.01) | 0.99 (0.96, 1.02) | 0.99 (0.95, 1.03) | 0.99 (0.95, 1.03) | 1.00 (0.96, 1.04) | 0.99 (0.94, 1.04) |
| Myocardial infarction | **0.97 (0.95, 0.99)** | 0.97 (0.92, 1.02) | 0.98 (0.92, 1.03) | 1.01 (0.95, 1.08) | 1.02 (0.95, 1.08) | 1.02 (0.95, 1.10) | 1.03 (0.94, 1.11) | 1.01 (0.92, 1.12) |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S6**. Binary exposure results by age group: ORs (95% CI) for ED visit categories by age group and lag day for wildfire smoke vs non-wildfire smoke days

| Outcome Category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| Respiratory | | | | | | | | |
| 0-4 | 0.98 (0.92, 1.04) | 1.00 (0.94, 1.06) | 0.97 (0.91, 1.03) | 0.95 (0.90, 1.01) | 0.97 (0.92, 1.03) | 0.97 (0.92, 1.03) | | |
| 5-18 | 0.99 (0.93, 1.04) | 1.00 (0.95, 1.06) | 1.02 (0.96, 1.07) | 1.01 (0.95, 1.06) | 1.00 (0.94, 1.06) | 1.00 (0.94, 1.05) | | |
| 19-44 | **1.07 (1.04, 1.10)** | **1.07 (1.04, 1.10)** | **1.09 (1.06, 1.12)** | **1.07 (1.04, 1.10)** | **1.06 (1.03, 1.09)** | **1.06 (1.03, 1.09)** | | |
| 45-64 | 1.03 (1.00, 1.06) | **1.04 (1.00, 1.07)** | **1.03 (1.00, 1.06)** | **1.04 (1.01, 1.07)** | 1.03 (1.00, 1.06) | 1.01 (0.98, 1.04) | | |
| 65-84 | **0.96 (0.93, 0.99)** | 0.97 (0.94, 1.00) | 0.99 (0.96, 1.02) | 0.99 (0.96, 1.02) | 1.00 (0.97, 1.03) | 1.01 (0.98, 1.04) | | |
| 85+ | 0.96 (0.91, 1.02) | 0.98 (0.92, 1.04) | 1.00 (0.95, 1.07) | 1.03 (0.97, 1.09) | 1.06 (1.00, 1.12) | 1.04 (0.98, 1.10) | | |
| Asthma | | | | | | | | |
| 0-4 | 1.05 (0.87, 1.28) | 1.04 (0.87, 1.26) | 0.99 (0.83, 1.2) | 0.96 (0.8, 1.15) | 1.04 (0.87, 1.24) | 0.97 (0.81, 1.17) | | |
| 5-18 | **1.17 (1.05, 1.3)** | **1.14 (1.02, 1.27)** | 1.1 (0.99, 1.22) | 1.09 (0.98, 1.21) | 1.09 (0.98, 1.21) | 1.05 (0.94, 1.16) | | |
| 19-44 | **1.18 (1.12, 1.24)** | **1.18 (1.12, 1.24)** | **1.18 (1.13, 1.25)** | **1.13 (1.08, 1.19)** | **1.08 (1.03, 1.14)** | **1.07 (1.01, 1.12)** | | |
| 45-64 | **1.14 (1.07, 1.23)** | **1.14 (1.06, 1.22)** | **1.13 (1.05, 1.21)** | **1.09 (1.02, 1.17)** | **1.11 (1.03, 1.19)** | 1.07 (0.99, 1.14) | | |
| 65-84 | 1.03 (0.94, 1.13) | 1.03 (0.94, 1.14) | 1.05 (0.95, 1.15) | 1.05 (0.96, 1.16) | 1.05 (0.95, 1.15) | 1.02 (0.93, 1.12) | | |
| 85+ | 0.86 (0.69, 1.08) | 0.9 (0.72, 1.12) | 0.9 (0.72, 1.12) | 1.03 (0.83, 1.28) | 1.05 (0.85, 1.31) | 1.02 (0.82, 1.26) | | |
| Cardiovascular | | | | | | | | |
| 19-44 | 1.02 (0.99, 1.05) | 1.01 (0.99, 1.04) | 1.02 (0.99, 1.05) | 1.01 (0.99, 1.04) | **1.03 (1.00, 1.06)** | **1.03 (1.00, 1.06)** | **1.03 (1.00, 1.06)** | 1.01 (0.98, 1.04) |
| 45-64 | 0.99 (0.97, 1.01) | 1.00 (0.98, 1.02) | 0.99 (0.96, 1.01) | 1.00 (0.98, 1.02) | 1.01 (0.98, 1.03) | 1.01 (0.99, 1.03) | 1.02 (1.00, 1.04) | 1.02 (1.00, 1.04) |
| 65-84 | **0.96 (0.94, 0.98)** | **0.97 (0.95, 0.99)** | **0.97 (0.95, 0.99)** | 0.99 (0.97, 1.01) | 0.99 (0.97, 1.02) | 0.99 (0.97, 1.01) | 1.01 (0.99, 1.03) | 1.01 (0.99, 1.03) |
| 85+ | 0.99 (0.95, 1.03) | 0.98 (0.95, 1.02) | 0.99 (0.95, 1.03) | 0.99 (0.95, 1.03) | 1.00 (0.97, 1.04) | 1.00 (0.96, 1.03) | 1.01 (0.97, 1.05) | 0.97 (0.93, 1.01) |
| Myocardial infarction | | | | | | | | |
| 19-44 | 1.01 (0.97, 1.06) | 1.02 (0.97, 1.07) | **1.06 (1.01, 1.10)** | **1.08 (1.03, 1.13)** | **1.10 (1.05, 1.15)** | **1.10 (1.05, 1.15)** | **1.09 (1.04, 1.14)** | **1.05 (1.01, 1.10)** |
| 45-64 | 0.99 (0.95, 1.03) | 0.99 (0.95, 1.04) | 0.97 (0.93, 1.01) | 1.02 (0.98, 1.07) | 1.04 (1.00, 1.08) | 1.03 (0.99, 1.07) | 1.02 (0.98, 1.07) | **1.05 (1.00, 1.09)** |
| 65-84 | **0.92 (0.88, 0.97)** | **0.95 (0.91, 1.00)** | 0.96 (0.91, 1.00) | 0.99 (0.94, 1.04) | 0.99 (0.94, 1.03) | 1.00 (0.96, 1.05) | 1.04 (0.99, 1.09) | 1.03 (0.98, 1.08) |
| 85+ | **0.85 (0.77, 0.95)** | **0.86 (0.77, 0.95)** | **0.85 (0.76, 0.94)** | **0.86 (0.77, 0.96)** | 0.92 (0.82, 1.02) | 0.95 (0.86, 1.06) | 0.96 (0.86, 1.07) | 0.93 (0.83, 1.03) |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S7**. Results using HMS exposure compared to results with primary exposure: ORs and 95% CIs for ED visits by lag day for wildfire smoke vs non-wildfire smoke days using HMS smoke plumes

| Outcome | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| **Exposure: HMS Smoke Plumes** | | | | | | | | |
| All non-traumatic | **0.99 (0.98, 1.00)** | 0.99 (0.99, 1.00) | **0.99 (0.98, 1.00)** | **0.99 (0.99, 1.00)** | 1.00 (0.99, 1.01) | **1.01 (1.00, 1.02)** | | |
| Respiratory | 1.02 (1.00, 1.04) | **1.02 (1.00, 1.04)** | 1.02 (1.00, 1.03) | 1.01 (1.00, 1.03) | **1.02 (1.01, 1.04)** | 1.01 (1.00, 1.03) | | |
| Asthma | **1.11 (1.07, 1.15)** | **1.11 (1.07, 1.16)** | **1.11 (1.06 1.15)** | **1.08 (1.04, 1.13)** | **1.05 (1.01, 1.10)** | 1.02 (0.98, 1.06) | | |
| Cardiovascular | **0.98 (0.97, 0.99)** | **0.98 (0.97, 0.99)** | 0.99 (0.98, 1.00) | 1.00 (0.98, 1.01) | 1.01 (1.00, 1.03) | **1.02 (1.01, 1.04)** | **1.00 (0.99, 1.01)** | 1.00 (0.99, 1.02) |
| Myocardial infarction | 0.98 (0.96, 1.01) | 1.00 (0.97, 1.02) | 1.01 (0.98, 1.04) | **1.03 (1.00, 1.06)** | **1.05 (1.02, 1.08)** | **1.07 (1.04, 1.10)** | **1.05 (1.02, 1.08)** | **1.04 (1.01, 1.07)** |
| **Exposure: Primary exposure surface** | | | | | | | | |
| All non-traumatic | 0.99 (0.99, 1.00) | 1.00 (0.99, 1.00) | **0.99 (0.99, 1.00)** | 1.00 (0.99, 1.00) | 1.00 (0.99, 1.00) | 1.00 (0.99, 1.01) | | |
| Respiratory | 1.01 (1.00, 1.03) | **1.02 (1.00, 1.03)** | **1.03 (1.01, 1.04)** | **1.02 (1.00, 1.04)** | **1.02 (1.00, 1.04)** | **1.02 (1.00, 1.04)** | | |
| Asthma | **1.13 (1.10, 1.17)** | **1.13 (1.09, 1.17)** | **1.13 (1.09, 1.17)** | **1.10 (1.06, 1.13)** | **1.08 (1.04, 1.12)** | **1.05 (1.02, 1.09)** | | |
| Cardiovascular | **0.98 (0.97, 1.00)** | 0.99 (0.98, 1.00) | **0.99 (0.98, 1.00)** | 1.00 (0.99, 1.01) | 1.01 (1.00, 1.02) | 1.01 (0.99, 1.02) | **1.02 (1.00, 1.03)** | 1.01 (1.00, 1.02) |
| Myocardial infarction | **0.97 (0.95, 0.99)** | 0.98 (0.96, 1.01) | 0.99 (0.97, 1.02) | 1.03 (1.00, 1.05) | **1.04 (1.01, 1.06)** | **1.04 (1.02, 1.07)** | **1.04 (1.02, 1.07)** | **1.04, 1.01, 1.06)** |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S8**. Binary exposure results excluding COVID-19 diagnoses at the visit-level: ORs and 95% CIs for ED visit categories by lag day for wildfire smoke vs non-wildfire smoke days, excluding COVID diagnoses

| Outcome category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 |
|---|---|---|---|---|---|---|
| All non-traumatic | 0.99 (0.99, 1.00) | 1.00 (0.99, 1.00) | **0.99 (0.99, 1.00)** | 1.00 (0.99, 1.00) | 1.00 (0.99, 1.00) | 1.00 (0.99, 1.01) |
| Respiratory | 1.01 (1.00, 1.03) | **1.02 (1.00, 1.04)** | **1.03 (1.01, 1.05)** | **1.02 (1.01, 1.04)** | **1.02 (1.01, 1.04)** | **1.02 (1.00, 1.04)** |
| Asthma | **1.14 (1.10, 1.18)** | **1.14 (1.10, 1.18)** | **1.13 (1.09, 1.17)** | **1.10 (1.06, 1.14)** | **1.08 (1.04, 1.12)** | **1.05 (1.02, 1.09)** |

Note: Bolded ORs are statistically significant at the p<0.05 level

**Table S9**. Binary exposure results excluding 2017: ORs and 95% CIs for ED visit categories by lag day for wildfire smoke vs non-wildfire smoke days

| Outcome Category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 1.00 (0.99, 1.01) | 1.01 (1.00, 1.01) | 1.00 (1.00, 1.01) | **1.01 (1.00, 1.02)** | **1.01 (1.00, 1.02)** | **1.01 (1.00, 1.02)** | | |
| Respiratory | **1.02 (1.01, 1.04)** | **1.04 (1.02, 1.05)** | **1.05 (1.03, 1.07)** | **1.05 (1.03, 1.07)** | **1.05 (1.04, 1.07)** | **1.05 (1.03, 1.07)** | | |
| Asthma | **1.14 (1.10, 1.18)** | **1.15 (1.10, 1.19)** | **1.14 (1.10, 1.19)** | **1.12 (1.08, 1.17)** | **1.11 (1.07, 1.15)** | **1.08 (1.04, 1.12)** | | |
| Cardiovascular | 0.99 (0.98, 1.00) | 1.00 (0.98, 1.01) | 0.99 (0.98, 1.01) | 1.01 (0.99, 1.02) | **1.02 (1.00, 1.03)** | **1.02 (1.00, 1.03)** | **1.03 (1.01, 1.04)** | **1.02 (1.01, 1.04)** |
| Myocardial infarction | 0.98 (0.95, 1.01) | 1.00 (0.97, 1.02) | 1.01 (0.98, 1.03) | **1.04 (1.01, 1.07)** | **1.06 (1.03, 1.09)** | **1.07 (1.04, 1.10)** | **1.07 (1.04, 1.10)** | **1.07 (1.04, 1.10)** |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S10**. Binary results stratified by year: ORs and 95% CIs for ED visit categories by lag day for wildfire smoke vs non-wildfire smoke days

| Outcome Category | | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|---|
| All non-traumatic | | | | | | | | | |
| | 2017 | 0.97 (0.95, 0.98) | 0.97 (0.96, 0.98) | 0.96 (0.95, 0.97) | 0.96 (0.94, 0.97) | 0.96 (0.94, 0.97) | 0.96 (0.95, 0.97) | | |
| | 2018 | 1.01 (1.00, 1.02) | 1.00 (0.99, 1.02) | 1.00 (0.99, 1.02) | 1.00 (0.99, 1.02) | 1.01 (1.00, 1.02) | 1.01 (1.00, 1.02) | | |
| | 2019 | 0.97 (0.95, 1.00) | 0.97 (0.94, 0.99) | 0.96 (0.93, 0.99) | 0.98 (0.96, 1.01) | 0.98 (0.95, 1.00) | 0.98 (0.95, 1.00) | | |
| | 2020 | 1.00 (0.99, 1.01) | **1.01 (1.00, 1.02)** | 1.01 (1.00, 1.02) | **1.02 (1.01, 1.03)** | **1.02 (1.01, 1.03)** | **1.02 (1.01, 1.03)** | | |
| Respiratory | | | | | | | | | |
| | 2017 | 0.96 (0.93, 0.99) | 0.96 (0.93, 0.99) | 0.96 (0.93, 0.99) | 0.94 (0.91, 0.97) | 0.93 (0.90, 0.96) | 0.93 (0.90, 0.96) | | |
| | 2018 | 1.02 (0.99, 1.05) | 1.02 (0.99, 1.05) | **1.04 (1.01, 1.07)** | **1.03 (1**.00**, 1.06)** | **1.04 (1.01, 1.06)** | **1.03 (1.01, 1.06)** | | |
| | 2019 | 0.95 (0.89, 1.02) | 0.98 (0.92, 1.05) | 1.00 (0.93, 1.06) | 1.01 (0.95, 1.08) | 0.99 (0.93, 1.05) | 0.96 (0.9, 1.03) | | |
| | 2020 | **1.04 (1.01, 1.06)** | **1.05 (1.03, 1.08)** | **1.06 (1.04, 1.09)** | **1.07 (1.05, 1.10)** | **1.08 (1.05, 1.10)** | **1.07 (1.05, 1.10)** | | |
| Asthma | | | | | | | | | |
| | 2017 | **1.11 (1.02, 1.22)** | 1.06 (0.97, 1.15) | 1.04 (0.96, 1.13) | 0.98 (0.90, 1.06) | 0.94 (0.87, 1.02) | 0.94 (0.87, 1.02) | | |
| | 2018 | **1.22 (1.14, 1.31)** | **1.23 (1.15, 1.32)** | **1.23 (1.15, 1.32)** | **1.17 (1.09, 1.26)** | **1.16 (1.08, 1.25)** | **1.14 (1.07, 1.23)** | | |
| | 2019 | 0.97 (0.83, 1.14) | 1.08 (0.92, 1.26) | 1.07 (0.91, 1.25) | 1.07 (0.91, 1.24) | 1.11 (0.95, 1.29) | 0.95 (0.82, 1.12) | | |
| | 2020 | **1.12 (1.07, 1.17)** | **1.12 (1.07, 1.17)** | **1.12 (1.06, 1.17)** | **1.11 (1.06, 1.16)** | **1.09 (1.04, 1.14) *** | **1.06 (1.01, 1.11)** | | |
| Cardiovascular | | | | | | | | | |
| | 2017 | 0.95 (0.92, 0.98) | 0.95 (0.93, 0.98) | 0.96 (0.93, 0.98) | 0.97 (0.94, 1.00) | 0.97 (0.94, 1.00) | 0.97 (0.94, 0.99) | 0.96 (0.94, 0.99) | 0.95 (0.92, 0.98) |
| | 2018 | 0.99 (0.97, 1.02) | 0.99 (0.96, 1.01) | 0.98 (0.96, 1.01) | 0.99 (0.97, 1.01) | 0.99 (0.97, 1.02) | 1.00 (0.97, 1.02) | 1.01 (0.99, 1.04) | 1.00 (0.98, 1.02) |
| | 2019 | 0.94 (0.9, 0.99) | 0.93 (0.89, 0.98) | 0.93 (0.88, 0.97) | 0.96 (0.92, 1.01) | 0.95 (0.90, 0.99) | 0.95 (0.91, 1.00) | 0.96 (0.92, 1.01) | 0.95 (0.91, 1.00) |
| | 2020 | 0.99 (0.98, 1.01) | 1.01 (1.00, 1.03) | 1.01 (0.99, 1.03) | **1.02 (1.01, 1.04)** | **1.04 (1.02, 1.06)** | **1.04 (1.02, 1.06)** | **1.05 (1.03, 1.07)** | **1.05 (1.03, 1.07)** |
| Myocardial infarction | | | | | | | | | |
| | 2017 | 0.94 (0.88, 0.99) | 0.94 (0.89, 1.00) | 0.94 (0.89, 0.99) | 0.96 (0.91, 1.01) | 0.98 (0.93, 1.03) | 0.96 (0.91, 1.01) | 0.96 (0.91, 1.01) | 0.92 (0.88, 0.97) |
| | 2018 | 0.99 (0.95, 1.03) | 0.96 (0.92, 1.00) | 0.97 (0.93, 1.01) | 1.00 (0.95, 1.04) | 1.00 (0.95, 1.04) | 1.00 (0.96, 1.05) | 1.01 (0.97, 1.06) | 0.99 (0.95, 1.04) |
| | 2019 | 0.92 (0.83, 1.01) | 0.92 (0.83, 1.01) | 0.92 (0.84, 1.02) | 0.99 (0.9, 1.09) | 0.98 (0.89, 1.08) | 1.00 (0.91, 1.10) | 1.00 (0.9, 1.10) | 0.98 (0.89, 1.08) |
| | 2020 | 0.98 (0.94, 1.02) | 1.04 (1.00, 1.08) | **1.05 (1.01, 1.09)** | **1.09 (1.05, 1.14)** | **1.12 (1.08, 1.16)** | **1.13 (1.09, 1.18)** | **1.14 (1.10, 1.19)** | **1.17 (1.12, 1.21)** |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S11**. Binary exposure results from quasi-Poisson models: ORs and 95% CIs for ED visit categories by lag day for wildfire smoke vs non-wildfire smoke days

| Outcome Category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 1.00 (0.98, 1.01) | 1.00 (0.99, 1.02) | 1.00 (0.99, 1.02) | 1.01 (1.00, 1.03) | **1.02 (1.00, 1.03)** | **1.02 (1.00, 1.03)** | | |
| Respiratory | **1.03 (1.00, 1.06)** | **1.04 (1.01, 1.07)** | **1.05 (1.03, 1.08)** | **1.06 (1.03, 1.08)** | **1.06 (1.03, 1.09)** | **1.06 (1.03, 1.08)** | | |
| Asthma | **1.15 (1.10, 1.21)** | **1.14 (1.09, 1.20)** | **1.14 (1.09, 1.19)** | **1.12 (1.07, 1.17)** | **1.10 (1.05, 1.15)** | **1.08 (1.03, 1.13)** | | |
| Cardiovascular | 0.99 (0.97, 1.02) | 1.00 (0.98, 1.02) | 0.99 (0.97, 1.02) | 1.01 (0.98, 1.03) | 1.01 (0.99, 1.03) | 1.01 (0.99, 1.03) | 1.01 (0.99, 1.04) | 1.01 (0.99, 1.03) |
| Myocardial infarction | 0.97 (0.94, 1.01) | 0.99 (0.95, 1.02) | 1.00 (0.96, 1.03) | 1.02 (0.99, 1.06) | **1.03 (1.00, 1.07)** | **1.04 (1.00, 1.07)** | **1.03 (1.00, 1.07)** | 1.02 (0.99, 1.06) |

Note: Bolded ORs are statistically significant at the p<0.05 level.

**Table S12**. Lag-specific exposure results: ORs and 95% CIs for ED visits per 10 μg/m$^3$ increase in smoke-impacted $PM_{2.5}$, from a distributed lag model

| Outcome category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 0.999 (0.998, 1.000) | 1.000 (1.000, 1.000) | **1.001 (1.000, 1.001)** | **1.001 (1.000, 1.001)** | 1.000 (1.000, 1.001) | 1.000 (0.999, 1.001) | | |
| Respiratory | 1.001 (0.999, 1.002) | **1.001 (1.001, 1.002)** | **1.001 (1.000, 1.002)** | **1.001 (1.000, 1.002)** | **1.001 (1.000, 1.001)** | 1.000 (0.999, 1.002) | | |
| Asthma | **1.009 (1.006, 1.013)** | **1.003 (1.002, 1.004)** | 0.999 (0.997, 1.001) | 0.998 (0.996, 1.000) | 1.000 (0.999, 1.001) | 1.003 (0.999, 1.007) | | |
| Cardiovascular | 1.000 (0.999, 1.001) | 1.000 (0.999, 1.000) | 1.000 (1.000, 1.000) | 1.000 (1.000, 1.001) | 1.000 (1.000, 1.001) | **1.001 (1.000, 1.001)** | **1.001 (1.000, 1.001)** | **1.001 (1.000, 1.002)** |
| Myocardial infarction | 0.999 (0.997, 1.001) | 1.000 (0.999, 1.001) | 1.001 (1.000, 1.001) | 1.001 (1.000, 1.002) | **1.002 (1.000, 1.003)** | **1.002 (1.001, 1.002)** | **1.002 (1.001, 1.003)** | **1.002 (1.000, 1.004)** |

Note: Bolded ORs are statistically significant at the p<0.05 level

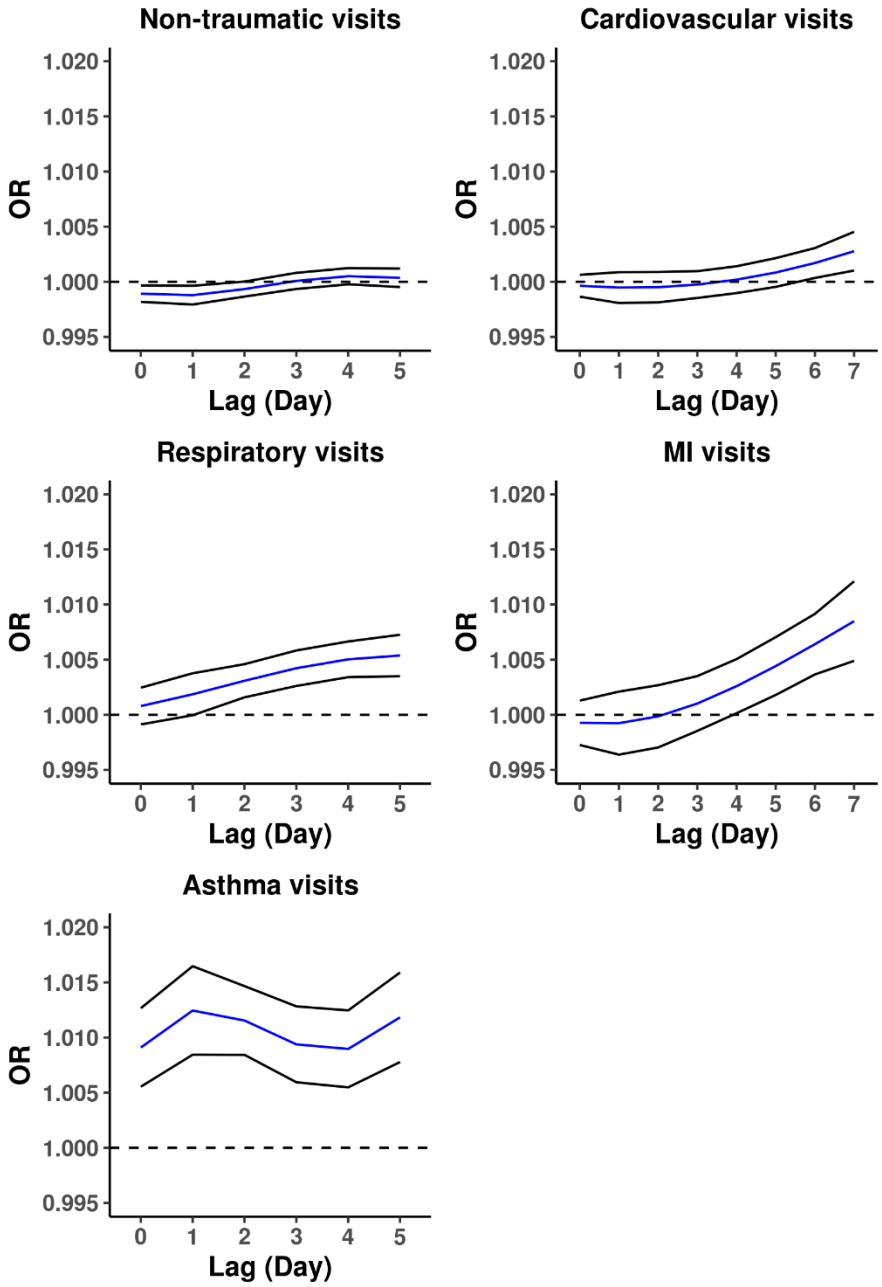


**Figure S2**. Cumulative odds ratios of outcome-specific ED visits for a 10 µg/m3 increase in wildfire smoke PM2.5. Exposure is modeled with a distributed lag model. ORs are plotted on the same scale for comparison across outcome groups.

**Table S13**. Cumulative exposure results: ORs and 95% CIs for ED visits cumulatively across all preceding lag days per 10 μg/$m^3$ increase in smoke-impacted $PM_{2.5}$, from a distributed lag model

| Outcome category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | **0.999 (0.998, 1.000)** | **0.999 (0.998, 1.000)** | 0.999 (0.999, 1.000) | 1.000 (0.999, 1.001) | **1.001 (1.000, 1.001)** | **1.000 (1.000, 1.001)** | | |
| Respiratory | 1.001 (0.999, 1.002) | **1.002 (1.000, 1.004)** | **1.003 (1.002, 1.005)** | **1.004 (1.003, 1.006)** | **1.005 (1.003, 1.007)** | **1.005 (1.004, 1.007)** | | |
| Asthma | **1.009 (1.006, 1.013)** | **1.012 (1.008, 1.016)** | **1.012 (1.008, 1.015)** | **1.009 (1.006, 1.013)** | **1.009 (1.005, 1.012)** | **1.012 (1.008, 1.016)** | | |
| Cardiovascular | 1.000 (0.999, 1.001) | 0.999 (0.998, 1.001) | 1.000 (0.998, 1.001) | 1.000 (0.999, 1.001) | 1.000 (0.999, 1.001) | **1.001 (1.000, 1.002)** | **1.002 (1.000, 1.003)** | **1.003 (1.001, 1.005)** |
| Myocardial infarction | 0.999 (0.997, 1.001) | 0.999 (0.996, 1.002) | 1.000 (0.997, 1.003) | 1.001 (0.999, 1.004) | **1.003 (1.000, 1.005)** | **1.004 (1.002, 1.007)** | **1.006 (1.004, 1.009)** | **1.008 (1.005, 1.012)** |

Note: Bolded ORs are statistically significant at the p<0.05 level

**Table S14**. Lag-specific exposure results from quasi-Poisson models: ORs and 95% CIs for ED visits per 10 μg/$m^3$ increase in smoke-impacted $PM_{2.5}$, from a distributed lag model

| Outcome category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 1.001 (0.999, 1.003) | 1.000 (1.000, 1.001) | 0.999 (0.998, 1.001) | 0.999 (0.998, 1.000) | 0.999 (0.998, 1.000) | 0.999 (0.997, 1.001) | | |
| Respiratory | 1.001 (0.998, 1.004) | **1.001 (1.000, 1.002)** | 1.001 (0.999, 1.003) | 1.001 (0.999, 1.003) | **1.001 (1.000, 1.002)** | 1.001 (0.998, 1.004) | | |
| Asthma | **1.008 (1.004, 1.012)** | **1.002 (1.001, 1.004)** | 0.998 (0.995, 1.002) | 0.997 (0.994, 1.000) | 0.999 (0.998, 1.001) | 1.002 (0.998, 1.007) | | |
| Cardiovascular | 0.999 (0.997, 1.000) | 0.999 (0.999, 1.000) | 1.000 (0.999, 1.001) | 1.000 (0.999, 1.001) | 1.001 (1.000, 1.002) | **1.001 (1.000, 1.001)** | 1.001 (1.000, 1.001) | 1.000 (0.999, 1.002) |
| Myocardial infarction | 0.998 (0.996, 1.000) | 0.999 (0.999, 1.000) | 1.001 (1.000, 1.001) | **1.002 (1.001, 1.003)** | **1.002 (1.001, 1.003)** | **1.002 (1.001, 1.003)** | **1.002 (1.001, 1.002)** | 1.001 (0.999, 1.003) |

Note: Bolded ORs are statistically significant at the p<0.05 level

**Table S15**. Cumulative exposure results from quasi-Poisson models: ORs and 95% CIs for ED visits cumulatively across all preceding lag days per 10 μg/m$^3$ increase in smoke-impacted $PM_{2.5}$, from a distributed lag model

| Outcome category | Lag 0 | Lag 1 | Lag 2 | Lag 3 | Lag 4 | Lag 5 | Lag 6 | Lag 7 |
|---|---|---|---|---|---|---|---|---|
| All non-traumatic | 1.001 (0.999, 1.003) | 1.001 (0.999, 1.004) | 1.001 (0.999, 1.003) | 1.000 (0.998, 1.002) | 0.999 (0.997, 1.001) | 0.999 (0.996, 1.001) | | |
| Respiratory | 1.001 (0.998, 1.004) | 1.002 (0.999, 1.005) | **1.003 (1.000, 1.006)** | **1.004 (1.001, 1.007)** | **1.005 (1.002, 1.008)** | **1.006 (1.003, 1.009)** | | |
| Asthma | **1.008 (1.004, 1.012)** | **1.010 (1.006, 1.015)** | **1.009 (1.005, 1.013)** | **1.006 (1.002, 1.011)** | **1.005 (1.001, 1.010)** | **1.008 (1.002, 1.013)** | | |
| Cardiovascular | 0.999 (0.997, 1.000) | 0.998 (0.996, 1.001) | 0.998 (0.996, 1.001) | 0.998 (0.996, 1.001) | 0.999 (0.997, 1.001) | 1.000 (0.997, 1.002) | 1.000 (0.998, 1.003) | 1.001 (0.998, 1.003) |
| Myocardial infarction | 0.998 (0.996, 1.000) | 0.997 (0.995, 1.000) | 0.998 (0.996, 1.001) | 1.000 (0.997, 1.002) | 1.002 (0.999, 1.004) | **1.003 (1.001, 1.006)** | **1.005 (1.002, 1.007)** | **1.006 (1.003, 1.009)** |

Note: Bolded ORs are statistically significant at the p<0.05 level

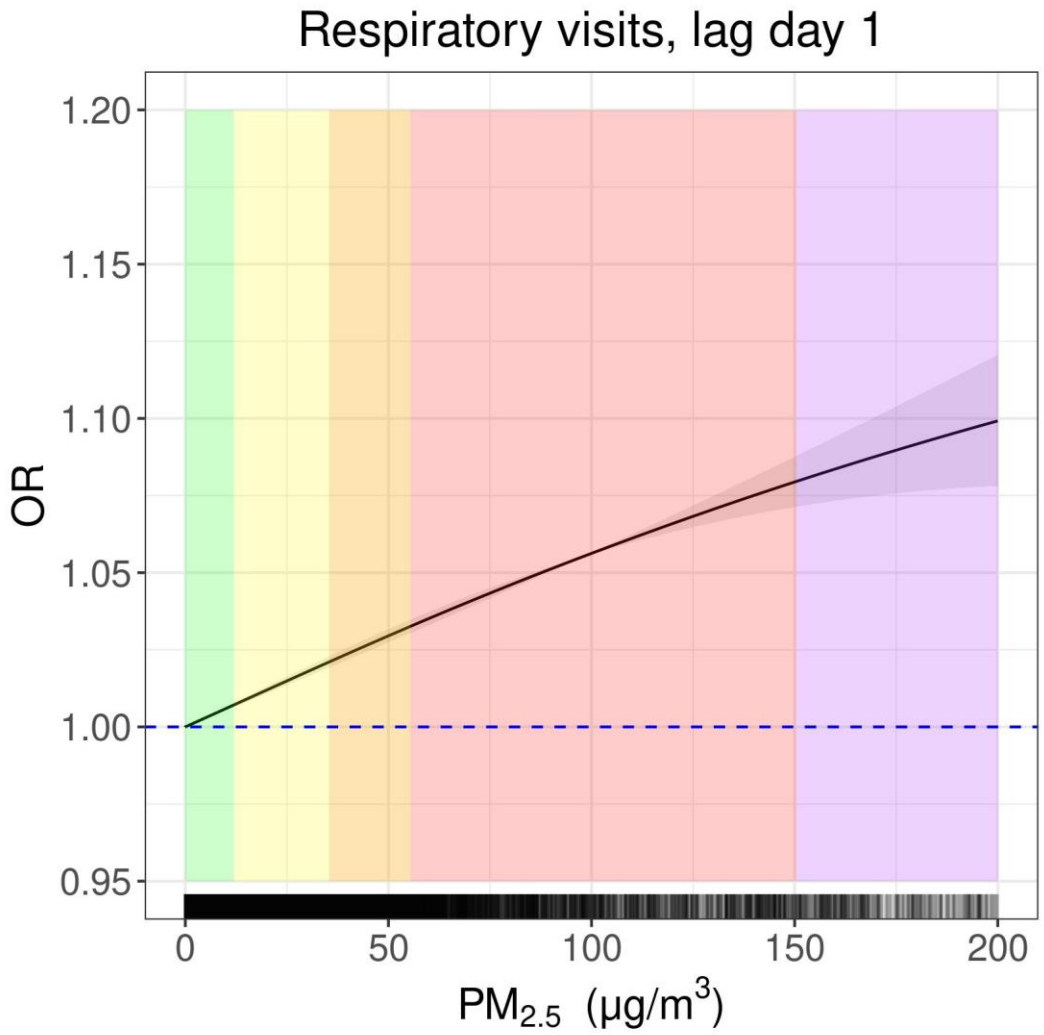


**Figure S3**. Exposure response curve for all-cause respiratory visits on lag day 1 as a function of total $PM_{2.5}$. Exposure is modeled as a natural cubic spline with one knot at 55 µg/m$^3$ (corresponding to the cut point between Unhealthy for Sensitive Groups and Unhealthy) with no distributed lag. Colors correspond to AQI groups (Green = Good (0-12 µg/m$^3$); Yellow = Moderate (12.1-35.4 µg/m$^3$); Orange = Unhealthy for Sensitive Groups (35.5-55.4 µg/m$^3$); Red = Unhealthy (55.5-150.4 µg/m$^3$); Purple = Very Unhealthy (150.5-250.4 µg/m$^3$)). Shading corresponds to the 95% CI. The rug plot on the bottom of the figure corresponds to the distribution of total $PM_{2.5}$ by zip code days in the study period.

# Chapter 5: Conclusion

This dissertation advances our scientific understanding of two key components of fine particulate matter, ultrafine particles and wildfire smoke, and increases our understanding of their myriad health effects. Specifically, this dissertation examines how to utilize on-road mobile monitoring data for traffic-related air pollution epidemiology studies (Aim 1), adds to the small evidence base on understanding the impacts of traffic-related air pollution and decline in physical functioning (Aim 2), and contributes to the growing body of literature on the health impacts of wildfire smoke exposure (Aim 3).

In Aim 1, we developed a method for lessening the contribution of on-road plumes in data collected in mobile monitoring campaigns for use in epidemiology and developed predictions using these on-road plume-adjusted measurements. We found these data pick up additional spatial information and variation compared to using a model with data collected at short-term stationary sites only. We recommend mobile monitoring campaigns correct mobile measurements prior to creating exposure predictions for epidemiology. However, further research is needed to validate our on-road plume adjustment algorithm in different campaigns and locations, and to examine how on-road plume adjusted mobile data impact epidemiologic inference. This will deepen our understanding of how to optimize the design of mobile monitoring campaigns for epidemiologic inference and increase our understanding of how to best use on-road data collected in mobile monitoring campaigns for epidemiology.

In Aim 2, we examined the longitudinal association between exposure to traffic-related air pollutants, including ultrafine particles, and a measure of physical functioning in a cohort of older adults. We observed an overall decline in physical functioning due to higher levels of TRAPs at baseline, and a faster decline in physical functioning over 5-years due to increased levels of BC, PNC in the range of 56-178nm, and non-time varying $PM_{2.5}$. Future research could examine additional exposure measures, including time varying UFP and BC measures, as data

are available, and additional measures of physical decline including measures of frailty to increase confidence in our findings. Additional research is needed in more racially and ethnically diverse cohorts to broaden our understanding of the population impact of exposure to TRAPs on physical decline and other aging-related outcomes.

In Aim 3, we observed an increase in respiratory-related emergency department visits due to wildfire smoke exposure and intensity shortly following initial exposure, and a lagged effect of wildfire smoke exposure and intensity on cardiovascular-related visits. We also observed a near linear dose-response when examining total $PM_{2.5}$ and respiratory emergency department visits one day following initial wildfire smoke exposure, indicating that the risk of negative health outcomes increases with higher concentrations of $PM_{2.5}$. Future research could examine the impact of co-exposure with extreme heat or with COVID-19; consider using the EPA's Air Quality Index (AQI) as the exposure metric; examine how health effects change after interventions; or examine the impact of long-term exposure to wildfire smoke in a cohort study setting. Additional research could also examine different exposure windows and time-averages for wildfire smoke exposure, including hourly exposures, maximum hourly average each day, cumulative exposure over several days, or cumulative exposure over an entire wildfire season.

Overall, this dissertation advances the science on several topics of traffic-related air pollution and wildfire smoke exposure and epidemiology and provides support for further inquiry. We support future research that targets policy and intervention-relevant questions, and research that supports the continued protection of human health from both traffic-related air pollutants and wildfire smoke.

## References


1. Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, et al. An Association Between Air Pollution and Mortality in Six U.S. Cities. N Engl J Med. 1993;329(24).

2. Koenig JQ. Health Effects of Particulate Matter. 2013.

3. U.S. EPA. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2019). Washington, DC;

4. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The Lancet. 2017;389(10082):1907–18.

5. Aguilera R, Corringham T, Gershunov A, Leibel S, Benmarhnia T. Fine Particles in Wildfire Smoke and Pediatric Respiratory Health in California. Pediatrics. 2021;147(4).

6. Ohlwein S, Kappeler R, Kutlar Joss M, Künzli N, Hoffmann B. Health effects of ultrafine particles: a systematic literature review update of epidemiological evidence. Int J Public Health. 2019;64(4):547–59.

7. Moreno-Ríos AL, Tejeda-Benítez LP, Bustillo-Lecompte CF. Sources, characteristics, toxicity, and control of ultrafine particles: An overview. Geoscience Frontiers [Internet]. 2022 Jan;13(1):101147. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1674987121000116

8. Zhu Y, Hinds WC, Kim S, Sioutas C. Concentration and Size Distribution of Ultrafine Particles Near a Major Highway. J Air Waste Manage Assoc. 2002;52(9):1032–42.

9. Austin E, Xiang J, Gould TR, Shirai JH, Yun S, Yost MG, et al. Distinct Ultrafine Particle Profiles Associated with Aircraft and Roadway Traffic. Environ Sci Technol. 2021;55(5):2847–58.

10. Hagler GSW, Baldauf RW, Thoma ED, Long TR, Snow RF, Kinsey JS, et al. Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants. Atmos Environ. 2009;43(6):1229–34.

11. Zhou Y, Levy JI. Factors influencing the spatial extent of mobile source air pollution impacts: A meta-analysis. BMC Public Health. 2007;7:1–11.

12. Apte JS, Messier KP, Gani S, Brauer M, Kirchstetter TW, Lunden MM, et al. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. Environ Sci Technol [Internet]. 2017 Jun 20;51(12):6999–7008. Available from: https://pubs.acs.org/doi/10.1021/acs.est.7b00891

13. Weichenthal S, Van Ryswyk K, Goldstein A, Shekarrizfard M, Hatzopoulou M. Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: A land use regression model. Environmental Pollution. 2016;208:241–8.

14. Sabaliauskas K, Jeong CH, Yao X, Reali C, Sun T, Evans GJ. Development of a land-use regression model for ultrafine particles in Toronto, Canada. Atmos Environ. 2015;110:84–92.

15. Hankey S, Marshall JD. Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, $PM_{2.5}$, Particle Size) Using Mobile Monitoring. Environ Sci Technol [Internet]. 2015 Aug 4;49(15):9194–202. Available from: https://pubs.acs.org/doi/10.1021/acs.est.5b01209

16. Van den Bossche J, Peters J, Verwaeren J, Botteldooren D, Theunis J, De Baets B. Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. Atmos Environ. 2015;105:148–61.

17. Hatzopoulou M, Valois MF, Levy I, Mihele C, Lu G, Bagg S, et al. Robustness of Land-Use Regression Models Developed from Mobile Air Pollutant Measurements. Environ Sci Technol. 2017;51(7):3938–47.

18. Alexeeff SE, Roy A, Shan J, Liu X, Messier K, Apte JS, et al. High-resolution mapping of traffic related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. Environ Health. 2018;17(1):1–13.

19. Southerland VA, Anenberg SC, Harris M, Apte J, Hystad P, van Donkelaar A, et al. Assessing the distribution of air pollution health risks within cities: A neighborhood-scale analysis leveraging high-resolution data sets in the bay area, california. Environ Health Perspect. 2021;129(3).

20. Kerckhoffs J, Hoek G, Messier KP, Brunekreef B, Meliefste K, Klompmaker JO, et al. Comparison of ultrafine particle and black carbon concentration predictions from a mobile and short-term stationary land-use regression model. Environ Sci Technol. 2016;50(23):12894–902.

21. Minet L, Liu R, Valois MF, Xu J, Weichenthal S, Hatzopoulou M. Development and Comparison of Air Pollution Exposure Surfaces Derived from On-Road Mobile Monitoring and Short-Term Stationary Sidewalk Measurements. Environ Sci Technol. 2018;52(6):3512–9.

22. Chambliss SE, Preble C V., Caubel JJ, Cados T, Messier KP, Alvarez RA, et al. Comparison of Mobile and Fixed-Site Black Carbon Measurements for High-Resolution Urban Pollution Mapping. Environ Sci Technol. 2020;54(13):7848–57.

23. Kerckhoffs J, Hoek G, Vlaanderen J, van Nunen E, Messier K, Brunekreef B, et al. Robustness of intra urban land-use regression models for ultrafine particles and black carbon based on mobile monitoring. Environ Res. 2017;159(August):500–8.

24. Blanco MN, Gassett A, Gould T, Doubleday A, Slager DL, Austin E, et al. Mobile monitoring of traffic-related air pollution for a prospective cohort study in the greater Seattle area. medRxiv. 2021;1–38.

25. Schmid O, Möller W, Semmler-Behnke M, A. Ferron G, Karg E, Lipka J, et al. Dosimetry and toxicology of inhaled ultrafine particles. Biomarkers. 2009 Jul 15;14(sup1):67–73.

26. Block ML, Calderón-Garcidueñas L. Air Pollution: Mechanisms of Neuroinflammation & CNS Disease. Trends Neurosci. 2009;32(9):506–19.

27. CDC. Prevalence and Most Common Causes of Disability Among Adults --- United States, 2005. 2009.

28. Weuve J, Kaufman JD, Szpiro AA, Curl C, Puett RC, Beck T, et al. Exposure to traffic-related air pollution in relation to progression in physical disability among older adults. Environ Health Perspect. 2016;124(7):1000–8.

29. De Zwart F, Brunekreef B, Timmermans E, Deeg D, Gehring U. Air pollution and performance-based physical functioning in dutch older adults. Environ Health Perspect. 2018;126(1):1–9.

30. Lin H, Guo Y, Zheng Y, Zhao X, Cao Z, Rigdon SE, et al. Exposure to ambient PM 2 . 5 associated with overall and domain-specific disability among adults in six low- and middle-income countries. Environ Int. 2017;104(April):69–75.

31. Wen M, Gu D. Air Pollution Shortens Life Expectancy and Health Expectancy for Older Adults : The Case of China. 2012;

32. Wang L, Larson EB, Bowen JD, Van Belle G. Performance-based physical function and future dementia in older people. Arch Intern Med. 2006;166(10):1115–20.

33. Wilkins CH, Roe CM, Morris JC, Galvin JE. Mild Physical Impairment Predicts Future Diagnosis of Dementia of the Alzheimer Type. J Am Geriatr Soc. 2013;61(7).

34. Seeman TE, Merkin SS, Crimmins EM, Karlamangla AS. Disability Trends Among Older Americans: National Health and Nutrition Examination Surveys, 1988–1994 and 1999–2004. Am J Public Health [Internet]. 2010 Jan;100(1):100–7. Available from: https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2008.157388

35. Guralnik JM;, Patel K, Ferrucci L. Assessing Functional Status and Disability in Epidemiologic Studies. In: Newman AB;, Cauley JA, editors. The Epidemiology of Aging. Springer; 2012. p. 91–117.

36. Delfino RJ, Brummel S, Wu J, Stern H, Ostro B, Lipsett M, et al. The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. Occup Environ Med. 2009;66(3):1–24.

37. Gan RW, Ford B, Lassman W, Pfister G, Vaidyanathan A, Fischer E, et al. Comparison of wildfire smoke estimation methods and associations with cardiopulmonary-related hospital admissions. Vol. 1, GeoHealth. 2017. p. 122–36.

38. Martin KL, Hanigan IC, Morgan GG, Henderson SB, Johnston FH. Air pollution from bushfires and their association with hospital admissions in Sydney, Newcastle and Wollongong, Australia 1994-2007. Aust N Z J Public Health. 2013;37(3):238–43.

39. Alman BL, Pfister G, Hao H, Stowell J, Hu X, Liu Y, et al. The association of wildfire smoke with respiratory and cardiovascular emergency department visits in Colorado in 2012: A case crossover study. Environ Health. 2016;15(1):1–9.

40. Wettstein ZS, Hoshiko S, Fahimi J, Harrison RJ, Cascio WE, Rappold AG. Cardiovascular and Cerebrovascular Emergency Department Visits Associated With Wildfire Smoke Exposure in California in 2015. J Am Heart Assoc. 2018;7(8):e007492.

41. Kiser D, Metcalf WJ, Elhanan G, Schnieder B, Schlauch K, Joros A, et al. Particulate matter and emergency visits for asthma: A time-series study of their association in the presence and absence of wildfire smoke in Reno, Nevada, 2013-2018. Environ Health. 2020;19(1):1–12.

42. Rappold AG, Stone SL, Cascio WE, Neas LM, Kilaru VJ, Carraway MS, et al. Peat Bog Wildfire Smoke Exposure in Rural North Carolina Is Associated with Cardiopulmonary Emergency Department Visits Assessed through Syndromic Surveillance. Enviromental Health Perpectives. 2011;119(10):1415–20.

43. Gan RW, Ford B, Lassman W, Pfister G, Vaidyanathan A, Fischer E, et al. Comparisons of wildfire smoke estimation methods and associations with cardiopulmonary-related hospital admissions. Geohealth. 2017;1(3):122–36.

44. Martin KL, Hanigan IC, Morgan GG, Henderson SB, Johnston FH. Air pollution from bushfires and their association with hospital admissions in Sydney, Newcastle and Wollongong, Australia 1994-2007. Vol. 37, Australian and New Zealand Journal of Public Health. 2013. p. 238–43.

45. Alman BL, Pfister G, Hao H, Stowell J, Hu X, Liu Y, et al. The association of wildfire smoke with respiratory and cardiovascular emergency department visits in Colorado in 2012: a case crossover study. Environ Health. 2016 Jun;15(1):64.

46. Wettstein ZS, Hoshiko S, Fahimi J, Harrison RJ, Cascio WE, Rappold AG. Cardiovascular and Cerebrovascular Emergency Department Visits Associated With Wildfire Smoke Exposure in California in 2015. J Am Heart Assoc. 2018 Apr;7(8).

47. Rappold AG, Stone SL, Cascio WE, Neas LM, Kilaru VJ, Carraway MS, et al. Peat Bog Wildfire Smoke Exposure in Rural North Carolina is associated with Cardiopulmonary Emergency Department Visits assessed through Syndromic Surveillance. Enviromental Health Perpectives. 2011;119(10):1415–20.

48. Thelen B, French NH, Koziol BW, Billmire M, Owen RC, Johnson J, et al. Modeling acute respiratory illness during the 2007 San Diego wildland fires using a coupled emissions-transport system and generalized additive modeling. Environmental Health. 2013 Dec 5;12(1):94.

49. Washington State Department of Health. Rapid Health Information Network (RHINO) [Internet]. Available from: https://www.doh.wa.gov/ForPublicHealthandHealthcareProviders/HealthcareProfessionsandFacilities/PublicHealthMeaningfulUse/RHINO

50. Doubleday A, Schulte J, Sheppard L, Kadlec M, Dhammapala R, Fox J, et al. Mortality associated with wildfire smoke exposure in Washington state, 2006-2017: a case-crossover study. Environ Health. 2020 Jan;19(1):4.

51. Hoekman SK, Welstand JS. Vehicle Emissions and Air Quality: The Early Years (1940s–1950s). Atmosphere (Basel) [Internet]. 2021 Oct 16;12(10):1354. Available from: https://www.mdpi.com/2073-4433/12/10/1354

52. Wen Y, Wang H, Larson T, Kelp M, Zhang S, Wu Y, et al. On-highway vehicle emission factors, and spatial patterns, based on mobile monitoring and absolute principal component score. Science of the Total Environment. 2019;676:242–51.

53. Larson T, Gould T, Riley EA, Austin E, Fintzi J, Sheppard L, et al. Ambient air quality measurements from a continuously moving mobile platform: Estimation of area-wide, fuel-based, mobile source emission factors using absolute principal component scores. Atmos Environ. 2017;152:201–11.

54. Blanco MN, Doubleday A, Austin E, Marshall JD, Seto E, Larson T V., et al. Design and evaluation of short-term monitoring campaigns for long-term air pollution exposure assessment. J Expo Sci Environ Epidemiol [Internet]. 2022 Aug 31; Available from: https://www.nature.com/articles/s41370-022-00470-5

55. Blanco MN, Bi J, Austin E, Larson T V., Marshall JD, Sheppard L. Impact of Mobile Monitoring Network Design on Air Pollution Exposure Assessment Models. Environ Sci Technol [Internet]. 2023 Jan 10;57(1):440–50. Available from: https://pubs.acs.org/doi/abs/10.1021/acs.est.2c05338

56. Blanco MN, Gassett A, Gould T, Doubleday A, Slager DL, Austin E, et al. Characterization of Annual Average Traffic-Related Air Pollution Concentrations in the Greater Seattle Area from a Year-Long Mobile Monitoring Campaign. Environ Sci Technol [Internet]. 2022 Aug 16;56(16):11460–72. Available from: https://pubs.acs.org/doi/10.1021/acs.est.2c01077

57. Kerckhoffs J, Hoek G, Gehring U, Vermeulen R. Modelling nationwide spatial variation of ultrafine particles based on mobile monitoring. Environ Int. 2021;154(2):106569.

58. Keller JP, Olives C, Kim SY, Sheppard L, Sampson PD, Szpiro AA, et al. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. Environ Health Perspect. 2015;123(4):301–9.

59. Bi J, Carmona N, Blanco MN, Gassett AJ, Seto E, Szpiro AA, et al. Publicly available low-cost sensor measurements for PM2.5 exposure modeling: Guidance for monitor deployment and data selection. Environ Int [Internet]. 2022 Jan;158:106897. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0160412021005225

60. Kwon HS, Ryu MH, Carlsten C. Ultrafine particles: unique physicochemical properties relevant to health and disease. Exp Mol Med. 2020;52(3):318–28.

61. Thurston GD, Spengler JD. A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. Atmospheric Environment (1967). 1985;19(1):9–25.

62. Kelp M, Gould T, Austin E, Marshall JD, Yost M, Simpson C, et al. Sensitivity analysis of area-wide, mobile source emission factors to high-emitter vehicles in Los Angeles. Atmos

Environ [Internet]. 2020;223(November 2019):117212. Available from: https://doi.org/10.1016/j.atmosenv.2019.117212

63. Tessum MW, Larson T, Gould TR, Simpson CD, Yost MG, Vedal S. Mobile and Fixed-Site Measurements To Identify Spatial Distributions of Traffic-Related Pollution Sources in Los Angeles. Environ Sci Technol. 2018;52(5):2844–53.

64. Revelle W. Package ‘ psych .’ 2022;

65. Gilbert P, Jennrich R. Package ‘ GPArotation .’ 2022;

66. R Development Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2018. Available from: http://www.r-project.org

67. Appendix E: Census Feature Class Codes (CFCC) [Internet]. TIGER/Line Files. 1992 [cited 2022 Sep 22]. Available from: https://www2.census.gov/geo/tiger/TIGER1992/Documentation/APPENDXE.txt

68. University of Washington Department of Environmental and Occupational Health Sciences. Mobile ObserVations of Ultrafine Particles: The MOV-UP study report. Seattle; 2019.

69. Brantley HL, Hagler GSW, Kimbrough ES, Williams RW, Mukerjee S, Neas LM. Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. Atmos Meas Tech. 2014;7(7):2169–83.

70. Drewnick F, Böttger T, von der Weiden-Reinmüller SL, Zorn SR, Klimach T, Schneider J, et al. Design of a mobile aerosol research laboratory and data processing tools for effective stationary and mobile field measurements. Atmos Meas Tech. 2012;5:443–1457.

71. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Measurement Error in Nonlinear Models: A Modern Perspective. Second Edi. London: Chapman and Hall; 2006.

72. Szpiro AA, Paciorek CJ. Measurement error in two-stage analyses, with application to air pollution epidemiology. Environmetrics. 2013;24(8):501–17.

73. Zhu Y, Yu N, Kuhn T, Hinds WC. Field Comparison of P-Trak and Condensation Particle Counters. Aerosol Science and Technology. 2006 Jul;40(6):422–30.

74. MESA Air. Data Organization and Operating Procedures (DOOP) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MEAS Air) and Associated Studies. 2019;1–77.

75. Carroll RJ. Measurement error in nonlinea models: a modern perspective. 2nd ed. Boca Raton, FL; 2006.

76. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. Journals of Gerontology. 1994;49(2).

77. Wang H, Liu H, Guo F, Li J, Li P, Guan T, et al. Association Between Ambient Fine Particulate Matter and Physical Functioning in Middle-Aged and Older Chinese Adults: A Nationwide Longitudinal Study. The Journals of Gerontology: Series A. 2022 May 5;77(5):986–93.

78. Freiberger E, de Vreede P, Schoene D, Rydwik E, Mueller V, Frandin K, et al. Performance-based physical function in older community-dwelling persons: a systematic review of instruments. Age Ageing. 2012;41(6):712–21.

79. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir G V, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci. 2000;55(4):M221-31.

80. Kukull WA, Higdon R, Bowen JD, McCormick WC, Teri L, Schellenberg GD, et al. Dementia and Alzheimer Disease Incidence: A Prospective Cohort Study. Arch Neurol. 2002;59(11):1737–46.

81. Wang R, Henderson SB, Sbihi H, Allen RW, Brauer M. Temporal stability of land use regression models for traffic-related air pollution. Atmos Environ. 2013 Jan;64:312–9.

82. Eeftens M, Beelen R, Fischer P, Brunekreef B, Meliefste K, Hoek G. Stability of measured and modelled spatial contrasts in NO2 over time. Occup Environ Med. 2011 Oct 1;68(10):765–70.

83. Cesaroni G, Porta D, Badaloni C, Stafoggia M, Eeftens M, Meliefste K, et al. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. Environmental Health. 2012 Dec 18;11(1):48.

84. Shaffer RM, Blanco MN, Li G, Adar SD, Carone M, Szpiro AA, et al. Fine Particulate Matter and Dementia Incidence in the Adult Changes in Thought Study. Environ Health Perspect. 2021 Aug;129(8):087001.

85. Kaufman JD, Adar SD, Barr RG, Budoff M, Burke GL, Curl CL, et al. Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the Multi-Ethnic Study of Atherosclerosis and Air Pollution): a longitudinal cohort study. The Lancet [Internet]. 2016 Aug;388(10045):696–704. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673616003780

86. Christine PJ, Auchincloss AH, Bertoni AG, Carnethon MR, Sánchez BN, Moore K, et al. Longitudinal associations between neighborhood physical and social environments and incident type 2 diabetes mellitus: The Multi-Ethnic Study of Atherosclerosis (MESA). JAMA Intern Med. 2015 Aug 1;175(8):1311–20.

87. Weuve J, Tchetgen ET, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, et al. Accounting for bias due to selective attrition: The example of smoking and cognitive decline. Epidemiology. 2012;23(1):119–28.

88. Vu T V., Delgado-Saborit JM, Harrison RM. Review: Particle number size distributions from seven major sources and implications for source apportionment studies. Atmos Environ. 2015 Dec;122:114–32.

89. Hopke PK, Feng Y, Dai Q. Source apportionment of particle number concentrations: A global review. Science of The Total Environment. 2022 May;819:153104.

90. Hajat A, Diez-Roux A V., Adar SD, Auchincloss AH, Lovasi GS, O'Neill MS, et al. Air Pollution and Individual and Neighborhood Socioeconomic Status: Evidence from the Multi-Ethnic Study of Atherosclerosis (MESA). Environ Health Perspect [Internet]. 2013 Nov;121(11–12):1325–33. Available from: https://ehp.niehs.nih.gov/doi/10.1289/ehp.1206337

91. Hajat A, MacLehose RF, Rosofsky A, Walker KD, Clougherty JE. Confounding by Socioeconomic Status in Epidemiological Studies of Air Pollution and Health: Challenges and Opportunities. Environ Health Perspect. 2021 Jun;129(6):065001.

92. Clark TG, Altman DG, Stavola BL De. Quantification of the completeness of follow-up. The Lancet. 2002 Apr;359(9314):1309–10.

93. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. J Gerontol A Biol Sci Med Sci [Internet]. 2001 Mar 1;56(3):M146–57. Available from: https://academic.oup.com/biomedgerontology/article-lookup/doi/10.1093/gerona/56.3.M146

94. Jung HW, Baek JY, Jang IY, Guralnik JM, Rockwood K, Lee E, et al. Short Physical Performance Battery as a Crosswalk Between Frailty Phenotype and Deficit Accumulation Frailty Index. Newman AB, editor. The Journals of Gerontology: Series A [Internet]. 2021 Nov 15;76(12):2249–55. Available from: https://academic.oup.com/biomedgerontology/article/76/12/2249/6199834

95. U.S. EPA. Timeline of Major Accomplishments in Transportation, Air Pollution, and Climate Change [Internet]. 2022 [cited 2023 Apr 5]. Available from: https://www.epa.gov/transportation-air-pollution-and-climate-change/timeline-major-accomplishments-transportation-air

96. Abatzoglou JT, Williams AP. Impact of anthropogenic climate change on wildfire across western US forests. PNAS. 2016;113(42):11770–1175.

97. Radeloff VC, Helmers DP, Anu Kramer H, Mockrin MH, Alexandre PM, Bar-Massada A, et al. Rapid growth of the US wildland-urban interface raises wildfire risk. Proc Natl Acad Sci U S A. 2018;115(13):3314–9.

98. Jaffe DA, O'Neill SM, Larkin NK, Holder AL, Peterson DL, Halofsky JE, et al. Wildfire and prescribed burning impacts on air quality in the United States. J Air Waste Manage Assoc. 2020;70(6):583–615.

99. Parisien MA, Barber QE, Hirsch KG, Stockdale CA, Erni S, Wang X, et al. Fire deficit increases wildfire risk for many communities in the Canadian boreal forest. Nat Commun. 2020;11(1).

100. Liu JC, Mickley LJ, Sulprizio MP, Dominici F, Yue X, Ebisu K, et al. Particulate air pollution from wildfires in the Western US under climate change. Clim Change. 2016;138(3–4).

101. Ford B, Val Martin M, Zelasky SE, Fischer E V., Anenberg SC, Heald CL, et al. Future Fire Impacts on Smoke Concentrations, Visibility, and Health in the Contiguous United States. Geohealth. 2018;2(8):229–47.

102. Childs ML, Li J, Wen J, Heft-neal S, Driscoll A, Wang S, et al. Daily Local-Level Estimates of Ambient Wildfire Smoke PM 2.5 for the Contiguous US. 2022;

103. Wegesser TC, Pinkerton KE, Last JA. California wildfires of 2008: Coarse and fine particulate matter toxicity. Environ Health Perspect. 2009;117(6):893–7.

104. Kim YH, Warren SH, Krantz QT, King C, Jaskot R, Preston WT, et al. Mutagenicity and lung toxicity of smoldering vs. Flaming emissions from various biomass fuels: Implications for health effects from wildland fires. Environ Health Perspect. 2018;126(1):1–14.

105. Malig BJ, Fairley D, Pearson D, Wu X, Ebisu K, Basu R. Examining fine particulate matter and cause-specific morbidity during the 2017 North San Francisco Bay wildfires. Science of the Total Environment. 2021;787.

106. Hahn MB, Kuiper G, O'Dell K, Fischer E V., Magzamen S. Wildfire Smoke Is Associated With an Increased Risk of Cardiorespiratory Emergency Department Visits in Alaska. Geohealth. 2021;5(5):1–15.

107. Henderson SB, Brauer M, Kennedy S, MacNab Y. Three Measures of Forest Fire Smoke Exposure and Their Association with Respiratory and Cardiovascular Physician Visits and Hospital Admissions. Epidemiology. 2009;20(9):S82.

108. Yao J, Brauer M, Wei J, McGrail KM, Johnston FH, Henderson SB. Sub-daily exposure to fine particulate matter and ambulance dispatches during wildfire seasons: A case-crossover study in British Columbia, Canada. Environ Health Perspect. 2020;128(6):1–10.

109. Hospital emergency room patient care information—Data collection, maintenance, analysis, and dissemination—Rules. 2017.

110. Lipner EM, O'Dell K, Brey SJ, Ford B, Pierce JR, Fischer E V., et al. The Associations Between Clinical Respiratory Outcomes and Ambient Wildfire Smoke Exposure Among Pediatric Asthma Patients at National Jewish Health, 2012–2015. Geohealth. 2019;3(6):146–59.

111. Magzamen S, Gan RW, Liu J, Dell KO, Ford B, Berg K, et al. Differential Cardiopulmonary Health Impacts of Local and Long-Range Transport of Wildfire Smoke. Geohealth. 2021;1–18.

112. National Oceanic and Atmospheric Administration. Hazard Mapping System Fire and Smoke Product [Internet]. 2018. Available from: https://www.ospo.noaa.gov/Products/land/hms.html

113. Masterton J, Richardson F. Humidex: a method of quantifying human discomfort due to excessive heat and humidity. Downsview, Ontario: Environment Canada, Atmospheric Environment; 1979.

114. Deflorio-Barker S, Crooks J, Reyes J, Rappold AG. Cardiopulmonary effects of fine particulate matter exposure among older adults, during wildfire and non-wildfire periods, in the United States 2008–2010. Environ Health Perspect. 2019;127(3).

115. Gan RW, Liu J, Ford B, O'Dell K, Vaidyanathan A, Wilson A, et al. The association between wildfire smoke exposure and asthma-specific medical care utilization in Oregon during the 2013 wildfire season. J Expo Sci Environ Epidemiol. 2020;618–28.

116. Rice K, Higgins JPT, Lumley T. A re-evaluation of fixed effect(s) meta-analysis. J R Stat Soc Ser A Stat Soc. 2018;181(1):205–27.

117. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. Stat Med. 2010;29(21):2224–34.

118. Abdo M, Ward I, O'dell K, Ford B, Pierce JR, Fischer E V., et al. Impact of wildfire smoke on adverse pregnancy outcomes in Colorado, 2007–2015. Int J Environ Res Public Health. 2019;16(19).

119. Stowell JD, Geng G, Saikawa E, Chang HH, Fu J, Yang CE, et al. Associations of wildfire smoke PM2.5 exposure with cardiorespiratory events in Colorado 2011–2014. Environ Int. 2019;133(August):105151.

120. Mustafić H, Jabre P, Caussin C, Murad MH, Escolano S, Tafflet M, et al. Main Air Pollutants and Myocardial Infarction. JAMA. 2012 Feb 15;307(7):713.

121. Farhadi Z, Abulghasem Gorgi H, Shabaninejad H, Aghajani Delavar M, Torani S. Association between PM2.5 and risk of hospitalization for myocardial infarction: A systematic review and a meta-analysis. BMC Public Health. 2020 Mar 12;20(1).

122. Luo C, Zhu X, Yao C, Hou L, Zhang J, Cao J, et al. Short-term exposure to particulate air pollution and risk of myocardial infarction: a systematic review and meta-analysis. Environmental Science and Pollution Research. 2015 Oct 1;22(19):14651–62.

123. Geng J, Liu H, Ge P, Hu T, Zhang Y, Zhang X, et al. PM2.5 promotes plaque vulnerability at different stages of atherosclerosis and the formation of foam cells via TLR4/MyD88/NFκB pathway. Ecotoxicol Environ Saf. 2019 Jul 30;176:76–84.

124. Kiss P, Carcel C, Hockham C, Peters SAE. The impact of the COVID-19 pandemic on the care and management of patients with acute cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes. 2021 Jan 25;7(1):18–27.

125. Giannouchos T v., Brooks JM, Andreyeva E, Ukert B. Frequency and factors associated with foregone and delayed medical care due to COVID-19 among nonelderly US adults from August to December 2020. J Eval Clin Pract. 2022 Feb 15;28(1):33–42.

126. Casey JA, Kioumourtzoglou M anna, Elser H, Walker D, Taylor S, Adams S, et al. Wildfire particulate matter in Shasta County, California and respiratory and circulatory disease-related emergency department visits and mortality, 2013 – 2018. Environmental Epidemiology. 2020;

127. Henderson SB, Morrison KT, Mclean KE, Ding Y, Yao J, Shaddick G, et al. Staying Ahead of the Epidemiologic Curve : Evaluation of the British Columbia Asthma Prediction

System ( BCAPS ) During the Unprecedented 2018 Wildfire Season. 2021;9(March):1–13.

128. O'Dell K, Bilsback K, Ford B, Martenies SE, Magzamen S, Fischer E V., et al. Estimated Mortality and Morbidity Attributable to Smoke Plumes in the United States : Not Just a Western US Problem GeoHealth. Geohealth. 2021;1–17.

129. 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;

130. Schulte J. 2018 Ambient Air Monitoring Network Plan. 2018.

131. Lassman W, Ford B, Gan RW, Pfister G, Magzamen S, Fischer E V., et al. Spatial and temporal estimates of population exposure to wildfire smoke during the Washington state 2012 wildfire season using blended model, satellite, and in situ data. Geohealth. 2017;1(3):106–21.

132. Washington State University. AIRPACT [Internet]. 2019. Available from: http://lar.wsu.edu/airpact/

133. Pilz J, Spöck G. Why do we need and how should we implement Bayesian kriging methods. Stochastic Environmental Research and Risk Assessment. 2008;22(5):621–32.

134. Management O of F. Population & demographics [Internet]. 2017. Available from: https://www.ofm.wa.gov/washington-data-research/population-demographics