ORIGINAL RESEARCH

Associations Between Long-Term Fine Particulate Matter Exposure and Mortality in Heart Failure Patients

Cavin K. Ward-Caviness, PhD; Anne M. Weaver, PhD; Matthew Buranosky, BS; Emily R. Pfaff, MS; Lucas M. Neas, PhD; Robert B. Devlin, PhD; Joel Schwartz, PhD; Qian Di, PhD; Wayne E. Cascio, MD; David Diaz-Sanchez, MD

***Background*—**Environmental health risks for individuals with heart failure (HF) have been inadequately studied, as these individuals are not well represented in traditional cohort studies. To address this we studied associations between long-term air pollution exposure and mortality in HF patients.

***Methods and Results*—**The study population was a hospital-based cohort of individuals diagnosed with HF between July 1, 2004 and December 31, 2016 compiled using electronic health records. Individuals were followed from 1 year after initial diagnosis until death or the end of the observation period (December 31, 2016). We used Cox proportional hazards models to evaluate the association of annual average fine particulate matter ($PM_{2.5}$) exposure at the time of initial HF diagnosis with all-cause mortality, adjusted for age, race, sex, distance to the nearest air pollution monitor, and socioeconomic status indicators. Among 23 302 HF patients, a 1 $\mu g/m^3$ increase in annual average $PM_{2.5}$ was associated with an elevated risk of all-cause mortality (hazard ratio 1.13; 95% CI, 1.10–1.15). As compared with people with exposures below the current national $PM_{2.5}$ exposure standard (12 $\mu g/m^3$), those with elevated exposures experienced 0.84 (95% CI, 0.73–0.95) years of life lost over a 5-year period, an observation that persisted even for those residing in areas with $PM_{2.5}$ concentrations below current standards.

***Conclusions*—**Residential exposure to elevated concentrations of $PM_{2.5}$ is a significant mortality risk factor for HF patients. Elevated $PM_{2.5}$ exposures result in substantial years of life lost even at concentrations below current national standards. **(*J Am Heart Assoc*. 2020;9:e012517. DOI: 10.1161/JAHA.119.012517.)**

Key Words: air pollution • electronic health record • heart failure • mortality • $PM_{2.5}$

Long-term air pollution exposure remains one of the most significant risk factors for mortality worldwide, a trend that may be increasing despite improvements in air quality in many nations.[1] Reasons for this increasing mortality burden from poor air quality include advancing age and the increasing prevalence of chronic disease conditions, both of which can increase sensitivity to poor air quality. Heart failure (HF) is one of the most severe forms of cardiovascular disease, with HF patients having a median life expectancy <5 years.[2,3] Improvements in care, combined with an aging population and stable incidence, result in an increasing prevalence of HF in the United States, with a projected 8 million adults to have HF in 2030 (46% increase over 2012).[4] Still, recent gains in survival rates for HF patients lag behind those for many other common chronic diseases.[5]

Individuals with HF have substantial environmental health risks,[6-10] which are generally elevated as compared with the general population or with individuals with less severe heart disease.[10-12] HF-related deaths may account for nearly 1 out of every 3 deaths due to airborne particulate matter exposure,[11] and elevated particulate matter air pollution (primarily particles <2.5 μm in diameter [$PM_{2.5}$]) contributes to hundreds of millions of dollars in HF-related morbidity and mortality costs and several thousand HF-related hospital admissions.[9]

Most environmental health studies of individuals with HF have focused on short-term exposures, and, to date, no study has examined associations with long-term air pollution

From the Center for Public Health and Environmental Assessment, US Environmental Protection Agency, Chapel Hill, NC (C.K.W.-C., A.M.W., L.M.N., R.B.D., W.E.C., D.D.-S.); NC Translational and Clinical Sciences Institute, University of North Carolina–Chapel Hill, Chapel Hill, NC (E.R.P.); Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA (J.S.); Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA (J.S.); Research Center for Public Health, School of Medicine, Tsinghua University, Beijing, China (Q.D.)

Accompanying Data S1, Tables S1 through S5, and Figures S1, S2 are available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.119.012517

Correspondence to: Cavin K. Ward-Caviness PhD, US EPA, Public Health & Integrated Toxicology Division, Human Studies Facility, 104 Mason Farm Rd, Chapel Hill, NC 27514. E-mail: ward-caviness.cavin@epa.gov

Received May 6, 2019; accepted January 8, 2020.

Clinical Perspective

What Is New?

- The environmental health risks of heart failure have been understudied despite evidence that these patients are at increased risk relative to the general population.
- We observe that among individuals with heart failure, long-term exposure to particulate matter air pollution at their primary residence conferred a significant mortality risk and is associated with substantial years of life lost.
- Mortality risk differed slightly by heart failure subtype (systolic versus diastolic) in some models.

What Are the Clinical Implications?

- The robust association between increasing mortality and increasing concentrations of ambient particulate matter raises the question of whether clinical interventions to lower residential exposures to particulate matter would result in improved clinical outcomes.

exposures exclusively in the HF patient community.[13] Existing knowledge on the long-term exposure effects is currently extrapolated from studies of the general population that conduct secondary stratified (ie, subgroup) analyses on HF deaths.[12] This results in a lack of knowledge of health risks based on exposures at the time of HF diagnosis and on the modification of these health risks by factors such as systolic versus diastolic HF, preexisting conditions such as type 2 diabetes mellitus or hypertension, and demographic factors such as age and sex. The Environmental Protection Agency Clinical and Archived Records Research for Environmental Studies (EPA CARES) research program is a collaborative effort between the EPA and the North Carolina Translational and Clinical Sciences Institute at the University of North Carolina at Chapel Hill (UNC) to utilize electronic health records (EHRs) to examine health effects in the hospital-going population of North Carolina. Using EPA CARES, we have examined the relationship between long-term exposure to $PM_{2.5}$ and mortality among individuals with HF.

Methods

The data that support the findings of this study are available from the corresponding author on reasonable request and appropriate approval from the UNC–Chapel Hill Institutional Review Board.

Study Cohort

All participants in this study were sourced from the Carolina Data Warehouse for Health, the enterprise data warehouse for the UNC Health Care System. The Carolina Data Warehouse for Health contains clinical and administrative data from both UNC Health Care System's Epic EHR (2014–present) and its homegrown legacy EHR, WebCIS (2004–2014). This study was approved by UNC–Chapel Hill Institutional Review Board (IRB 17-0150). As a secondary analysis of already collected data, written, informed consent was waived.

The study cohort was composed of individuals with a recorded diagnosis of HF between July 1, 2004 and December 31, 2016. HF was defined according to the *International Classification of Diseases, Ninth Revision* (*ICD-9*) codes 428.x and the *International Classification of Diseases, Tenth Revision* (*ICD-10*) codes I50.x. We also identified whether individuals were diagnosed with diastolic HF (*ICD-10* I50.3, I50.81; *ICD-9* 428.3) versus systolic HF (*ICD-10* I50.1, I50.2; *ICD-9* 428.1, 428.2). Individuals were then linked to demographics, address history, hospital and state death records, and disease diagnosis history. Records with inconsistent birth, death, or visit dates were removed. Addresses were cleaned for spelling mistakes as well as addresses that could not be geocoded. We considered an address to be successfully geocoded at the zip code level or better and successfully geocoded 99.9% of all addresses. The full geocoding procedure is given in Data S1 (Expanded Methods—Geocoding Procedure). We assigned annual average $PM_{2.5}$ values based on the nearest ground-based monitor to the listed primary residential address at the time of the heart failure diagnosis, defined as the first visit where heart failure was listed in the medical history, as well as based on an ensemble model of daily $PM_{2.5}$ concentrations at a 1×1-km resolution, which were then averaged to create an annual average.[14] We restricted the study to participants who resided in North Carolina. To remove HF cases likely to be due to congenital disease, we removed any participants diagnosed with HF at age 20 or younger.

Air Pollution Data

Daily $PM_{2.5}$ values were obtained from the EPA National Ambient Air Quality Standards (NAAQS) ground-based monitoring network. Data were obtained for the period July 1, 2003 through December 31, 2016. After geocoding the study cohort and determining the nearest monitor, the annual average was defined as the average daily $PM_{2.5}$ exposure over the 365 days preceding the examination based on the date and primary address at the time of initial heart failure, based on the patient's EHR. Monitors were required to have a minimum of 100 days measured for analysis. We also obtained daily air quality data from an ensemble model incorporating satellite measurements, land use regression variables, and several other predictors. The model uses 3 machine learning algorithms to integrate predictors and estimate daily $PM_{2.5}$ exposure at 1×1-km grids for the

continental United States. The model had an r^2 of 0.89 for the Middle Atlantic region of the United States (which would include North Carolina) for the years 2000-2015.[14] Data were obtained for the years 2003-2016 for the state of North Carolina.

In addition to the annual average air pollution, we also examined associations with an indicator variable for residence within areas with annual averages below the current NAAQS for annual average $PM_{2.5}$ of (12 $\mu g/m^3$). Because the implementation of national regulations has caused air quality to trend downwards over the study period in a roughly linear fashion, air quality data were detrended before analyses to remove the linear time trend and potential confounding from non–air quality–related variables that also trended over time.

Statistical Analysis

We used a Cox proportional hazards model to model the association between annual average $PM_{2.5}$ exposure at the time of HF diagnosis and all-cause mortality. We used a nested confounder adjustment strategy composed of a basic model, which adjusted for age, sex, race, and distance to the nearest monitor, and a full model that adjusted for age, sex, race, distance to the nearest monitor, and neighborhood level socioeconomic variables assessed at the census block group based on the 2000 Census. The socioeconomic variables included were median income, median house value, percentage of individuals on public assistance, urbanicity, and percentage of households below the federal poverty line. Confounders were chosen a priori. In visualizing the follow-up time for participants, an excess of participants had a recorded death with a follow-up time <1 year. This excess does not follow typical mortality patterns for HF patients[5] and is likely due either to late diagnosis of HF in patients who had had the disease for years, a failure of the EHR to properly capture earlier diagnoses, or patients diagnosed at other hospital systems and coming to the UNC Health Care System only as their condition worsened. When models were compared with all patients, excluding deaths (and associated person-time) within 1 year, it was seen that the proportional hazards assumption was better met in models where the year after diagnosis was excluded as an at-risk time and was consistently violated in models with all person-time included. Thus, we excluded the year after diagnosis as time at risk for all individuals. This had the effect of potentially making estimates applicable only to medium- to long-term mortality as opposed to potential mortality immediately after diagnosis.

Our primary outcome was all-cause mortality. We used stratified (ie, subgroup) analyses to examine potential effect modification by age at HF diagnosis (20-50, 50-65, and 65+), sex, race, and previous diagnosis of chronic obstructive pulmonary disorder, type 2 diabetes mellitus, ischemic heart disease, lipid disorders, pulmonary disease, primary essential hypertension, and peripheral arterial disease. Definitions of the disease outcomes by *ICD-9* and *ICD-10* codes appear in Data S1 (Expanded Methods, comorbid disease *ICD-9* and *ICD-10* codes). We also stratified individuals on diagnosis of systolic versus diastolic HF to determine if associations differed by HF subtype. Because the 1×1-km grids had complete representation across the state, whereas the monitors were typically placed in urban centers, we used the modeled $PM_{2.5}$ data to examine associations in urban versus rural areas, defining urban areas as year 2000 census block groups that were 100% urban (which captured slightly more than the top third of the distribution), and rural areas as those in the bottom third of the urbanicity distribution (<37% urban).

To examine the sensitivity of our results to individuals residing far from a ground-based monitor, we considered 2 levels of monitor radius restriction, individuals within 30 km of a monitor versus individuals within 8 km. Results were equivalent for the primary analyses, so we present results for the 30-km restriction. We also used the 1×1-km modeled $PM_{2.5}$ data as a robustness check for this, as the modeled data cover all participants. Potential geographic heterogeneity was evaluated via a mixed-effects Cox regression model that included a random intercept for county of residence. This type of mixed-effects Cox regression model is also called a shared frailty model. To evaluate effects below the current $PM_{2.5}$ annual average NAAQS, we performed an analysis restricted to those individuals with annual average $PM_{2.5}$ exposure <12 $\mu g/m^3$.

We estimated years of life lost by first estimating the baseline hazard for the participants. We then model associations using a binary indicator for annual average $PM_{2.5}$ <12 $\mu g/m^3$ to estimate the mortality risk conferred by residing in areas above this cutoff. The years of life lost were determined by the excess number of deaths seen from 2 to 5 years (excluding year 1, as this was not counted in person-time at risk in models) based on a mortality hazard proportionally increased by the amount estimated from the aforementioned model. We evaluated sensitivity to diagnoses near the December 31, 2016 study end date by examining a subcohort of those with a HF diagnosis before January 1, 2014. Failing to observe a death can result in "immortal person time," which could bias analyses. To examine this potential bias, we performed 1 sensitivity analysis that removed individuals who were outliers (±3SD) for age at the end of observation (death or December 31, 2016) and 1 in which we removed individuals with an age at end of observation >100 years. Results are reported as the hazard ratio (HR) and associated 95% CI.

Software

All analyses were run in R (R Foundation, Vienna, Austria) version 3.5.3.[15] A relational database constructed in part using FDTool (https://github.com/USEPA/FDTool)[16] and implemented in Microsoft SQL Server was used to manage the data, with database documentation managed using Dataedo (Gdansk, Poland). Maps were created using ArcGIS version 10.7 (Esri, Redlands, CA), and geocoding was performed using SAS version 9.4 (SAS Institute, Cary, NC).

Results

Figure 1 shows the distribution of study participants in North Carolina. A total of 35 084 HF patients had sufficient data for analyses; however, after the restriction of observing events 1 year after HF diagnosis had been applied, only 23 302 individuals contributed person-time at risk and to the analysis. The clinical covariates of these individuals are given in Table 1. The distribution of $PM_{2.5}$ in North Carolina using the 1×1-km resolution modeled values for 2004 and 2016 is given in Figure S1.

We report results for the full model here, with results for the basic and full model appearing in Table S1. Results were concordant for the basic and full models, indicating little evidence of confounding by the additional covariates in the full model. A 1 $\mu g/m^3$ higher annual average $PM_{2.5}$ exposure was associated with a HR of 1.13 (95% CI, 1.10–1.15) for all-cause mortality (Figure 2; Table 2; Table S1). Although there was a slight trend for increasing mortality risk with increasing age at diagnosis, there was no substantial difference in risks when data were stratified on sex, race, or preexisting comorbidities including ischemic heart disease.

When we examined associations using the modeled air pollution data at 1×1-km resolution, we observed similar, often stronger, associations than when we used the monitor data (Table S2). Use of the 1×1-km model allowed us to assess if associations differed by urban versus rural areas. As compared with 2000 Census block groups in the bottom third of the urbanicity distribution (urbanicity <37%), individuals in block groups that were 100% urban (39% of data) had slightly stronger associations with mortality, but the 95% CIs overlapped substantially. Rural residents had an HR of 1.19 (95% CI, 1.14–1.24), whereas urban residents had a HR of 1.22 (95% CI, 1.17–1.27) (Table S2).

Using the regulatory, ground-based monitor data, we observed no differences in the associations for the sensitivity analyses restricted by HF diagnosis date or age at end of observation or when restricted to individuals with at least 1 prior visit (Table S3). Associations for patients diagnosed with systolic versus diastolic HF were generally similar except in the mixed-effects model (Table 2), where associations were stronger for those with diastolic HF. This indicates that after accounting for differences in baseline hazard by county, individuals with diastolic HF might be more sensitive to long-term $PM_{2.5}$ exposure. Clinical covariates for systolic versus diastolic HF patients are in Table S4.

When the study was restricted to individuals residing in areas with annual average $PM_{2.5}$ <12 $\mu g/m^3$, we still observed significant associations equivalent to those seen in the full $PM_{2.5}$ distribution (Figure 2 bottom; Table 2). HF patients residing in areas with annual average $PM_{2.5}\geq12$ $\mu g/m^3$ had an almost 2-fold higher mortality risk than HF patients with lower exposures (HR 1.77; 95% CI, 1.66–1.89; Table S5).

Use of a mixed-effects model with a random intercept for county to adjust for heterogeneity between counties revealed no evidence of confounding by county-level geographic heterogeneity (Table 2). A concentration-response curve using either the modeled 1×1-km $PM_{2.5}$ data (Figure 3) or the monitor data (Figure S2) for all-cause mortality indicated an approximately linear concentration-response function for annual average $PM_{2.5}$, particularly in the 8- to 12-$\mu g/m^3$ range. There was some thresholding outside this range, but whether this indicates a leveling off of risk or is more driven by a relative lack of observations will need to be evaluated in cohorts with broader ranges of observed concentrations. We used the data from the regulatory $PM_{2.5}$ monitoring network in a mixed effects model with county-specific intercepts to estimate the years of life lost for someone residing in an area with $PM_{2.5}$ $\geq$12 $\mu g/m^3$ as compared with someone living below this level to be 0.84 years (95% CI, 0.73–0.95) over a 5-year period. This effect persisted even when exposures were restricted to below 12 $\mu g/m^3$. Stratifying the median exposure (10.1 $\mu g/m^3$) for individuals with $PM_{2.5}$ concentrations <12 $\mu g/m^3$, we observed 0.54 (95% CI, 0.40–0.69) years of life lost for above-median exposures as compared with below-median exposures.

Discussion

The increasing prevalence of HF in the United States,[4] its high risk of mortality, and its societal burden create an imperative to identify its modifiable risk factors. The observation of a significant association between annual average $PM_{2.5}$ exposure and mortality in HF patients implicates long-term exposure to poor air quality as a risk factor for poor clinical outcomes among HF patients. These data provide further evidence that long-term exposure to ambient air particle pollution contributes to heart disease and mortality[17,18] and underscore the need to better understand environmental exposures as potentially modifiable risk factors in highly vulnerable populations. To date, most manuscripts in air pollution epidemiology have focused on atherosclerotic

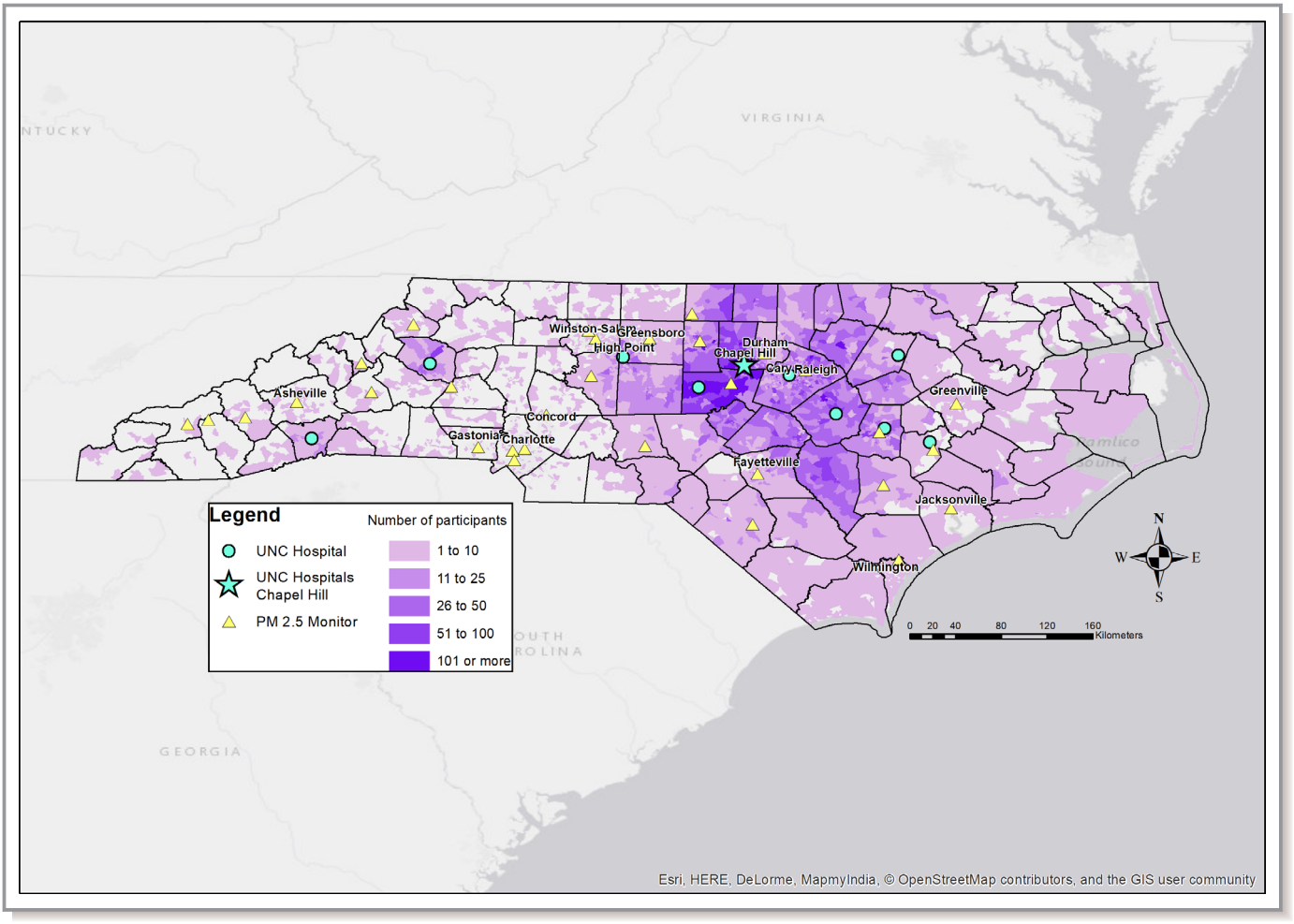


Figure 1. Map of HF patients. A map of the density of HF patients in CARES by census block group. Darker shading indicates a higher density of observed HF patients. Also indicated are the UNC-affiliated hospitals (blue circles) with the flagship hospital, located in Chapel Hill, NC, given as a blue star. EPA $PM_{2.5}$ monitors are represented as yellow triangles. CARES indicates Clinical and Archived Records Research for Environmental Studies; EPA, Environmental Protection Agency; HF, heart failure; $PM_{2.5}$, particulate matter <2.5 μm in diameter; UNC, University of North Carolina.

disease such as coronary heart disease. Here we focus on heart failure, an understudied patient group. Patients with both heart failure and atherosclerotic disease did not differ in their associations from the overall analysis, and adjusting for ischemic heart disease did not substantially alter models (data not shown). Associations observed here are similar in magnitude to associations between $PM_{2.5}$ and myocardial infarction in patients with atherosclerotic disease,[19] suggesting that both HF and atherosclerotic disease patients may have high sensitivity to air pollution exposure.

Although air quality has improved substantially over the past 4 decades in the United States, millions of people live, work, and play in areas where the annual $PM_{2.5}$ concentration is near or exceeds EPA's NAAQS for $PM_{2.5}$. The need to understand that the effects of air quality on health are compounded by recent wildland fire–driven trends in air pollution that have seen annual average $PM_{2.5}$ increases in the United States.[20] Because wildland fire–generated $PM_{2.5}$ increases HF emergency department visits,[21] more frequent and larger wildland fires may exacerbate health effects from rising HF prevalence. Previous studies of short-term exposures have indicated that associations between air pollution and mortality are higher for HF deaths than other causes.[11,12] To date, few studies have specifically examined air pollution exposure impacts within a HF cohort. Existing studies have focused on daily mortality with short-term exposures,[7,9] leaving questions on the relationship between long-term exposure and mortality unanswered.

In this study an increase of 1 $\mu g/m^3$ in annual average $PM_{2.5}$ at the patient's residence was associated with a 13% higher mortality risk (95% CI, 10–15). Residential exposure to annual average $PM_{2.5}$ above 12 $\mu g/m^3$, the current $PM_{2.5}$ NAAQS, was associated with 0.84 years (95% CI, 0.73–0.95) of life lost as compared with exposures below this level. Even

Table 1. Clinical Covariates

Clinical Covariates (N=23 302)	Mean	SD	IQR
Age (y)	66.9	15.2	22.2
Follow-up time (y)	4.23	3.14	4.05
Distance to monitor (km)	21.6	18.3	21.9
$PM_{2.5}$ (Monitor) ($\mu g/m^3$)	10.2	2.11	3.36
$PM_{2.5}$ (1×1-km model) ($\mu g/m^3$)	10.3	1.70	2.45
Households below federal poverty line (%)	18.6	15.0	19.4
Median home value ($)	177 041	106 435	114 100
Median household income ($)	51 881	25 924	29 962
Urbanicity (%)	62.7	42.2	89.9
Households receiving public assistance (%)	2.00	3.01	2.94
	N	%	
Black	7063	30.3	
White	14 216	61.0	
Other race	2023	8.7	
Male	11 224	48.2	
Within 30 km of monitor	17 212	73.9	
Within 8 km of monitor	4849	20.8	
Death (all cause)	4496	19.3	
Type 2 diabetes mellitus	7853	33.7	
IHD	13 260	56.9	
COPD	8293	35.6	
PAD	8195	35.2	
Hypertension	17 027	73.1	
Dyslipidemia	17 215	73.9	
Systolic HF	7120	30.6	
Diastolic HF	6385	29.3	

Clinical covariates for the HF patients. Units of measurement for continuous variables given in parentheses. Households below federal poverty line, median home value, median household income, urbanicity, and households receiving public assistance assessed at the block-group level based on the 2000 US Census. Numbers of systolic and diastolic HF patients do not sum to the total because many had an unspecified HF subtype in their medical record. COPD indicates chronic obstructive pulmonary disorder; HF, heart failure; IHD, ischemic heart disease; IQR, interquartile range; PAD, peripheral arterial disease; $PM_{2.5}$, particulate matter <2.5 μm in diameter.

when we restricted data collection to HF patients with residential exposures below the current $PM_{2.5}$ standard, we observed substantial years of life lost for those with elevated exposures (above the median; 10.1 $\mu g/m^3$) as compared with patients with lower exposure. This suggests that the protective effect of lower residential $PM_{2.5}$ exposure persists at levels well below current $PM_{2.5}$ standards.

Diastolic and systolic HF differ substantially in clinical presentation. Diastolic HF is characterized by improper filling of the left ventricle during diastole, typically due to a stiffening and abnormal relaxation of the left ventricle, which result in ventricular pressure overload despite the ejection fraction being largely preserved. In contrast, systolic HF is characterized by inefficient contraction during systole due to decreased contractility (typically with enlargement) of the left ventricle. In models accounting for differential baseline mortality rates by county, we observed differences in $PM_{2.5}$-related mortality risks by systolic versus diastolic HF diagnosis with $PM_{2.5}$ potentially more strongly associated with all-cause mortality among those with diastolic HF than among those with systolic HF. Several mechanisms may account for links between $PM_{2.5}$ exposure and mortality among those with HF, and in particular diastolic HF. One such mechanism goes through the differential role of reactive oxygen species (ROS) in each HF subtype. ROS dysfunction is typical in HF, and individuals with diastolic HF may have ROS dysfunction primarily linked to endothelial cells, which can then lead to prohypertrophic signaling and sarcomere stiffness. In systolic HF, ROS dysfunction may more closely be tied to myocyte function/dysfunction.[22] $PM_{2.5}$ exposure is strongly linked to oxidative stress and ROS production[23–26] and may exacerbate ROS dysfunction in HF patients. Autonomic dysfunction may be an additional mechanism linking $PM_{2.5}$ exposure and mortality among HF patients. In a study of 31 HF patients, higher $PM_{2.5}$ exposure was associated with lower oxygen saturation (measured by pulse oximetry) and higher heart rate,[27] both of which would be impacted by autonomic dysfunction. Other mechanisms linking $PM_{2.5}$ exposure and mortality in individuals with HF are dysfunction in vasoconstriction and blood pressure regulation. Long-term air pollution exposure is associated with elevated systolic and diastolic blood pressure.[28-30] HF, particularly diastolic HF, is marked by blood pressure regulation dysfunction and is often accompanied by hypertension, and $PM_{2.5}$ exposure may worsen this dysfunction. We also cannot ignore the possibility of metabolic dysfunction. Type 2 diabetes mellitus is a risk factor for HF,[31] and $PM_{2.5}$ exposure is associated with metabolic dysfunction,[32-34] providing another means by which exposure may exacerbate existing organ system dysfunction, thereby increasing mortality risk. Both hypertension and type 2 diabetes mellitus were more prevalent in diastolic HF patients, offering another potential means by which associations may differ by HF subtypes. Because much of our population resided in urban areas, combustion-related exposures may also have played a large role. Combustion-related $PM_{2.5}$, typically derived from traffic in urban areas, is strongly associated with cardiovascular disease[35] and many of the risk factors related to HF.

Strengths

This large study of over 23 000 patients is the largest environmental health study to focus exclusively on HF patients. As far as the authors are aware, this is the first

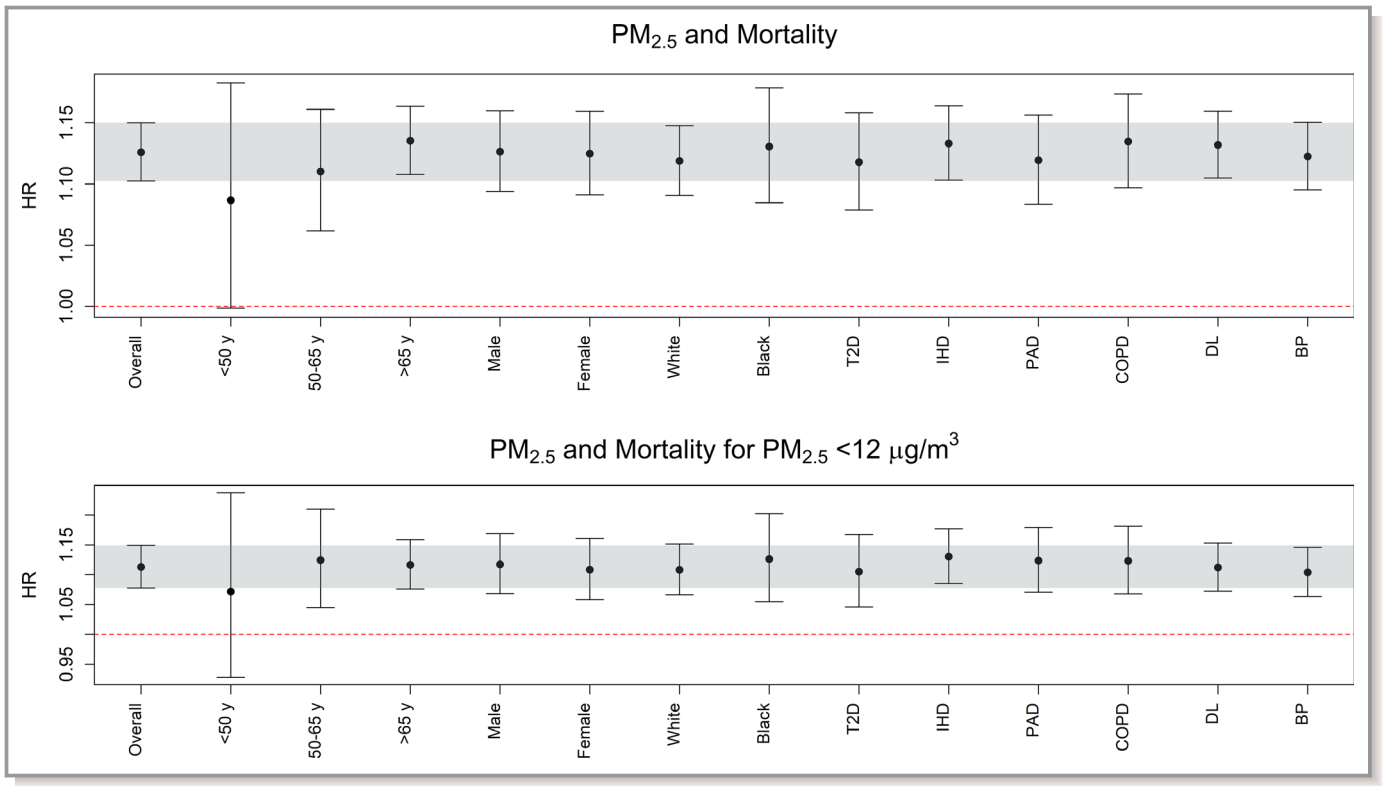


Figure 2. Associations between annual average $PM_{2.5}$ exposure and all-cause mortality among HF patients. The top graph gives the association for all exposure levels, and the bottom graph is restricted to residences with exposure below the 12 μg/m^3 NAAQS for annual average $PM_{2.5}$. The full-adjustment model was used, and the association in the entire population is given on the far left ("Overall"). The gray bar indicates the 95% CI for the "overall" association, and stratified analyses are listed based on age, sex, race, and existing comorbidities. Tabular results are given in Table S3. BP indicates hypertension; COPD, chronic obstructive pulmonary disorder; DL, dyslipidemia; HF, heart failure; HR, hazard ratio; IHD, ischemic heart disease; NAAQS, National Ambient Air Quality Standard; PAD, peripheral arterial disease; $PM_{2.5}$, particulate matter <2.5 μm in diameter; T2D, type 2 diabetes mellitus.

study to estimate mortality risks for HF patients stratified on demographic and clinical characteristics. The use of EHRs makes extensive information available on disease diagnosis history, which has allowed us to examine differences by HF subtype and to estimate mortality risks based on residential exposure at initial indication of HF in the medical record. Much of North Carolina is exposed to $PM_{2.5}$ near or below the EPA-mandated annual average standard, which gives us the ability to explore associations at and below the mandated ambient concentrations.

Table 2. Sensitivity Analyses for Association Between $PM_{2.5}$ Exposure Estimated From Regulatory Monitoring Network and All-Cause Mortality Stratified by HF Subtype

	All Individuals		Diastolic HF		Systolic HF	
	HR (95% CI)	N (Deaths)	HR (95% CI)	N (Deaths)	HR (95% CI)	N (Deaths)
Primary analysis	1.13 (1.10–1.15)	23 012 (4445)	1.09 (1.05–1.13)	6315 (1449)	1.07 (1.03–1.11)	7041 (1055)
Restricted to exposures with $PM_{2.5}$ <12 μg/m^3	1.11 (1.08–1.15)	18 055 (2151)	1.08 (1.02–1.15)	4853 (687)	1.08 (1.02–1.15)	6228 (696)
Restricted to participants <30 km from monitor	1.12 (1.09–1.15)	16 787 (3460)	1.07 (1.03–1.12)	4883 (1224)	1.08 (1.03–1.13)	5113 (836)
Primary analysis with random intercept for county	1.16 (1.13–1.19)	23 012 (4445)	1.12 (1.07–1.16)	6315 (1449)	1.07 (1.03–1.11)	7041 (1055)

We conducted a series of sensitivity analyses to understand the robustness of our primary analysis for all individuals, those with diastolic HF, and those with systolic HF. All results are from the full model adjusted for age, sex, race, distance to the nearest monitor, median income, median house value, percentage of individuals on public assistance, urbanicity, and percentage of households below the federal poverty line. HF indicates heart failure; HR, hazard ratio; $PM_{2.5}$, particulate matter <2.5 μm in diameter.

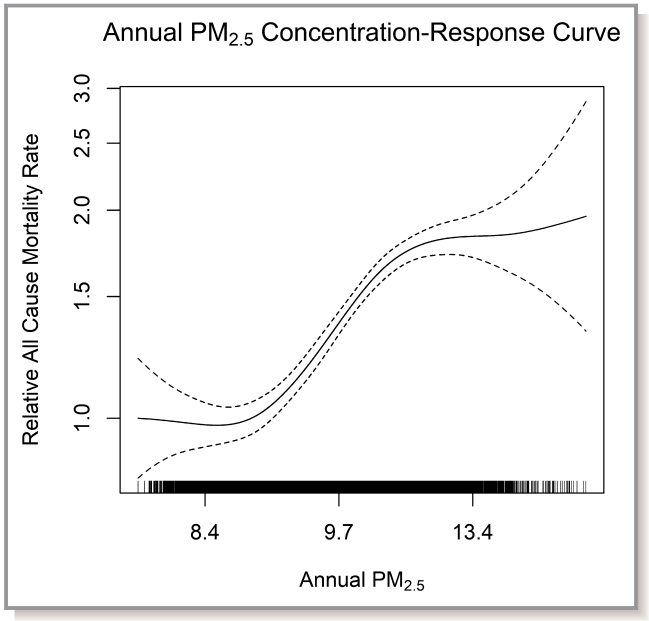


Figure 3. Concentration-response curve for all-cause mortality for annual average $PM_{2.5}$ exposure based on modeled $PM_{2.5}$ concentrations at 1×1-km resolution.[14] Concentration-response curve limited to $PM_{2.5}$ concentrations within the inner 95% of the distribution (values >8.0 μg/m^3 and <14 μg/m^3) as the confidence intervals widened considerably beyond this range. The curve is broadly similar to that seen when using the monitors (particularly for concentrations from 9 to 12 μg/m^3; Figure S2) with perhaps better behavior toward the upper end of the concentration distribution. $PM_{2.5}$ indicates particulate matter <2.5 μm in diameter.

Limitations

A primary limitation of this study is the incomplete capture of individual-level confounders in EHR systems. This means that potentially important confounders such as smoking were not included in the models. This is an unfortunate but common limitation of studies such as these. However, previous studies of air quality and mortality, using similar resources, have indicated that associations are robust to the omission of such confounders.[36] Another limitation of this study is that patients may use multiple hospital systems, resulting in uncertainty in the determination of the initial HF diagnosis. This date of diagnosis uncertainty is unlikely to be related to air quality and thus would not be expected to bias our associations, although it would still increase the standard error of the estimated HRs. Bias introduced by misclassification of initial HF diagnosis date would be expected to be larger for those diagnosed later in life, where there is a greater chance of previous diagnosis in a different hospital system. Additionally, sensitivity analyses restricted to individuals with a visit before their HF diagnosis, which should limit the potential for date-of-diagnosis misclassification, showed similar association as the primary model (Table S3). Still, analyses would be improved by expanding existing EHR resources to incorporate medical records over the complete life course.

As a hospital-based cohort, this population may not generalize to the entire population of North Carolina; however, as most people with HF are likely to visit the hospital, this cohort is largely reflective of the HF population, particularly for more urban areas where much of the study cohort resides (Figure 1). Limited individual-level information on socioeconomic status is available in the medical record; however, publicly available data on the socioeconomic condition of the census block group of participants were used as proxy for individual-level socioeconomic status and thus supplemented the information available in the medical record. A final limitation was our use of ground-based monitors for air quality assessments. Although monitor data are widely utilized for air pollution studies, they offer incomplete geographic coverage. Both the monitors and hospitals are often located near urban areas, and most of our population lived within 30 km of a monitor. However, we overcame this limitation by examining associations using modeled air quality at 1-km resolution for the entire study period. Associations were similar and often even stronger using the modeled air quality data, something that has been observed in other studies.[37]

Conclusions

Long-term residential exposure to $PM_{2.5}$ is associated with all-cause mortality in HF patients, even for exposures below the current $PM_{2.5}$ NAAQS. Mortality risks may be elevated for patients with diastolic HF as compared with systolic HF. Given the magnitude of the associations and associated estimated years of life lost, understanding environmental risks and mitigating harmful exposures may provide substantial benefits to the HF patient population. With current nationwide efforts to reduce cardiovascular disease burden through education of patients and clinicians, as with Million Hearts, it is imperative that we continue to improve our understanding of the contribution of environmental exposures to cardiovascular mortality.

Acknowledgments

This work does not necessarily represent the views or policy of the US EPA. Any mention of trade names does not constitute endorsement.

Sources of Funding

All funding for this study was provided by the US Environmental Protection Agency. The funder had no role in the study design, data collection, analysis, interpretation, or manuscript writing.

Disclosures

None.

References

1. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K, Brunekreef B, Dandona L, Dandona R, Feigin V, Freedman G, Hubbell B, Jobling A, Kan H, Knibbs L, Liu Y, Martin R, Morawska L, Pope CA, Shin H, Straif K, Shaddick G, Thomas M, van Dingenen R, van Donkelaar A, Vos T, Murray CJL, Forouzanfar MH. 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. *Lancet*. 2017;389:1907–1918.
2. Ko DT, Alter DA, Austin PC, You JJ, Lee DS, Qiu F, Stukel TA, Tu JV. Life expectancy after an index hospitalization for patients with heart failure: a population-based study. *Am Heart J*. 2008;155:324–331.
3. MacIntyre K, Capewell S, Stewart S, Chalmers J, Boyd J, Finlayson A, Redpath A, Pell J, McMurray J. Evidence of improving prognosis in heart failure: trends in case fatality in 66 547 patients hospitalized between 1986 and 1995. *Circulation*. 2000;102:1126–1131.
4. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, Ikonomidis JS, Khavjou O, Konstam MA, Maddox TM, Nichol G, Pham M, Pina IL, Trogdon JG. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail*. 2013;6:606–619.
5. Taylor CJ, Ordóñez-Mena JM, Roalfe AK, Lay-Flurrie S, Jones NR, Marshall T, Hobbs FDR. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000–2017: population based cohort study. *BMJ*. 2019;364:l223.
6. Wellenius GA, Yeh GY, Coull BA, Suh HH, Phillips RS, Mittleman MA. Effects of ambient air pollution on functional status in patients with chronic congestive heart failure: a repeated-measures study. *Environ Health*. 2007;6:26.
7. Goldberg MS, Burnett RT, Valois M-F, Flegel K, Bailar Iii JC, Brook J, Vincent R, Radon K. Associations between ambient air pollution and daily mortality among persons with congestive heart failure. *Environ Res*. 2003;91:8–20.
8. Wellenius GA, Schwartz J, Mittleman MA. Particulate air pollution and hospital admissions for congestive heart failure in seven United States cities. *Am J Cardiol*. 2006;97:404–408.
9. Shah ASV, Langrish JP, Nair H, McAllister DA, Hunter AL, Donaldson K, Newby DE, Mills NL. Global association of air pollution and heart failure: a systematic review and meta-analysis. *Lancet*. 2013;382:1039–1048.
10. Goldberg MS, Burnett RT, Bailar JC, Tamblyn R, Ernst P, Flegel K, Brook J, Bonvalot Y, Singh R, Valois MF, Vincent R. Identification of persons with cardiorespiratory conditions who are at risk of dying from the acute effects of ambient air particles. *Environ Health Perspect*. 2001;109:487–494.
11. Hoek G, Brunekreef B, Fischer P, van Wijnen J. The association between air pollution and heart failure, arrhythmia, embolism, thrombosis, and other cardiovascular causes of death in a time series study. *Epidemiology*. 2001;12:355–357.
12. Pope CA III, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, Godleski JJ. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation*. 2004;109:71–77.
13. U.S. Environmental Protection Agency. U.S. EPA. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2009). Washington, DC, EPA/600/R-08/139F, 2009.
14. Di Q, Amini H, Shi L, Kloog I, Silvern R, Kelly J, Sabath MB, Choirat C, Koutrakis P, Lyapustin A, Wang Y, Mickley LJ, Schwartz J. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environ Int*. 2019;130:104909.
15. R Core Team. R: A language and environment for statistical computing. 2018.
16. Buranosky M, Stellnberger E, Pfaff E, Diaz-Sanchez D, Ward-Caviness C. FDTool: a Python application to mine for functional dependencies and candidate keys in tabular data [version 2; peer review: 2 approved]. *F1000Research*. 2019;7.
17. Kaufman JD, Adar SD, Barr RG, Budoff M, Burke GL, Curl CL, Daviglus ML, Roux AVD, Gassett AJ, Jacobs DR, Kronmal R, Larson TV, Navas-Acien A, Olives C, Sampson PD, Sheppard L, Siscovick DS, Stein JH, Szpiro AA, Watson KE. 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. *Lancet*. 2016;388:696–704.
18. Brook RD, Rajagopalan S, Pope CA 3rd, Brook JR, Bhatnagar A, Diez-Roux AV, Holguin F, Hong Y, Luepker RV, Mittleman MA, Peters A, Siscovick D, Smith SC Jr, Whitsel L, Kaufman JD. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*. 2010;121:2331–2378.
19. McGuinn LA, Ward-Caviness C, Neas LM, Schneider A, Di Q, Chudnovsky A, Schwartz J, Koutrakis P, Russell AG, Garcia V, Kraus WE, Hauser ER, Cascio W, Diaz-Sanchez D, Devlin RB. Fine particulate matter and cardiovascular disease: comparison of assessment methods for long-term exposure. *Environ Res*. 2017;159:16–23.
20. McClure CD, Jaffe DA. US particulate matter air quality improves except in wildfire-prone areas. *Proc Natl Acad Sci USA*. 2018;115:7901–7906.
21. 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:e007492. DOI:10.1161/JAHA.117.007492.
22. Münzel T, Gori T, Keaney JJF, Maack C, Daiber A. Pathophysiological role of oxidative stress in systolic and diastolic heart failure and its therapeutic implications. *Eur Heart J*. 2015;36:2555–2564.
23. Riva DR, Magalhães CB, Lopes AA, Lanças T, Mauad T, Malm O, Valença SS, Saldiva PH, Faffe DS and Zin WA. Low dose of fine particulate matter (PM2.5) can induce acute oxidative stress, inflammation and pulmonary impairment in healthy mice. *Inhalation Toxicol*. 2011;23:257–267.
24. Rui W, Guan L, Zhang F, Zhang W and Ding W. $PM_{2.5}$-induced oxidative stress increases adhesion molecules expression in human endothelial cells through the ERK/AKT/NF-κB-dependent pathway. *J Appl Toxicol*. 2016;36:48–59.
25. Araujo JA, Barajas B, Kleinman M, Wang X, Bennett BJ, Gong KW, Navab M, Harkema J, Sioutas C, Lusis AJ, Nel AE. Ambient particulate pollutants in the ultrafine range promote early atherosclerosis and systemic oxidative stress. *Circ Res*. 2008;102:589–596.
26. Miller MR, Shaw CA, Langrish JP. From particles to patients: oxidative stress and the cardiovascular effects of air pollution. *Future Cardiol*. 2012;8:577–602.
27. Goldberg MS, Giannetti N, Burnett RT, Mayo NE, Valois M-F, Brophy JM. Shortness of breath at night and health status in congestive heart failure: effects of environmental conditions and health-related and dietary factors. *Environ Res*. 2009;109:166–174.
28. Chan SH, Hee VCV, Bergen S, Szpiro AA, DeRoo LA, London SJ, Marshall JD, Kaufman JD, Sandler DP. Long-term air pollution exposure and blood pressure in the sister study. *Environ Health Perspect*. 2015;123:951–958.
29. Chuang K-J, Yan Y-H, Cheng T-J. Effect of air pollution on blood pressure, blood lipids, and blood sugar: a population-based approach. *J Occup Environ Med*. 2010;52:258–262.
30. Fuks K, Moebus S, Hertel S, Viehmann A, Nonnemacher M, Dragano N, Möhlenkamp S, Jakobs H, Kessler C, Erbel R, Hoffmann B. Long-term urban particulate air pollution, traffic noise, and arterial blood pressure. *Environ Health Perspect*. 2011;119:1706–1711.
31. Nichols GA, Gullion CM, Koro CE, Ephross SA, Brown JB. The incidence of congestive heart failure in type 2 diabetes: an update. *Diabetes Care*. 2004;27:1879–1884.
32. Ward-Caviness CK, Kraus WE, Blach C, Haynes CS, Dowdy E, Miranda ML, Devlin RB, Diaz-Sanchez D, Cascio WE, Mukerjee S, Stallings C, Smith LA, Gregory SG, Shah SH, Hauser ER, Neas LM. Association of roadway proximity with fasting plasma glucose and metabolic risk factors for cardiovascular disease in a cross-sectional study of cardiac catheterization patients. *Environ Health Perspect*. 2015;123:1007–1014.
33. Rajagopalan S, Brook RD. Air pollution and type 2 diabetes. *Mechanistic Insights*. 2012;61:3037–3045.
34. Krämer U, Herder C, Sugiri D, Strassburger K, Schikowski T, Ranft U, Rathmann W. Traffic-related air pollution and incident type 2 diabetes: results from the SALIA Cohort study. *Environ Health Perspect*. 2010;118:1273–1279.
35. Bourdrel T, Bind M-A, Béjot Y, Morel O, Argacha J-F. Cardiovascular effects of air pollution. *Arch Cardiovasc Dis*. 2017;110:634–642.
36. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, Dominici F, Schwartz JD. Air pollution and mortality in the Medicare population. *N Engl J Med*. 2017;376:2513–2522.
37. Di Q, Dai L, Wang Y, Zanobetti A, Choirat C, Schwartz JD, Dominici F. Association of short-term exposure to air pollution with mortality in older adults. *JAMA*. 2017;318:2446–2456. DOI: 10.1001/jama.2017.17923.

SUPPLEMENTAL MATERIAL

Data S1.

Expanded Methods

Geocoding Procedure:

We obtained patient addresses from two medical records datasets: the Epic® electronic health record (EHR) dataset contained 112,519 addresses from visits occurring from 2014 to 2016; the homegrown WebCIS (legacy) EHR dataset contained 63,856 addresses from visits prior to 2014. We began cleaning data by standardizing variable formats. The variables included two text variables for address lines 1 (street address) and 2 (apartment or unit number), a text variable for city, a text variable for state, and a numeric variable for zip/postal code. We then removed duplicate addresses: the exact same text entered as an address two or more times for a single participant ("Main St" and "Main Street" would not be considered duplicates at this stage). This reduced our number of addresses to 86,616 in the Epic® EHR dataset and 58,905 in legacy. Both datasets underwent initial geocoding using proc geocode in SAS version 9.4 (SAS Institute Inc. Cary, NC). We matched addresses to latitude and longitude using street lookup data from SAS MapsOnline, which was generated from the US Census Bureau's TIGER/Line shapefiles. SAS proc geocode detects and corrects for minor textual variants, such as common spelling errors and abbreviations. Epic® and legacy data were then merged based on level of initial geocoding match: street level (n = 101 594), zip code level (n = 43 285), city level (n = 500), or unmatched (n = 142).

We then cleaned data to improve the level of matching for addresses not matched to the street level, where possible. We began by cleaning data with addresses that were unmatched. Various

indicators of null address elements (“unknown”, “none”, “NA”, etc.) were standardized to all read “NULL”. Then, observations were separated into those with no valid address information and those with partial information (i.e. only zip code). Those with partial address information were merged to the city matched dataset, while those with no valid address information were kept as unmatched.

For those addresses in the city matched dataset, we standardized null address elements as above, standardized the spelling of PO Box, and removed punctuation from text fields. PO Boxes are Post Office Boxes, for individuals who receive mail at a Post Office instead of their home. PO Box addresses can be geocoded to zip code, but not street address. We assumed that PO Boxes were located in the same zip code where a patient resided. We then identified likely mistakes in which the listed city was a near but not exact match to the listed zip code and corrected the zip code. We then re-geocoded city matched addresses. Those who matched to zip code at this point were merged with the zip code matched dataset.

For those addresses matched to zip code, we standardized null address elements as above, standardized the spelling of PO Box, removed punctuation, corrected obvious misspellings and unconventional abbreviations, and standardized indicators of homelessness. Many individuals experiencing homelessness have listed a zip code where they typically stay, although they may be transient. Some participants listed both a PO Box and a street address. For those individuals, we attempted to geocode the street address. We then standardized text fields of data entries including apartments. If street address and apartment number were entered in an unconventional format (i.e. listing apartment number before street address), we attempted to extract the street

address and re-geocode the address. Then, for any addresses with text on the second address line, we attempted to geocode the second address line. Any addresses that were re-geocoded to street level at this stage were merged to street level matches.

For street level matches, we corrected the address of a common senior living community that geocoded incorrectly in some instances. At this point, we removed duplicate addresses: multiple addresses matched to the same street (for street level matches) or zip code (for zip code level matches) for the same participant. We then matched participants to their "best" address(es), by source and by matching level. We created a dummy dataset with one observation per participant (n = 41 913). We first matched participants to addresses from Epic® that were street matches. If there were multiple unique addresses per participant, we kept all addresses (n = 32 614 addresses). If a participant was not matched to street in Epic™, we repeated the same procedure for addresses from Epic® that were zip code matches (n = 3 857 addresses). For participants still not matched to an address, we matched them to addresses from legacy files that were street matched (n = 9 006 addresses), keeping only the most recent based on account number, an approximate estimation of chronology. For participants still not matched, we repeated the procedure in order for the following sources and types of match, keeping only the most recent instance: legacy zip code match (n = 2 118 addresses), Epic® city match (n=47 addresses), legacy city match (n = 2 addresses), and unmatched (n = 17 participants).

Once participants' address(es) were established, we endeavored to match participants to their nearest EPA $PM_{2.5}$ monitor. We limited to participants who resided in NC and matched addresses to all visit dates. For each visit date, we calculated the date 365 prior to the visit. We then took

all data from EPA monitors across the state of NC during the period from January 1, 2003 until December 31, 2016. We calculated the first and last dates within that period when each monitor had valid data. We matched all monitors that were active during the entire 365 days prior to participants' visits to their addresses. We calculated the distances between the monitor and the addresses using the latitude and longitude of the addresses and monitors and the geodist function in SAS 9.4. We then kept the nearest monitor to each address for each visit date. We then averaged the daily mean $PM_{2.5}$ value for the 365 days prior to each visit.

Co-morbid disease ICD-9 and ICD-10 codes:

International Classification of Disease (ICD)-9 and ICD-10 codes used to determine existing co-morbidities are given below. A * is used as a wildcard character which was used to capture all subcodes for relevant disease definitions.

Type 2 diabetes: 250, 250.0, 250.00, 250.02, 250.1, 250.10, 250.12, 250.2, 250.20, 250.22, 250.3, 250.30, 250.32, 250.4, 250.40, 250.42, 250.5, 250.50, 250.52, 250.6, 250.60, 250.62, 250.7, 250.70, 250.72, 250.8, 250.80, 250.82, 250.9, 250.90, 250.92, E11.*

Hypertension: 401*, I10*

Dyslipidemia: 272*, E78*

Peripheral Arterial Disease: 443*, I73*

Chronic Obstructive Pulmonary Disorder (bronchitis): 491*, J44*

Emphysema: 492*, J43*

Ischemic Heart Disease: 414*, I20, I21, I22, I23, I24, I25

Table S1. Associations between all-cause mortality and $PM_{2.5}$.

Stratification	Model	HR	LCI	UCI	N	N deaths
None	basic	1.12	1.10	1.15	23302	4496
None	full	1.13	1.10	1.15	23012	4445
Males	basic	1.12	1.09	1.16	11224	2229
Males	full	1.13	1.09	1.16	11075	2208
Females	basic	1.12	1.09	1.16	12078	2267
Females	full	1.12	1.09	1.16	11937	2237
European Americans	basic	1.11	1.09	1.14	14216	2953
European Americans	full	1.12	1.09	1.15	14090	2926
African Americans	basic	1.13	1.09	1.18	7063	1322
African Americans	full	1.13	1.08	1.18	6934	1300
<50 y	basic	1.10	1.01	1.19	3253	383
<50 y	full	1.09	1.00	1.18	3182	376
50-65 y	basic	1.11	1.06	1.16	6943	1126
50-65 y	full	1.11	1.06	1.16	6825	1110
≥65 y	basic	1.13	1.11	1.16	13106	2987
≥65 y	full	1.14	1.11	1.16	13005	2959
$PM_{2.5} < 12$ μg/m^3	basic	1.11	1.07	1.14	18268	2166
$PM_{2.5} < 12$ μg/m^3	full	1.11	1.08	1.15	18055	2151
Type 2 diabetes	basic	1.12	1.08	1.16	7853	1768
Type 2 diabetes	full	1.12	1.08	1.16	7741	1743
PAD	basic	1.12	1.08	1.15	8195	1767
PAD	full	1.12	1.08	1.16	8100	1749
COPD	basic	1.13	1.09	1.17	8293	1772
COPD	full	1.13	1.10	1.17	8177	1747
Dyslipidemia	basic	1.13	1.10	1.16	17215	3402
Dyslipidemia	full	1.13	1.10	1.16	17017	3367
Hypertension	basic	1.12	1.09	1.15	17027	3482
Hypertension	full	1.12	1.10	1.15	16815	3442
IHD	basic	1.13	1.10	1.16	13260	2871
IHD	full	1.13	1.10	1.16	13093	2842

Associations are given per 1 μg/m^3 increase in $PM_{2.5}$. Associations were stratified on sex, race, Hispanic ethnicity, age at heart failure diagnosis, exposure level, and pre-exiting co-morbidities. The "basic" model adjusted for age, race, sex, and distance to the nearest monitor while the "full"

model adjusted for age, sex, race, distance to the nearest monitor, median income, median house value, percent of individuals on public assistance, urbanicity, and percent of households below federal poverty line.

* COPD = chronic obstructive pulmonary disorder; HR = hazard ratio; IHD = Ischemic heart disease; LCI = lower 95% confidence interval; PAD = peripheral arterial disease; UCI = upper 95% confidence interval

Table S2. Associations between all-cause mortality and $PM_{2.5}$ assessed using modeled $PM_{2.5}$ data at 1x1 km resolution.

Stratification	Model	HR	LCI	UCI	N	N deaths
None	basic	1.18	1.15	1.21	23127	4480
None	full	1.19	1.16	1.22	22841	4429
Males	basic	1.17	1.13	1.21	11146	2219
Males	full	1.18	1.14	1.22	11001	2198
Females	basic	1.19	1.15	1.23	11981	2261
Females	full	1.20	1.15	1.24	11840	2231
European Americans	basic	1.17	1.14	1.21	14113	2945
European Americans	full	1.18	1.15	1.22	13988	2918
African Americans	basic	1.18	1.13	1.24	7004	1315
African Americans	full	1.19	1.13	1.24	6877	1293
<50 y	basic	1.17	1.06	1.28	3214	381
<50 y	full	1.16	1.05	1.27	3146	374
50-65 y	basic	1.17	1.11	1.23	6883	1124
50-65 y	full	1.18	1.12	1.24	6766	1108
≥65 y	basic	1.19	1.15	1.22	13030	2975
≥65 y	full	1.19	1.16	1.23	12929	2947
$PM_{2.5}$ < 12 µg/m^3	basic	1.18	1.14	1.22	18109	2152
$PM_{2.5}$ < 12 µg/m^3	full	1.19	1.15	1.24	17900	2137
T2D	basic	1.18	1.14	1.23	7812	1765
T2D	full	1.19	1.14	1.24	7702	1740
PAD	basic	1.16	1.12	1.21	8135	1760
PAD	full	1.17	1.13	1.22	8042	1742
COPD	basic	1.19	1.15	1.24	8241	1767
COPD	full	1.20	1.15	1.25	8127	1742
Dyslipidemia	basic	1.18	1.15	1.21	17093	3391
Dyslipidemia	full	1.19	1.15	1.22	16899	3356
Hypertension	basic	1.17	1.13	1.20	16903	3468
Hypertension	full	1.18	1.14	1.21	16694	3428
IHD	basic	1.18	1.14	1.21	13169	2863
IHD	full	1.19	1.15	1.22	13003	2834
rural	full	1.19	1.14	1.24	7634	1534
urban	full	1.22	1.17	1.27	8840	1697

Associations are given per 1 $\mu g/m^3$ increase in $PM_{2.5}$ with PM2.5 as assessed using the ensemble model values which have a 1x1 km resolution. Associations were stratified on sex, race, Hispanic ethnicity, age at heart failure diagnosis, exposure level, pre-exiting co-morbidities, and urban (100% urban block group) vs rural (bottom 3^{rd} of urbanicity distribution; ~ < 39% urban) residence based on the 2000 Census block group urbanicity assessment. The "basic" model adjusted for age, race, sex, and distance to the nearest monitor while the "full" model adjusted for age, sex, race, distance to the nearest monitor, median income, median house value, percent of individuals on public assistance, urbanicity, and percent of households below federal poverty line.

* COPD = chronic obstructive pulmonary disorder; HR = hazard ratio; IHD = Ischemic heart disease; LCI = lower 95% confidence interval; PAD = peripheral arterial disease; UCI = upper 95% confidence interval

Table S3. Sensitivity analyses based on visit and age at end of observation.

Primary Analysis; HR (CI)	pre 1/1/2015; HR (CI)	Age Outlier Removed; HR (CI)	Age at end < 100; HR (CI)	At least 1 prior visit; HR (CI)
1.13 (1.10, 1.15)	1.10 (1.07, 1.12)	1.13 (1.10, 1.15)	1.13 (1.11, 1.16)	1.11 (1.08, 1.14)

Due to possible bias from lack of observation of deaths, which would result in individuals being "followed" indefinitely, we performed sensitivity analyses based on the age at the end of observation by removing outliers ($\pm$3*standard deviation; Age Outlier removed) and restricting to those < 100 years old at the end of observation time (Age at end < 100). We also performed a sensitivity analyses restricting to individuals seen before 1/1/2015 (pre 1/1/2015) as more individuals were seen in later years but due to the short follow up time would have had a much lower probability of death. Finally, we examined a sensitivity analysis restricting to individuals with at least one prior visit (At least 1 prior visit). No differences were seen across any of the sensitivity analyses. The full adjustment model was used. Also shown is are the results from the primary analysis for comparison.

* CI = 95% confidence interval; HR = hazard ratio

Table S4. Clinical covariates for individuals with diastolic and systolic heart failure.

	Diastolic HF (N = 6385)			Systolic HF (N = 7120)		
	Mean	SD	IQR	Mean	SD	IQR
Age (y)	69.7	14.0	20.7	64.6	15.1	21.7
Follow-up time (y)	4.3	3.12	4.07	3.69	2.69	3.06
Distance to monitor (km)	20.0	14.3	19.9	22.0	15.1	21.1
Monitor PM2.5 (ug/m3)	10.2	2.14	3.54	9.69	1.90	3.00
Model PM2.5 (ug/m3)	10.4	1.72	2.62	9.95	1.45	1.40
Households below federal poverty line (%)	18.0	14.9	18.6	18.5	14.8	19.5
Median Home Value ($)	190260	110753	122100	173793	105746	111900
Median Household Income ($)	54092	26870	30996	51707	26053	29709
Urbanicity (%)	64.1	41.7	83.6	61.1	42.5	94.8
Households receiving public assistance (%)	1.98	3.06	2.91	2.00	2.97	2.91
	N	%		N	%	
African American	1987	31.1		2177	30.6	
European American	3982	62.4		4341	61.0	
Other Race	416	6.5		602	8.5	
Male	2285	35.8		4379	61.5	
Within 30 km of monitor	4992	78.2		5255	73.8	
Death (All Cause)	1465	22.9		1062	14.9	
Type 2 diabetes	2574	40.3		2004	28.1	
IHD	3833	60.0		4214	59.2	
COPD	2786	43.6		2159	30.3	
PAD	2677	41.9		2226	31.3	
Hypertension	5315	83.2		4714	66.2	
Dyslipidemia	5303	83.1		5177	72.7	

BMI = body mass index; COPD = chronic obstructive pulmonary disorder; CVD = cardiovascular disease; IHD = Ischemic heart disease; PAD = peripheral arterial disease; SD = standard deviation

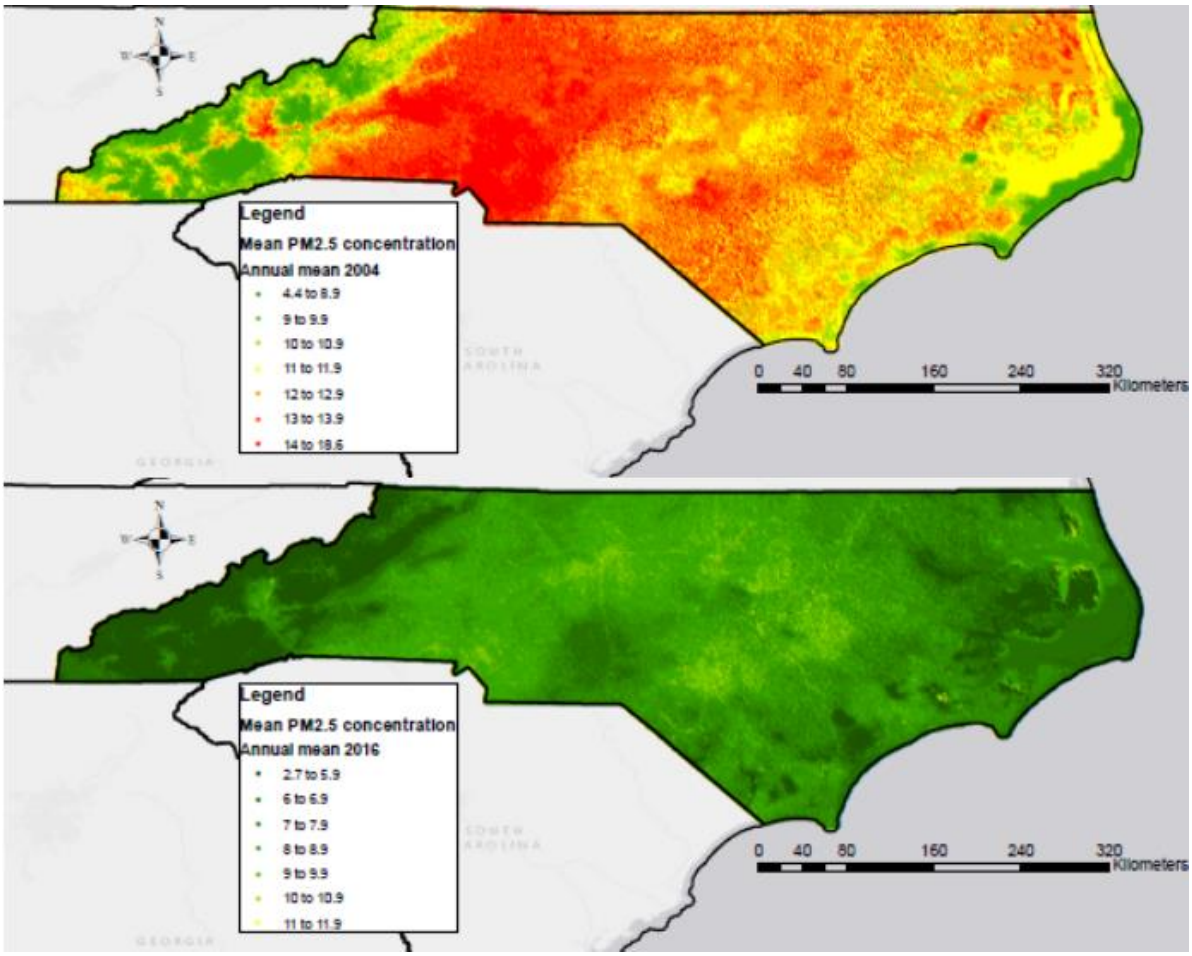
Table S5. Associations between all-cause mortality and binary $PM_{2.5}$ cutoff.

Stratification	Model	HR	LCI	UCI	N	N deaths
None	basic	1.79	1.68	1.91	23302	4496
None	full	1.77	1.66	1.89	23012	4445
Males	basic	1.76	1.61	1.93	11224	2229
Males	full	1.75	1.60	1.92	11075	2208
Females	basic	1.83	1.67	2.00	12078	2267
Females	full	1.79	1.64	1.96	11937	2237
European Americans	basic	1.78	1.64	1.92	14216	2953
European Americans	full	1.76	1.63	1.91	14090	2926
African Americans	basic	1.81	1.61	2.03	7063	1322
African Americans	full	1.76	1.57	1.98	6934	1300
<50 y	basic	2.03	1.63	2.51	3253	383
<50 y	full	1.96	1.58	2.44	3182	376
50-65 y	basic	1.74	1.53	1.97	6943	1126
50-65 y	full	1.72	1.51	1.96	6825	1110
≥65 y	basic	1.78	1.65	1.93	13106	2987
≥65 y	full	1.76	1.63	1.91	13005	2959
Type 2 diabetes	basic	1.80	1.63	1.99	7853	1768
Type 2 diabetes	full	1.77	1.60	1.96	7741	1743
PAD	basic	1.63	1.47	1.80	8195	1767
PAD	full	1.63	1.47	1.80	8100	1749
COPD	basic	1.83	1.66	2.02	8293	1772
COPD	full	1.81	1.64	2.00	8177	1747
Dyslipidemia	basic	1.78	1.65	1.91	17215	3402
Dyslipidemia	full	1.75	1.63	1.88	17017	3367
Hypertension	basic	1.78	1.65	1.91	17027	3482
Hypertension	full	1.75	1.63	1.88	16815	3442
IHD	basic	1.79	1.66	1.94	13260	2871
IHD	full	1.77	1.64	1.92	13093	2842

Associations are given based on a binary indicator for annual average $PM_{2.5}$ exposure < 12 $\mu g/m^3$ (the current EPA annual average standard for $PM_{2.5}$). Associations were stratified on sex, race, Hispanic ethnicity, age at HF diagnosis, exposure level, and pre-exiting co-morbidities.

* COPD = chronic obstructive pulmonary disorder; HR = hazard ratio; IHD = Ischemic heart disease; LCI = lower 95% confidence interval; PAD = peripheral arterial disease; UCI upper 95% confidence interval

Figure S1. $PM_{2.5}$ annual average concentration in 2004 & 2016 using modeled estimates at 1x1 km resolution.



Annual $PM_{2.5}$ concentration in NC for the years 2004 (top) and 2016 (bottom).

Figure S2. $PM_{2.5}$ concentration-response curve generated using ground-based $PM_{2.5}$ monitoring network.