

Air Pollution, Clustering of Particulate Matter Components, and Breast Cancer in the Sister Study: A U.S.-Wide Cohort

Alexandra J. White,¹ Joshua P. Keller,² Shanshan Zhao,³ Rachel Carroll,⁴ Joel D. Kaufman,⁵ and Dale P. Sandler¹

¹Epidemiology Branch, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), Research Triangle Park, North Carolina, USA

²Department of Statistics, Colorado State University, Fort Collins, Colorado, USA

³Biostatistics Branch, NIEHS, NIH, DHHS, Research Triangle Park, North Carolina, USA

⁴Department of Mathematics and Statistics, University of North Carolina at Wilmington, North Carolina, USA

⁵Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA

BACKGROUND: Particulate matter (PM) is a complex mixture. Geographic variations in PM may explain the lack of consistent associations with breast cancer.

OBJECTIVE: We aimed to evaluate the relationship between air pollution, PM components, and breast cancer risk in a United States-wide prospective cohort.

METHODS: We estimated annual average ambient residential levels of particulate matter <2.5 μm and <10 μm in aerodynamic diameter (PM_{2.5} and PM₁₀, respectively) and nitrogen dioxide (NO₂) using land-use regression for 47,433 Sister Study participants (breast cancer-free women with a sister with breast cancer) living in the contiguous United States. Cox proportional hazards regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for risk associated with an interquartile range (IQR) increase in pollutants. Predictive *k*-means were used to assign participants to clusters derived from PM_{2.5} component profiles to evaluate the impact of heterogeneity in the PM_{2.5} mixture. For PM_{2.5}, we investigated effect measure modification by component cluster membership and by geographic region without regard to air pollution mixture.

RESULTS: During follow-up (mean = 8.4 y), 2,225 invasive and 623 ductal carcinoma *in situ* (DCIS) cases were identified. PM_{2.5} and NO₂ were associated with breast cancer overall [HR = 1.05 (95% CI: 0.99, 1.11) and 1.06 (95% CI: 1.02, 1.11), respectively] and with DCIS but not with invasive cancer. Invasive breast cancer was associated with PM_{2.5} only in the Western United States [HR = 1.14 (95% CI: 1.02, 1.27)] and NO₂ only in the Southern United States [HR = 1.16 (95% CI: 1.01, 1.33)]. PM_{2.5} was associated with a higher risk of invasive breast cancer among two of seven identified composition-based clusters. A higher risk was observed [HR = 1.25 (95% CI: 0.97, 1.60)] in a California-based cluster characterized by low S and high Na and nitrate (NO₃⁻) fractions and for another Western United States cluster [HR = 1.60 (95% CI: 0.90, 2.85)], characterized by high fractions of Si, Ca, K, and Al.

CONCLUSION: Air pollution measures were related to both invasive breast cancer and DCIS within certain geographic regions and PM component clusters. <https://doi.org/10.1289/EHP5131>

Introduction

Air pollution is classified by the International Agency for Research on Cancer (IARC) as a Group 1 carcinogen (Loomis et al. 2013), consistent with the epidemiologic evidence for the role of air pollution in lung cancer incidence (Hamra et al. 2015). However, less is known about the association between air pollution and breast cancer. Air pollution exposure is widespread and thus has the potential to have a substantial impact on the incidence of breast cancer, which is the most common cancer diagnosed among women in the United States (Siegel et al. 2019).

Air pollution contains many carcinogens and other compounds that may act as endocrine disruptors—including polycyclic aromatic hydrocarbons (PAHs), metals, and benzene—which may influence breast cancer risk. Ecologic studies suggest that breast cancer risk is elevated in urban areas with higher air pollution in comparison with rural areas (Chen and Bina 2012; Wei et al. 2012). Some population studies have reported associations

between air pollution and breast cancer, as reviewed by White et al. 2018, especially in studies that consider markers of traffic-related pollution such as nitrogen dioxide (NO₂), nitrogen oxides (NO_x), and PAH exposure (Bonner et al. 2005; Hystad et al. 2015; Mordukhovich et al. 2016; Nie et al. 2007; Reding et al. 2015). In the Sister Study cohort, Reding et al. (2015) reported a modest association between residential NO₂ levels and risk of estrogen and progesterone receptor-positive (ER⁺PR⁺) breast cancer. However, associations with measures of particulate matter (PM) <2.5 μm and <10 μm in aerodynamic diameter (PM_{2.5} and PM₁₀, respectively) have not been consistently observed (Andersen et al. 2017a, 2017b; Hart et al. 2016; Reding et al. 2015; Villeneuve et al. 2018).

Fine particulate matter (PM_{2.5}) is a complex mixture that varies in composition geographically due to varying sources, differences in meteorology, and other factors (Bell et al. 2007). Regional differences in particulate matter have been shown to modify the association with breast density, an important predictor of breast cancer risk (DuPre et al. 2017). Associations between PM_{2.5} and health effects such as blood pressure (Keller et al. 2017), cardiovascular disease (Brook et al. 2010), and mortality (Franklin et al. 2008) have been shown to vary significantly by PM_{2.5} component profiles. In this report, we have extended our prior research on the relationship between air pollutants and breast cancer risk (Reding et al. 2015) with additional years of follow-up and case accrual and expanded this work to include consideration of effect measure modification by PM_{2.5} components and breast cancer risk using predictive *k*-means clusters (Keller et al. 2017). We hypothesized that air pollution would be related to breast cancer risk and that associations for PM_{2.5} would vary by PM_{2.5} component cluster. Breast cancer is a heterogeneous disease (Polyak 2011). Associations with established breast cancer risk factors have been shown to vary by hormone receptor

Address correspondence to Alexandra J. White, Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709-2233 USA. Telephone: (984) 287-3713. Email: Alexandra.white@nih.gov

Supplemental Material is available online (<https://doi.org/10.1289/EHP5131>).

The authors declare they have no actual or potential competing financial interests.

Received 31 January 2019; Revised 17 September 2019; Accepted 17 September 2019; Published 9 October 2019.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

status [often defined by the presence or absence of the estrogen receptor (ER) and progesterone receptor (PR)] (Anderson et al. 2014) as well as by menopausal status at diagnosis (White et al. 2015). In addition, risk factors may vary by whether the tumor is invasive or ductal carcinoma *in situ* (DCIS) (Barclay et al. 1997). Previous research on the association between air pollution and breast cancer has been inconclusive on whether associations vary by these different outcome classifications; therefore, we also evaluated the risk associated with air pollutant exposure considering these different outcome definitions.

Methods

Study Population

The Sister Study is a nationwide prospective cohort designed to investigate environmental and lifestyle risk factors for breast cancer (Sandler et al. 2017). During 2003–2009, 50,884 women in the United States and Puerto Rico were recruited through a multimedia campaign. Women were eligible if they were between 35 and 74 y of age and had a sister who had been diagnosed with breast cancer but had no history of breast cancer themselves. At baseline, study participants completed an extensive computer-assisted baseline telephone questionnaire that collected information on each study participant's demographics, medical and family history, and reproductive and lifestyle factors including information on their baseline residential characteristics. All participants provided signed informed consent, and the Sister Study was approved by the institutional review boards of the National Institute of Environmental Sciences, National Institutes of Health, and the Copernicus Group. This study relied on Sister Study Data Release 6.0, which included follow-up data through 15 September 2016. For this analysis, only women living in the contiguous United States were eligible ($n = 49,771$).

Outcome Classification

Sister Study participants are contacted annually for health updates, including for information on any incident breast cancer diagnoses. Participants additionally complete detailed follow-up questionnaires every 2–3 y to update lifestyle and risk factor information and to report any other health updates. Response rates have remained over 90% (i.e., 91–96%) throughout follow-up. We obtained medical records and pathology reports, from which tumor receptor information was obtained. Currently, over 80% of breast cancer diagnoses have been confirmed through medical records. Agreement between medical records and self-report of breast cancer and tumor characteristics is very high (D'Aloisio et al. 2017), with a positive predictive value over 99% for breast cancer overall. Invasive breast cancer and DCIS combined was the main outcome of interest *a priori*; however, we explored heterogeneity in the outcome by invasive versus DCIS, combined ER/PR status, and menopausal status at diagnosis. We excluded women with a breast cancer diagnosis prior to completion of all baseline data collection or an unknown time of diagnosis ($n = 62$).

Exposure Classification

As previously described (Reding et al. 2015), air pollution measures (PM_{2.5}, PM₁₀, and NO₂) were estimated for Sister Study participants based on the annual average concentrations at their addresses during the 12 months prior to enrollment, as derived using monitoring data from 2006 (for PM_{2.5} and NO₂) and 2000 (for PM₁₀). Annual averages of air pollution concentration were estimated at each participant's home using a validated regionalized universal kriging model with spatial smoothing, which incorporated information from regulatory monitors and a large

number of geographic covariates, including some derived from satellite observations, as previously described (Sampson et al. 2013; Young et al. 2016). NO₂ estimates could not be obtained for $n = 69$ participants whose addresses could not be geocoded or for locations in which there was incomplete satellite coverage.

For the PM_{2.5} component analysis, data were obtained from 130 U.S. EPA Air Quality System monitoring locations in 2010 that measured mass concentrations for 22 PM_{2.5} component species [elemental carbon (EC), organic carbon (OC), nitrate (NO₃⁻), sulfate (SO₄²⁻), Al, As, Br, Cd, Ca, Co, Cr, Cu, Fe, K, Mn, Na, S, Si, Se, Ni, V, and Zn]. Mass concentrations were converted to mass fractions by dividing the annual average of each species by the annual average PM_{2.5} at that location. The mass fractions were log transformed.

Statistical Analysis

We first evaluated the association between an interquartile range (IQR) increase in air pollutants in relation to incident breast cancer using Cox proportional hazards model to estimate hazard ratios (HRs) and 95% confidence intervals (CIs). The time scale for the Cox model was age, and the women were followed from age at study entry until age at breast cancer diagnosis or age at the end of follow-up, with censoring for death or loss to follow-up. We tested for deviations from the proportional hazards assumption by using likelihood ratio tests to compare models with and without interaction terms for air pollutants and time.

We considered whether associations varied for invasive breast cancer versus DCIS, whether the cancer was diagnosed pre- versus postmenopause, and by tumor subtype (defined using combined ER and PR status). In models evaluating the association for premenopausal breast cancer, we censored women at age at menopause. For postmenopausal breast cancer, women entered the Cox model at the age at which they enrolled in the study or at their age of menopause, whichever was later. For tumor subtype analyses, women were censored if they were diagnosed with another subtype. For example, when the outcome of interest was ER- and PR-positive (ER⁺ PR⁺) breast cancer, women who were diagnosed with ER- or PR-negative breast cancer were censored at their age of diagnosis.

To assess the impact of PM_{2.5} composition on breast cancer risk, we evaluated associations between 2010 PM_{2.5} and incident breast cancer stratified by PM_{2.5} components using previously developed predictive *k*-means clusters (Keller et al. 2017). PM_{2.5} component information was not available for this study for 2006, the year used in our primary analysis described above, so this analysis used exposure estimates from 2010, the year for which component data were available. Clustering is a method of dimension reduction that can be used to partition multi-pollutant observations into a prespecified number (*k*) of clusters. The covariate-adaptive approach used here clustered monitor locations using the multidimensional component mass fractions while also allowing the geographic covariates at each location to influence cluster membership, resulting in groups of monitor locations with similar component profiles. Cluster membership was then predicted for each study participant based on the geographic covariates at their residential location. Participants were assigned to the cluster to which they had the highest probability of belonging. This covariate-adaptive clustering method has been shown to provide better predictive accuracy and power for detecting effect modification than using traditional *k*-means clustering, which does not incorporate geographic covariates in cluster identification. The number of clusters and the covariates were selected by 10-fold cross-validation. The final selected model had eight clusters, as detailed previously (Keller et al. 2017). Cluster 8 (to which $n = 74$ participants belonged) was not included in the analyses due to its small

sample size. For this study, we estimated the association between 2010 PM_{2.5} and breast cancer risk stratified by cluster (Figure 1). We tested for effect modification using a likelihood ratio test to compare models with and without interaction terms between PM_{2.5} and indicator variables for the clusters.

The covariate adjustment set included age, race/ethnicity (non-Hispanic white, other), education (high school degree/equivalent or less, some college, 4-y degree or higher), smoking status (never, former, current), and menopausal hormone therapy (ever, never) to be consistent with our prior publication (Reding et al. 2015). As a secondary analysis, we included additional confounders including household income, census-tract income, marital status, parity, and body mass index (BMI). We evaluated effect measure modification by years spent living at the home (<10 y, ≥10 y), census-defined geographic region (Northeast, Midwest, South, West based on state of residence), degree of family history of breast cancer (1 first-degree family member, >1 first-degree family member), BMI (<25 kg/m², 25 to <30 kg/m², ≥30 kg/m²), and postmenopausal hormone use (ever, never) by including a cross-product term in the Cox model and using a likelihood ratio test. Given the correlation between region and PM_{2.5} component clusters, in analyses stratified by region we also considered adjustment for PM_{2.5} cluster and in analyses stratified by PM_{2.5} cluster we also considered adjustment for region. To evaluate whether differences by region were explained by other factors, we considered the inclusion of multiple additional interaction terms within a single model (between air pollutant and region, air pollutant and cluster, air pollutant and BMI, and air pollutant and education). Covariates had <4% missing data; therefore, we conducted a complete case analysis (excluding

those with missing values for the adjustment covariates), with a resulting sample size of *n* = 47,433.

All analyses were conducted using SAS (version 9.4; SAS Institute Inc.).

Results

During an average of 8.4 y of follow-up, there were 2,852 incident breast cancer cases (2,225 invasive and 623 DCIS). Study participant baseline characteristics have been previously published (Sandler et al. 2017). Briefly, the median age at enrollment was 55.6 y. Women in the study are predominately non-Hispanic white (83.7%), reported being married or living as married (74.7%), and over half have a bachelor's degree or higher. The Sister Study includes participants from each of the contiguous states, with representation ranging from 0.2% participants from Wyoming to 8.5% from California. Participant characteristics by geographic region are displayed in Table 1.

An IQR increase in NO₂ (5.8 ppb) was associated with breast cancer risk overall [HR = 1.08 (95% CI: 1.03, 1.13)] (Table 2). We observed substantial heterogeneity when stratifying by invasive disease versus DCIS and therefore show these results separately. This association was stronger for DCIS [HR = 1.23 (95% CI: 1.12, 1.35)] than for invasive breast cancer [HR = 1.02 (95% CI: 0.96, 1.07)]. Similarly, PM_{2.5} (IQR = 3.6 μg/m³) was positively associated with DCIS incidence [HR = 1.16 (95% CI: 1.02, 1.31)] but not invasive breast cancer [HR = 1.03 (95% CI: 0.96, 1.09)]. No elevated HRs were observed in relation to PM₁₀ (IQR = 5.8 μg/m³). Further adjustment for other known and

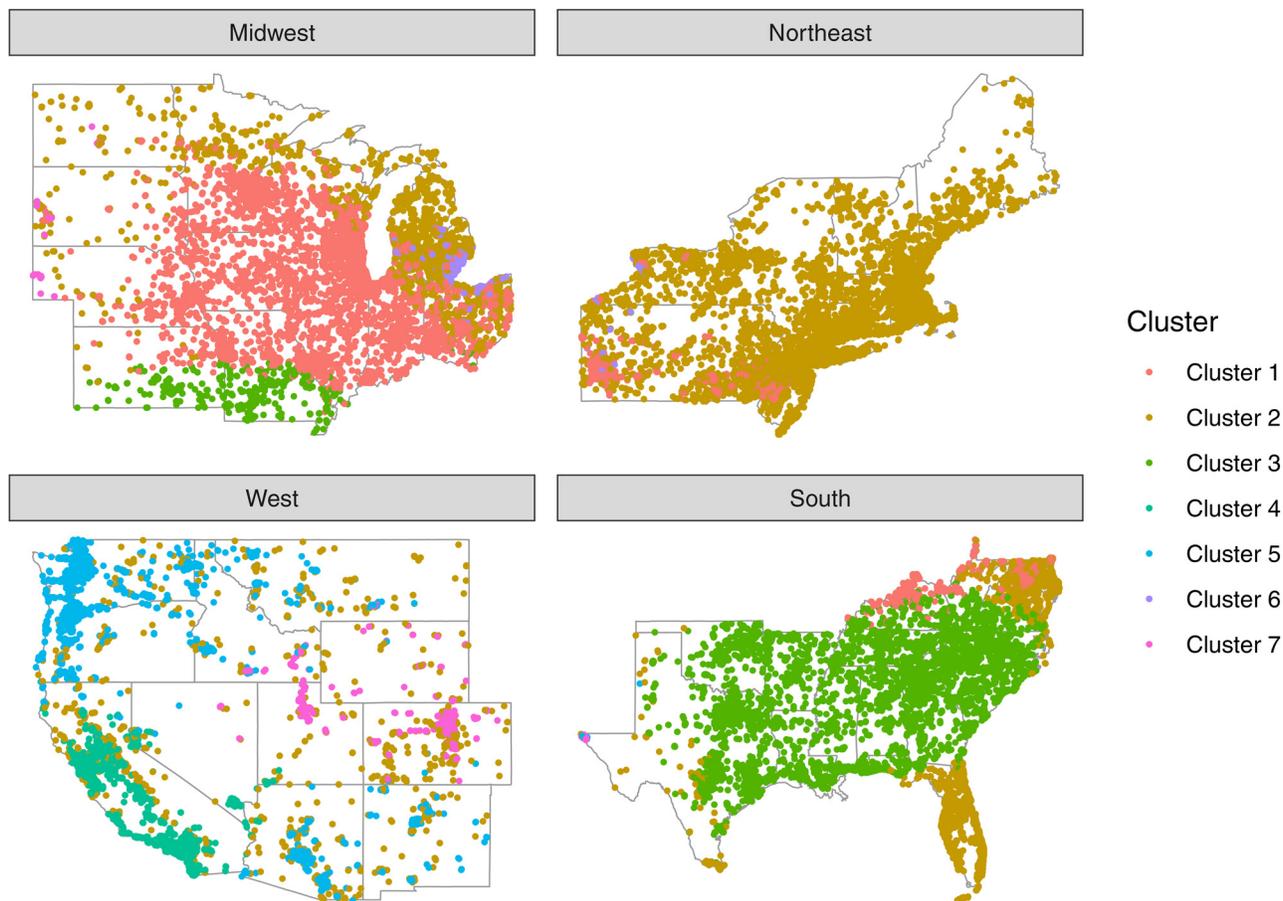


Figure 1. Predicted PM_{2.5} component cluster membership by geographic region (jittered to protect confidentiality), Sister Study, 2003–2009. Figure adapted from Keller et al. (2017). PM_{2.5}, particulate matter <2.5 μm in aerodynamic diameter.

Table 1. Study population characteristics by geographic region, Sister Study, 2003–2009.

Characteristic	Geographic Region							
	Midwest (n = 13,047)		Northeast (n = 8,082)		South (n = 15,960)		West (n = 10,344)	
	n	(%)	n	(%)	n	(%)	n	(%)
Age at baseline (y)								
≤45	1,845	14	1,139	14	2,123	13	1,279	12
46–49	2,097	16	1,344	17	2,429	15	1,517	15
50–54	2,581	20	1,594	20	3,179	20	1,960	19
55–59	2,620	20	1,535	19	3,147	20	2,094	20
60–64	1,877	14	1,167	14	2,454	15	1,586	15
≥65	2,027	16	1,303	16	2,628	16	1,908	18
Race								
Non-Hispanic white	12,000	89	7,365	91	12,000	77	9,006	87
Other	1,380	11	717	9	3,669	23	1,338	13
Education								
≥4-y college degree	6,132	47	4,469	55	8,140	51	5,491	53
≤High school degree or equivalent	2,337	18	1,197	15	2,354	15	1,247	12
Some college/technical school	4,578	35	2,416	30	5,466	34	3,606	35
Household income								
\$50,000–< \$100,000	5,781	44	3,280	41	6,369	40	4,118	40
<\$50,000	3,403	26	1,741	22	4,130	26	2,467	24
>\$100,000	3,863	30	3,061	38	5,461	34	3,759	36
Census tract–level income								
\$50,000–< \$100,000	8,137	62	5,115	63	8,317	52	6,295	61
<\$50,000	4,096	31	1,983	25	6,444	40	3,105	30
>\$100,000	814	6	984	12	1,199	8	944	9
Smoking status								
Never smoker	1,202	9	629	8	1,445	9	678	7
Former smoker	7,473	57	4,045	50	8,965	56	5,983	58
Current smoker	4,372	34	3,408	42	5,550	35	3,683	36
Marital status								
Married or living as married	10,000	77	6,022	75	12,000	73	7,783	75
Never married	650	5	543	7	888	6	498	5
Widowed, divorced, or separated	2,388	18	1,517	19	3,480	22	2,063	20
BMI (kg/m ²)								
<24.9	4,207	32	2,494	31	5,067	32	3,210	31
25–29.9	4,609	35	3,342	41	5,797	36	4,379	42
≥30	4,231	32	2,246	28	5,096	32	2,755	27
Ever HRT								
No	7,332	56	5,294	66	8,327	52	5,101	49
Yes	5,715	44	2,788	34	7,633	48	5,243	51
Parity								
None	2,032	16	1,570	19	2,920	18	2,176	21
1	1,659	13	1,124	14	2,662	17	1,553	15
2–3	7,754	59	4,661	58	8,997	56	5,594	54
>3	1,602	12	727	9	1,381	9	1,021	10
Mammographic screening in last 24 months								
No	936	8	493	6	1,247	8	764	8
Yes	11,000	92	7,159	94	14,000	92	8,951	92
Missing	746	—	430	—	1,116	—	629	—
Family history of breast cancer								
1 first-degree relative	9,476	73	6,031	75	12,000	73	7,539	73
>1 first-degree relative	3,571	27	2,051	25	4,239	27	2,805	27
Baseline menopausal status								
Postmenopausal	8,526	65	5,127	63	11,000	68	7,007	68
Premenopausal	4,468	34	2,927	36	5,038	32	3,295	32
Missing	53	—	28	—	46	—	42	—
Breast cancer characteristics								
Invasive	613	5	348	4	758	5	506	5
DCIS	165	1	113	1	203	1	142	1
ER ⁺ invasive	442	3	264	3	553	3	385	4
ER [−] invasive	93	1	42	1	90	1	53	1

Note: —, not applicable; BMI, body mass index; DCIS, ductal carcinoma *in situ*; ER, estrogen receptor; HRT, hormone replacement therapy.

established breast cancer risk factors and other markers of socioeconomic status, including household income, census-tract income, marital status, parity, and BMI, did not materially change the point estimates.

An IQR increase in NO₂ was inversely associated with ER[−]PR[−] breast cancer [HR = 0.87 (95% CI: 0.73, 1.04)] but not with ER⁺PR⁺ breast cancer [HR = 1.03 (95% CI: 0.95, 1.10)] (Table 3). Associations for PM_{2.5} and PM₁₀ did not vary by ER/

PR status of the tumor. We did not observe notable heterogeneity in the observed associations by menopausal status at diagnosis (see Table S1).

Associations for invasive breast cancer and exposure to PM_{2.5} ($p_{\text{heterogeneity}} = 0.04$), PM₁₀ ($p_{\text{heterogeneity}} = 0.04$), and NO₂ ($p_{\text{heterogeneity}} = 0.05$) all varied notably by geographic region (Table 4). An IQR increase in PM_{2.5} [HR = 1.14 (95% CI: 1.02, 1.27)] was associated with invasive breast cancer in women

Table 2. Air pollutants and risk of invasive breast cancer and DCIS, Sister Study, 2003–2009.

Air pollutant ^a	Overall breast cancer				Invasive breast cancer				DCIS			
	Cases (n)	Age-adjusted HR (95% CI)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c		
PM _{2.5}	2,820	1.05 (1.00, 1.11)	1.05 (0.99, 1.11)	1.04 (0.98, 1.10)	2,206	1.03 (0.96, 1.09)	1.02 (0.95, 1.08)	610	1.16 (1.02, 1.31)	1.15 (1.02, 1.30)		
PM ₁₀	2,820	1.01 (0.97, 1.05)	1.01 (0.97, 1.05)	1.01 (0.97, 1.05)	2,206	1.00 (0.96, 1.04)	1.00 (0.95, 1.04)	610	1.06 (0.99, 1.15)	1.06 (0.99, 1.15)		
NO ₂	2,817	1.08 (1.03, 1.13)	1.06 (1.02, 1.11)	1.06 (1.01, 1.11)	2,203	1.02 (0.96, 1.07)	1.01 (0.96, 1.07)	610	1.23 (1.12, 1.35)	1.23 (1.12, 1.36)		

Note: CI, confidence interval; DCIS, ductal carcinoma *in situ*; HR, hazard ratio; IQR, interquartile range; NO₂, nitrogen dioxide; PM_{2.5}, particulate matter <2.5 μm in aerodynamic diameter; PM₁₀, particulate matter <10 μm in aerodynamic diameter.

^aHR for a unit increase in the IQR difference: PM_{2.5} = 3.6 μg/m³, PM₁₀ = 5.8 μg/m³, and NO₂ = 5.8 ppb.

^bAdjusted for age, race, education, smoking status, and postmenopausal hormone use.

^cAdjusted for age, race, education, income, census tract-level income, marital status, parity, smoking status, body mass index, and postmenopausal hormone use.

residing in the West but not other geographic regions [Northeast HR = 0.89 (95% CI: 0.73, 1.07); Midwest HR = 0.93 (95% CI: 0.81, 1.08), South HR = 1.03 (95% CI: 0.90, 1.17)]. A similar trend, with a slightly higher HR among women in the Western United States was observed for PM₁₀ exposure. An IQR increase in NO₂ was similarly associated with breast cancer among women living in the West [HR = 1.09 (95% CI: 0.99, 1.21)] as well as for women residing in the South [HR = 1.16 (95% CI: 1.01, 1.33)]. For DCIS, in general we observed associations to be more pronounced in women living in the Northeast or the Midwest. For example, for an IQR increase in PM_{2.5}, we observed an HR = 1.35 (95% CI: 0.97, 1.88) for women in the Northeast and HR = 1.68 (95% CI: 1.21, 2.34) for women in the Midwest. The pattern was similar for PM₁₀ ($p_{\text{heterogeneity}} = 0.01$). For NO₂, risk of DCIS also varied by region ($p_{\text{heterogeneity}} = 0.01$), with the highest HRs observed in the Midwest [HR = 1.73 (95% CI: 1.39, 2.14)]. These associations persisted with further covariate adjustment and when including PM_{2.5} component clusters in the model. These associations were also robust to the inclusion of additional interaction terms with cluster, BMI, and education in the model (see Table S2).

Overall, the associations for PM_{2.5} using 2010 air pollution estimates (2010 IQR = 2.9 μg/m³) were similar to those from our main results using data from 2006 [e.g., 2010 invasive HR = 1.01 (95% CI: 0.95, 1.07) vs. 2006 invasive HR = 1.03 (95% CI: 0.96, 1.09)] (Table 5). Consistent with the results stratified by geographic region, invasive breast cancer risk also varied by PM_{2.5} component cluster ($p_{\text{heterogeneity}} = 0.3$) (Table 5). Specifically, we observed an elevated risk of invasive breast cancer associated with PM_{2.5} exposure for both Cluster 4 (California; Figure 1) and Cluster 7 (West; Figure 1) but no increase in risk for women in any of the other clusters. The California monitors were captured in Cluster 4 (Figure 1), which was characterized by having low S fractions and large fractions of Na and NO₃⁻ (Figure 2), indicating exposure to marine aerosols and agricultural emissions (Keller et al. 2017). For an IQR increase in PM_{2.5} for women who were assigned to Cluster 4, we observed a 25% higher risk of invasive breast cancer [HR = 1.25 (95% CI: 0.97, 1.60)]. Cluster 7 was also centered in the Western United States (Figure 1), and was defined by high fractions of Si, Ca, K, and Al (Figure 2), consistent with the surface soil in this geographic region (Shacklette and Boerngen 1984). For women in Cluster 7, we also observed an elevated risk associated with an IQR increase in PM_{2.5} [HR = 1.60 (95% CI: 0.90, 2.85)], but the estimate for this cluster was imprecise due to the small number of cases ($n = 59$). These associations remained similar with further adjustment for additional covariates and inclusion of geographic region in the adjustment set.

For DCIS, although sample sizes were small, there was less evidence of risk heterogeneity by cluster ($p_{\text{heterogeneity}} = 0.9$) (Table 5). Across the clusters, PM_{2.5} was positively associated with DCIS in all but Cluster 7. For example, a higher risk of DCIS in relation to an IQR increase in PM_{2.5} was observed for women in Cluster 1 [HR = 1.38 (95% CI: 1.02, 1.86)] and Cluster 2 [HR = 1.37 (95% CI: 1.03, 1.83)]. Cluster 1 is in the Midwest and Mid-Atlantic region (Figure 1) with above-average NO₃⁻ and SO₄²⁻ (Figure 2), which is consistent with high ambient ammonia levels from agriculture. Cluster 2 is in the Northeast (Figure 1) and is characterized by higher fractions of Cd, V, and Ni (Figure 2). Elevated HRs, but with wide CIs, were also observed for women in Cluster 3 [HR = 1.22 (95% CI: 0.75, 1.96)], Cluster 4 [HR = 1.33 (95% CI: 0.80, 2.22)], Cluster 5 [HR = 1.18 (95% CI: 0.67, 2.07)], and Cluster 6 [HR = 1.22 (95% CI: 0.35, 4.26)] in relation to DCIS.

We observed no significant effect measure modification of the associations between any of the air pollutants and breast cancer risk by time spent living at the baseline residence (see Table S3).

Table 3. Air pollutants and risk of invasive ER⁺PR⁺ and ER⁻PR⁻ breast cancer, Sister Study, 2003–2009.

Air pollutant ^a	ER ⁺ PR ⁺ invasive			ER ⁻ PR ⁻ invasive		
	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c
PM _{2.5}	1,347	1.01 (0.93, 1.10)	1.00 (0.92, 1.08)	253	0.94 (0.78, 1.13)	0.95 (0.79, 1.14)
PM ₁₀	1,347	1.00 (0.95, 1.06)	1.00 (0.95, 1.06)	253	0.89 (0.78, 1.02)	0.89 (0.78, 1.02)
NO ₂	1,346	1.03 (0.97, 1.11)	1.03 (0.95, 1.10)	253	0.86 (0.72, 1.02)	0.87 (0.73, 1.04)

Note: CI, confidence interval; ER, estrogen receptor; HR, hazard ratio; IQR, interquartile range; NO₂, nitrogen dioxide; PM_{2.5}, particulate matter <2.5 μm in aerodynamic diameter; PM₁₀, particulate matter <10 μm in aerodynamic diameter; PR, progesterone receptor.

^aHR for a unit increase in the IQR difference: PM_{2.5} = 3.6 μg/m³, and PM₁₀ = 5.8 μg/m³, NO₂ = 5.8 ppb.

^bAdjusted for age, race, education, smoking status, and postmenopausal hormone use.

^cAdjusted for age, race, education, income, census tract–level income, marital status, parity, smoking status, body mass index, and postmenopausal hormone use.

However, we did note an elevated HR for invasive breast cancer was observed for PM_{2.5} in women who lived in their residences for ≥10 y [HR = 1.07 (95% CI: 0.98, 1.17)]. We observed modification by obesity; women who had a BMI ≥30 kg/m² had a higher risk of invasive breast cancer associated with PM_{2.5} [HR = 1.19 (95% CI: 1.06, 1.34), *p*_{heterogeneity} = 0.02], and NO₂ [HR = 1.11 (95% CI: 1.01, 1.21), *p*_{heterogeneity} = 0.1] (see Table S4). We observed no significant effect measure modification of the associations for air pollutants and breast cancer risk by extent of breast cancer family history or hormone therapy use (see Tables S5 and S6). As expected, there was substantial overlap between clusters and geographic region (see Table S7).

Discussion

In this large, U.S.-wide prospective cohort study, we evaluated the association between air pollutants and breast cancer risk and demonstrated that air pollution levels were related to both invasive breast cancer and DCIS in certain geographic regions. For example, exposure to PM_{2.5} tended to be related to invasive breast cancer risk in the Western United States, whereas for DCIS, the associations were most evident among women in the Northeast and Midwest. These results were consistent with our analysis utilizing predictive *k*-means clustering to evaluate PM_{2.5} component mixtures in relation to breast cancer risk. PM_{2.5} levels in two Western-based clusters were related to the risk of invasive

breast cancer, whereas PM_{2.5} exposure in other clusters were more strongly related to the risk of DCIS. Together, these results suggest that consideration of geographic variability in air pollution is crucial when evaluating associations with breast cancer. This is the first U.S.-based study to evaluate the relationship between PM components and breast cancer risk.

Air pollution is plausibly related to breast cancer given that it is a complex mixture containing numerous carcinogens and endocrine disruptors (Loomis et al. 2013). In breast cancer cell lines, PM has been shown to have estrogenic properties and oxidative stress–related DNA-damaging activity (Chen et al. 2013). Inhaled toxicants can reach the breast tissue (Hill and Wynder 1979) and traffic-related air pollution has been associated with aberrant DNA methylation in breast cancer–related genes measured in tumor tissue (White et al. 2016). Air pollution has also been related to higher breast density (DuPre et al. 2017; White et al. 2019c; Yaghjian et al. 2017), a marker of breast cancer risk.

Markers of traffic pollution such as NO₂, NO_x, and PAH exposure have been found to be associated with breast cancer risk (Bonner et al. 2005; Hystad et al. 2015; Mordukhovich et al. 2016; Nie et al. 2007; Reding et al. 2015), whereas results for measures of PM have been mostly null (Andersen et al. 2017a, 2017b; Hart et al. 2016; Reding et al. 2015; Villeneuve et al. 2018). However, these studies have largely not considered the impact of geographic variability or PM heterogeneity. For example, although we too saw little consistent evidence of an association

Table 4. Air pollutants and risk of invasive breast cancer and DCIS by geographic region, Sister Study, 2003–2009.

Air pollutant ^a	Region	Invasive breast cancer				DCIS			
		Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Model 3 HR (95% CI) ^d	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Model 3 HR (95% CI) ^d
PM _{2.5}	Northeast	345	0.89 (0.73, 1.07)	0.86 (0.71, 1.04)	0.79 (0.64, 0.99)	111	1.35 (0.97, 1.88)	1.36 (0.97, 1.9)	1.43 (0.97, 2.09)
	Midwest	609	0.93 (0.81, 1.08)	0.89 (0.77, 1.04)	0.91 (0.77, 1.07)	161	1.68 (1.21, 2.34)	1.64 (1.17, 2.30)	1.56 (1.08, 2.26)
	South	753	1.03 (0.90, 1.17)	1.02 (0.89, 1.16)	1.07 (0.90, 1.27)	200	1.07 (0.83, 1.39)	1.08 (0.83, 1.40)	1.09 (0.78, 1.52)
	West	499	1.14 (1.02, 1.27)	1.12 (1.00, 1.26)	1.21 (1.03, 1.43)	138	1.10 (0.89, 1.36)	1.08 (0.86, 1.35)	1.21 (0.89, 1.65)
	<i>p</i> _{heterogeneity}		0.04	0.03	0.03		0.07	0.07	0.3
PM ₁₀	Northeast	345	0.90 (0.79, 1.02)	0.88 (0.77, 1.00)	0.86 (0.76, 0.99)	111	1.13 (0.91, 1.41)	1.15 (0.91, 1.44)	1.14 (0.91, 1.44)
	Midwest	609	0.92 (0.82, 1.04)	0.91 (0.80, 1.03)	0.92 (0.80, 1.06)	161	1.55 (1.22, 1.96)	1.55 (1.22, 1.97)	1.46 (1.13, 1.90)
	South	753	0.91 (0.80, 1.03)	0.91 (0.80, 1.03)	0.91 (0.80, 1.04)	200	1.00 (0.79, 1.28)	1.02 (0.80, 1.30)	1.07 (0.82, 1.38)
	West	499	1.04 (0.98, 1.10)	1.04 (0.98, 1.10)	1.04 (0.98, 1.11)	138	1.01 (0.90, 1.13)	1.01 (0.90, 1.13)	1.00 (0.89, 1.12)
	<i>p</i> _{heterogeneity}		0.04	0.04	0.07		0.01	0.01	0.02
NO ₂	Northeast	345	0.92 (0.82, 1.03)	0.89 (0.79, 1.01)	0.89 (0.79, 1.01)	111	1.16 (0.97, 1.39)	1.19 (0.98, 1.44)	1.19 (0.98, 1.44)
	Midwest	609	1.00 (0.88, 1.14)	0.97 (0.85, 1.11)	0.99 (0.86, 1.15)	161	1.73 (1.39, 2.14)	1.72 (1.38, 2.15)	1.69 (1.33, 2.14)
	South	750	1.16 (1.01, 1.33)	1.18 (1.03, 1.37)	1.20 (1.02, 1.41)	200	1.12 (0.86, 1.45)	1.14 (0.87, 1.50)	1.04 (0.75, 1.42)
	West	499	1.09 (0.99, 1.21)	1.09 (0.98, 1.20)	1.13 (1.00, 1.26)	138	1.17 (0.97, 1.41)	1.16 (0.95, 1.41)	1.14 (0.91, 1.41)
	<i>p</i> _{heterogeneity}		0.05	0.06	0.04		0.01	0.01	0.01

Note: CI, confidence interval; DCIS, ductal carcinoma *in situ*; HR, hazard ratio; IQR, interquartile range; NO₂, nitrogen dioxide; PM_{2.5}, particulate matter <2.5 μm in aerodynamic diameter; PM₁₀, particulate matter <10 μm in aerodynamic diameter.

^aHR for a unit increase in the IQR difference: PM_{2.5} = 3.6 μg/m³, PM₁₀ = 5.8 μg/m³, and NO₂ = 5.8 ppb.

^bAdjusted for age, race, education, smoking status, and postmenopausal hormone use.

^cAdjusted for age, race, education, income, census tract–level income, marital status, parity, smoking status, body mass index, and postmenopausal hormone use.

^dAdjusted for age, race, education, income, census tract–level income, marital status, parity, smoking status, body mass index, postmenopausal hormone use, and PM_{2.5} component clusters.

Table 5. PM_{2.5}, *k*-means clusters, and risk of invasive breast cancer and DCIS, Sister Study, 2003–2009.

Air pollutant ^a	Invasive breast cancer				DCIS			
	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Model 3 HR (95% CI) ^d	Cases (n)	Model 1 HR (95% CI) ^b	Model 2 HR (95% CI) ^c	Model 3 HR (95% CI) ^d
2010 PM _{2.5}	2,206	1.01 (0.95, 1.07)	1.00 (0.94, 1.06)	—	610	1.14 (1.01, 1.27)	1.13 (1.01, 1.27)	—
By clusters ^e								
1	607	0.95 (0.81, 1.10)	0.93 (0.80, 1.09)	0.93 (0.80, 1.09)	186	1.38 (1.02, 1.86)	1.38 (1.02, 1.85)	1.38 (1.02, 1.86)
2	649	0.99 (0.86, 1.14)	0.96 (0.83, 1.12)	0.95 (0.79, 1.13)	169	1.37 (1.03, 1.83)	1.30 (0.97, 1.76)	1.28 (0.9, 1.81)
3	438	0.95 (0.74, 1.22)	0.96 (0.75, 1.24)	0.95 (0.73, 1.23)	113	1.22 (0.75, 1.96)	1.27 (0.78, 2.07)	1.27 (0.77, 2.09)
4	203	1.25 (0.97, 1.60)	1.24 (0.96, 1.60)	1.24 (0.96, 1.60)	49	1.33 (0.80, 2.22)	1.32 (0.79, 2.21)	1.32 (0.79, 2.21)
5	203	1.00 (0.74, 1.36)	1.04 (0.77, 1.42)	1.05 (0.77, 1.42)	62	1.18 (0.67, 2.07)	1.31 (0.74, 2.32)	1.31 (0.74, 2.32)
6	47	0.82 (0.36, 1.87)	0.95 (0.40, 2.29)	0.97 (0.40, 2.37)	17	1.22 (0.35, 4.26)	1.32 (0.34, 5.14)	1.36 (0.34, 5.47)
7	59	1.60 (0.90, 2.85)	1.66 (0.92, 2.99)	1.71 (0.93, 3.14)	22	0.97 (0.40, 2.38)	0.81 (0.32, 2.04)	0.88 (0.34, 2.23)
<i>p</i> _{heterogeneity}		0.3	0.3	0.3		0.9	0.9	0.9

Note: —, not applicable; CI, confidence interval; DCIS, ductal carcinoma *in situ*; HR, hazard ratio; IQR, interquartile range; NO₂, nitrogen dioxide; PM_{2.5}, particulate matter <2.5 μm in aerodynamic diameter; PM₁₀, particulate matter <10 μm in aerodynamic diameter.

^aHR for a unit increase in the IQR difference: PM_{2.5} = 2.9 μg/m³, PM₁₀ = 5.8 μg/m³, and NO₂ = 5.8 ppb.

^bAdjusted for age, race, education, smoking status, and postmenopausal hormone use.

^cAdjusted for age, race, education, income, census tract–level income, marital status, parity, smoking status, body mass index, and postmenopausal hormone use.

^dAdjusted for age, race, education, income, census tract–level income, marital status, parity, smoking status, body mass index, postmenopausal hormone use, and geographic region.

^eCluster locations are provided in Figure 1.

with PM_{2.5} or PM₁₀ and invasive breast cancer in our nationwide study population, stratifying by region elucidated significant variability in the associations.

Air pollution is a complex mixture and it is important to address the heterogeneity of this exposure and to evaluate how that may impact breast cancer risk. Only one prior study has evaluated PM components with breast cancer. In a pooled analysis of European cohorts, Andersen et al. (2017b) considered PM components individually in relation to postmenopausal breast cancer risk. They observed a higher breast cancer risk for exposure to both PM_{2.5} and PM₁₀ V and PM₁₀ Ni levels. Importantly, considering a single PM component at a time does not address the correlated nature of the PM components. To better capture this heterogeneity, we utilized predictive *k*-means clustering, which is a data reduction technique that identifies subgroups of individuals who are exposed to similar combinations of PM components. This permits the identification of PM component mixtures and consideration of how these complex mixtures influence the association between PM_{2.5} and breast cancer risk.

We observed heterogeneity by geographic region and PM_{2.5} component cluster, individually and after simultaneous adjustment, in the associations between air pollutants and breast cancer risk. Although this geographic variability has not been explicitly considered previously in relation to breast cancer, DuPre et al. (2017) observed geographic variation in that PM_{2.5} in the Nurses' Health Study was related to breast density only among participants living in the Northeast. In our study, PM_{2.5} was related to DCIS across most of the clusters despite lower power to detect associations. In contrast, PM_{2.5} was associated with invasive breast cancer only in women assigned to two Western-based clusters (Clusters 4 and 7), consistent with our regional results finding a higher risk among women living in the Western United States. Cluster 4, which encompassed the California monitors, was characterized by having low fractions of S and large fractions of Na and NO₃⁻, indicative of marine aerosols and agricultural emissions. Airborne exposure to pesticides from agricultural practices may contribute to cancer risk (Engel et al. 2005; Lee et al. 2002; Lerro et al. 2015). Cluster 7, which was more widely spread across the Western

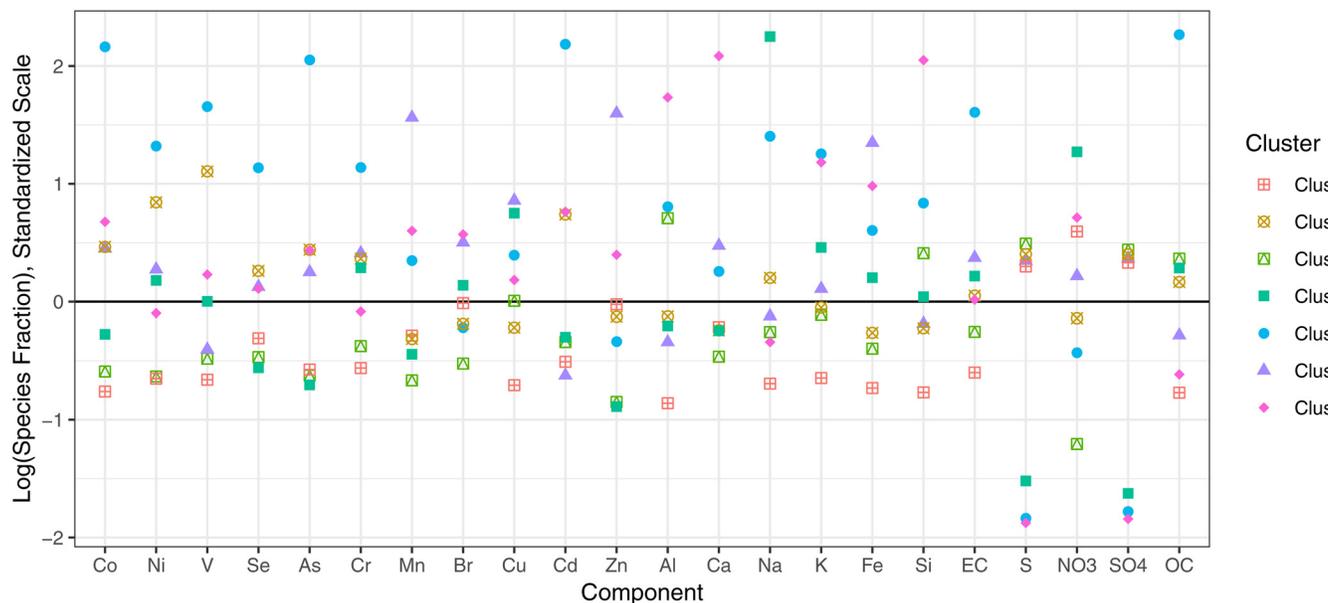


Figure 2. Relative composition by PM_{2.5} clusters. Clusters were identified using predictive *k*-means in the 2010 annual average PM_{2.5} component data. Species mass fractions were log transformed and then standardized. EC, elemental carbon; NO₃, nitrate; OC, organic carbon; PM_{2.5}, particulate matter <2.5 μm in aerodynamic diameter; SO₄, sulfate.

United States, had high fractions of Si, Ca, K, and Al, consistent with the surface soil in this region (Shacklette and Boerngen 1984). In a subset of our study population with DNA methylation data, among women in Clusters 4 and 7, PM_{2.5} was also associated with DNA methylation-based biologic age acceleration (White et al. 2019a), a marker of future breast cancer risk (Kresovich et al. 2019). These consistent findings support a role for these clusters of PM_{2.5} components in breast carcinogenesis.

Differences between overall results for invasive breast cancer and DCIS were unexpected. DCIS is generally thought to be a precursor to invasive breast cancer, and risk factor profiles for DCIS and invasive disease are similar although there are some differences (Reeves et al. 2012). However, it is possible that variation in socioeconomic status by region may have contributed to differences in access to health care that could have influenced the associations observed with DCIS, which is primarily detected by screening (Virmig et al. 2010). To address this, we further adjusted our models for risk of DCIS for individual and census tract-level socioeconomic variables, but we did not observe a change in results. It is unlikely that screening practices explain these results because over 92% of women in our study population were screened within the last 2 y. This high rate of screening may not be too surprising given that our study population consists of women with a family history of breast cancer among whom regular screening is very common. In addition, mammographic screening did not vary by geographic region, so geographic differences in screening behaviors or access cannot explain observed differences in associations by region or cluster. Despite extensive efforts to address potential residual confounding, it remains possible that there is some unaddressed confounding from other factors such as noise or other pollutants that may be driving the differences in DCIS/invasive disease risk by region. Another potential explanation is that these mixtures of pollutants simply contribute differently to breast cancer risk by stage of disease, perhaps by influencing tumor growth rate. Our results of a higher risk of DCIS in relation to air pollutants in the Northeast are consistent with results from a study of women on Long Island, New York, for whom higher vehicular traffic air pollution was similarly associated with DCIS (Mordukhovich et al. 2016).

We did not observe substantial evidence of variability in the associations of overall air pollutant exposures and breast cancer risk by menopausal status or by tumor subtype. However, a limitation of this study was that, despite our large sample size, we were unable to explore effect measure modification by cluster with consideration of tumor subtype.

We observed that invasive breast cancer risk associated with exposure to PM_{2.5} and NO₂ was higher among women with a BMI ≥ 30 kg/m², suggesting a possible synergistic relationship between obesity and air pollution. Components of air pollution, such as PAHs, are lipophilic (IARC 2010), whereas other components, such as metals, have been detected in visceral fat (Qin et al. 2010). Thus, fat tissue may serve as a possible reservoir for which the constituents of air pollution may accumulate. This finding is consistent with prior research on PAHs (Niehoff et al. 2017) and airborne metals (White et al. 2019b).

A strength of this study was the use of predictive *k*-means clustering to determine subgroups of women who were exposed to different PM_{2.5} component mixtures. Consideration of the mixture is important because PM is not a homogenous exposure and our approach permitted a more refined and nuanced exposure assessment. The predictive *k*-means approach used to identify and assign PM component clusters in the Sister Study was an unsupervised method, meaning that the clusters identified are useful for a public health-focused approach to identify existing air pollution mixtures and determine how they are related to health outcomes. However, given that breast cancer case status was not included in the

identification of these clusters, it is possible that there are some groups of pollutants that may be more strongly related to breast cancer risk that were not identified. Although these clusters incorporate 22 different PM_{2.5} components, it is possible that these clusters may be influenced by other correlated unmeasured air pollutants. In addition, the accuracy of the concentration measurements may vary for some of the PM_{2.5} components and thus may result in differential measurement error. Furthermore, we classified individuals into the cluster for which each person had the highest probability of membership, and there is uncertainty in the cluster predictions that could also lead to exposure measurement error. Finally, we cannot rule out the possibility of residual spatial confounding.

The Sister Study is a prospective cohort with extensive covariate information. A strength of this study is the use of land-use regression models with spatial smoothing to assess exposure to air pollution at the level of cohort enrollment residence. However, a limitation of this approach is that we used air pollution measures estimated around the time of enrollment in the study (on average, 8 y prior to breast cancer diagnosis). This measurement may not represent the most relevant time period of exposure with respect to breast cancer etiology. We did, however, consider duration of residence at the current residence. It is noteworthy that most results did not differ for women with <10 or ≥ 10 y at their enrollment address. It is possible that more long-term exposure, or exposure occurring during hypothesized susceptible windows of exposure including childhood (Bonner et al. 2005; Nie et al. 2007; Shmuel et al. 2017), or exposure during the reproductive time period may be more relevant.

In conclusion, in this large, prospective U.S.-wide cohort, we observed that measures of air pollution, including NO₂, PM_{2.5}, and PM₁₀, were related to both invasive and DCIS breast cancer when stratifying by geographic region. Using predictive *k*-means clusters to consider the potential modifying role of PM_{2.5} components, we observed that the risk of breast cancer varied based on PM_{2.5} component clusters, which were also correlated with geographic region. This study supports a relationship between air pollution and both invasive breast cancer and DCIS risk within certain geographic subgroups and emphasizes the need to consider variability in air pollution measures by geographic region and composition of the mixture, as well as by tumor staging, when assessing associated risks with breast cancer.

Acknowledgments

This research was supported by the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health (Z01-ES044005 and ES103332-01).

References

- Andersen ZJ, Ravnskjaer L, Andersen KK, Loft S, Brandt J, Becker T, et al. 2017a. Long-term exposure to fine particulate matter and breast cancer incidence in the Danish Nurse Cohort study. *Cancer Epidemiol Biomarkers Prev* 26(3):428–430, PMID: 27913396, <https://doi.org/10.1158/1055-9965.EPI-16-0578>.
- Andersen ZJ, Stafoggia M, Weinmayr G, Pedersen M, Galassi C, Jørgensen JT, et al. 2017b. Long-term exposure to ambient air pollution and incidence of postmenopausal breast cancer in 15 European cohorts within the ESCAPE project. *Environ Health Perspect* 125(10):107005, PMID: 29033383, <https://doi.org/10.1289/EHP1742>.
- Anderson KN, Schwab RB, Martinez ME. 2014. Reproductive risk factors and breast cancer subtypes: a review of the literature. *Breast Cancer Res Treat* 144(1):1–10, PMID: 24477977, <https://doi.org/10.1007/s10549-014-2852-7>.
- Barclay J, Ernster V, Kerlikowske K, Grady D, Sickles EA. 1997. Comparison of risk factors for ductal carcinoma in situ and invasive breast cancer. *J Natl Cancer Inst* 89(1):76–82, PMID: 8978410, <https://doi.org/10.1093/jnci/89.1.76>.
- Bell ML, Dominici F, Ebisu K, Zeger SL, Samet JM. 2007. Spatial and temporal variation in PM_{2.5} chemical composition in the United States for health effects studies. *Environ Health Perspect* 115(7):989–995, PMID: 17637911, <https://doi.org/10.1289/ehp.9621>.
- Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Muti P, et al. 2005. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total

- suspended particulates as a proxy measure. *Cancer Epidemiol Biomarkers Prev* 14(1):53–60, PMID: 15668476.
- Brook RD, Rajagopalan S, Pope CA III, Brook JR, Bhatnagar A, Diez-Roux AV, et al. 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation* 121(21):2331–2378, PMID: 20458016, <https://doi.org/10.1161/CIR.0b013e3181d8bec1>.
- Chen F, Bina WF. 2012. Correlation of white female breast cancer incidence trends with nitrogen dioxide emission levels and motor vehicle density patterns. *Breast Cancer Res Treat* 132(1):327–333, PMID: 22076479, <https://doi.org/10.1007/s10549-011-1861-z>.
- Chen ST, Lin CC, Liu YS, Lin C, Hung PT, Jao CW, et al. 2013. Airborne particulate collected from central Taiwan induces DNA strand breaks, poly(ADP-ribose) polymerase-1 activation, and estrogen-disrupting activity in human breast carcinoma cell lines. *J Environ Sci Health A Tox Hazard Subst Environ Eng* 48(2):173–181, PMID: 23043339, <https://doi.org/10.1080/10934529.2012.717809>.
- D'Aloisio AA, Nichols HB, Hodgson ME, Deming-Halverson SL, Sandler DP. 2017. Validity of self-reported breast cancer characteristics in a nationwide cohort of women with a family history of breast cancer. *BMC Cancer* 17(1):692, PMID: 29058598, <https://doi.org/10.1186/s12885-017-3686-6>.
- DuPre NC, Hart JE, Bertrand KA, Kraft P, Laden F, Tamimi RM. 2017. Residential particulate matter and distance to roadways in relation to mammographic density: results from the Nurses' Health Studies. *Breast Cancer Res* 19(1):124, PMID: 29169389, <https://doi.org/10.1186/s13058-017-0915-5>.
- Engel LS, Hill DA, Hoppin JA, Lubin JH, Lynch CF, Pierce J, et al. 2005. Pesticide use and breast cancer risk among farmers' wives in the Agricultural Health Study. *Am J Epidemiol* 161(2):121–135, PMID: 15632262, <https://doi.org/10.1093/aje/kwi022>.
- Franklin M, Koutrakis P, Schwartz J. 2008. The role of particle composition on the association between PM_{2.5} and mortality. *Epidemiology* 19(5):680–689, PMID: 18714438, <https://doi.org/10.1097/EDE.0b013e3181812bb7>.
- Hamra GB, Laden F, Cohen AJ, Raaschou-Nielsen O, Brauer M, Loomis D. 2015. Lung cancer and exposure to nitrogen dioxide and traffic: a systematic review and meta-analysis. *Environ Health Perspect* 123(11):1107–1112, PMID: 25870974, <https://doi.org/10.1289/ehp.1408882>.
- Hart JE, Bertrand KA, DuPre N, James P, Vieira VM, Tamimi RM, et al. 2016. Long-term particulate matter exposures during adulthood and risk of breast cancer incidence in the Nurses' Health Study II prospective cohort. *Cancer Epidemiol Biomarkers Prev* 25(8):1274–1276, PMID: 27257091, <https://doi.org/10.1158/1055-9965.EPI-16-0246>.
- Hill P, Wynder EL. 1979. Nicotine and cotinine in breast fluid. *Cancer Lett* 6(4–5):251–254, PMID: 436120, [https://doi.org/10.1016/s0304-3835\(79\)80042-7](https://doi.org/10.1016/s0304-3835(79)80042-7).
- Hystad P, Villeneuve PJ, Goldberg MS, Crouse DL, Johnson K, Canadian Cancer Registries Epidemiology Research Group. 2015. Exposure to traffic-related air pollution and the risk of developing breast cancer among women in eight Canadian provinces: a case-control study. *Environ Int* 74:240–248, PMID: 25454241, <https://doi.org/10.1016/j.envint.2014.09.004>.
- IARC (International Agency for Research on Cancer). 2010. Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. IARC Monogr Eval Carcinog Risks Hum 92:35–818, PMID: 21141735.
- Keller JP, Drton M, Larson T, Kaufman JD, Sandler DP, Szpiro AA. 2017. Covariate-adaptive clustering of exposures for air pollution epidemiology cohorts. *Ann Appl Stat* 11(1):93–113, PMID: 28572869, <https://doi.org/10.1214/16-AOAS392>.
- Kresovich JK, Xu Z, O'Brien KM, Weinberg CR, Sandler DP, Taylor JA. 2019. Methylation-based biological age and breast cancer risk. *J Natl Cancer Inst* djz020, PMID: 30794318, <https://doi.org/10.1093/jnci/djz020>.
- Lee S, McLaughlin R, Harnly M, Gunier R, Kreutzer R. 2002. Community exposures to airborne agricultural pesticides in California: ranking of inhalation risks. *Environ Health Perspect* 110(12):1175–1184, PMID: 12460795, <https://doi.org/10.1289/ehp.021101175>.
- Lerro CC, Koutros S, Andreotti G, Friesen MC, Alavanja MC, Blair A, et al. 2015. Organophosphate insecticide use and cancer incidence among spouses of pesticide applicators in the Agricultural Health Study. *Occup Environ Med* 72(10):736–744, PMID: 26150671, <https://doi.org/10.1136/oemed-2014-102798>.
- Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al. 2013. The carcinogenicity of outdoor air pollution. *Lancet Oncol* 14(13):1262–1263, PMID: 25035875, [https://doi.org/10.1016/S1470-2045\(13\)70487-X](https://doi.org/10.1016/S1470-2045(13)70487-X).
- Mordukhovich I, Beyea J, Herring AH, Hatch M, Stellman SD, Teitelbaum SL, et al. 2016. Vehicular traffic-related polycyclic aromatic hydrocarbon exposure and breast cancer incidence: the Long Island Breast Cancer Study Project (LIBCSP). *Environ Health Perspect* 124(1):30–38, PMID: 26008800, <https://doi.org/10.1289/ehp.1307736>.
- Nie J, Beyea J, Bonner MR, Han D, Vena JE, Rogerson P, et al. 2007. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer Causes Control* 18(9):947–955, PMID: 17632764, <https://doi.org/10.1007/s10552-007-9036-2>.
- Niehoff N, White AJ, McCullough LE, Steck SE, Beyea J, Mordukhovich I, et al. 2017. Polycyclic aromatic hydrocarbons and postmenopausal breast cancer: an evaluation of effect measure modification by body mass index and weight change. *Environ Res* 152:17–25, PMID: 27741445, <https://doi.org/10.1016/j.envres.2016.09.022>.
- Polyak K. 2011. Heterogeneity in breast cancer. *J Clin Invest* 121(10):3786–3788, PMID: 21965334, <https://doi.org/10.1172/JCI60534>.
- Qin YY, Leung CKM, Leung AOW, Wu SC, Zheng JS, Wong MH. 2010. Persistent organic pollutants and heavy metals in adipose tissues of patients with uterine leiomyomas and the association of these pollutants with seafood diet, BMI, and age. *Environ Sci Pollut Res* 17(1):229–240, <https://doi.org/10.1007/s11356-009-0251-0>.
- Reding KW, Young MT, Szpiro AA, Han CJ, DeRoo LA, Weinberg C, et al. 2015. Breast cancer risk in relation to ambient air pollution exposure at residences in the Sister Study cohort. *Cancer Epidemiol Biomarkers Prev* 24(12):1907–1909, PMID: 26464427, <https://doi.org/10.1158/1055-9965.EPI-15-0787>.
- Reeves GK, Pirie K, Green J, Bull D, Beral V, Million Women Study Collaborators. 2012. Comparison of the effects of genetic and environmental risk factors on *in situ* and invasive ductal breast cancer. *Int J Cancer* 131(4):930–937, PMID: 21952983, <https://doi.org/10.1002/ijc.26460>.
- Sampson PD, Richards M, Szpiro AA, Bergen S, Sheppard L, Larson TV, et al. 2013. A regionalized national universal kriging model using partial least squares regression for estimating annual PM_{2.5} concentrations in epidemiology. *Atmos Environ* (1994) 75:383–392, PMID: 24015108, <https://doi.org/10.1016/j.atmosenv.2013.04.015>.
- Sandler DP, Hodgson ME, Deming-Halverson SL, Juras PS, D'Aloisio AA, Suarez LM, et al. 2017. The Sister Study cohort: baseline methods and participant characteristics. *Environ Health Perspect* 125(12):127003, PMID: 29373861, <https://doi.org/10.1289/EHP1923>.
- Shacklette HT, Boerngen JG. 1984. Element concentrations in soils and other surficial materials of the conterminous United States. U.S. Geological Survey professional paper 1270. Washington, DC: U.S. Government Printing Office. <https://pubs.usgs.gov/pp/1270/> [accessed 23 September 2019].
- Shmuel S, White AJ, Sandler DP. 2017. Residential exposure to vehicular traffic-related air pollution during childhood and breast cancer risk. *Environ Res* 159:257–263, PMID: 28823803, <https://doi.org/10.1016/j.envres.2017.08.015>.
- Siegel RL, Miller KD, Jemal A. 2019. Cancer statistics, 2019. *CA Cancer J Clin* 69(1):7–34, PMID: 30620402, <https://doi.org/10.3322/caac.21551>.
- Villeneuve PJ, Goldberg MS, Crouse DL, To T, Weichenthal SA, Wall C, et al. 2018. Residential exposure to fine particulate matter air pollution and incident breast cancer in a cohort of Canadian women. *Environ Epidemiol* 2(3):e021, <https://doi.org/10.1097/EE9.0000000000000021>.
- Virrig BA, Wang S-Y, Shamilyan T, Kane RL, Tuttle TM. 2010. Ductal carcinoma *in situ*: risk factors and impact of screening. *J Natl Cancer Inst Monogr* 2010(41):113–116, PMID: 20956813, <https://doi.org/10.1093/jncimonographs/lqq024>.
- Wei Y, Davis J, Bina WF. 2012. Ambient air pollution is associated with the increased incidence of breast cancer in US. *Int J Environ Health Res* 22(1):12–21, PMID: 21644128, <https://doi.org/10.1080/09603123.2011.588321>.
- White AJ, Bradshaw PT, Hamra GB. 2018. Air pollution and breast cancer: a review. *Curr Epidemiol Rep* 5(2):92–100, PMID: 30271702, <https://doi.org/10.1007/s40471-018-0143-2>.
- White AJ, Chen J, Teitelbaum SL, McCullough LE, Xu X, Hee Cho Y, et al. 2016. Sources of polycyclic aromatic hydrocarbons are associated with gene-specific promoter methylation in women with breast cancer. *Environ Res* 145:93–100, PMID: 26671626, <https://doi.org/10.1016/j.envres.2015.11.033>.
- White AJ, Kresovich JK, Keller JP, Xu Z, Kaufman JD, Weinberg CR, et al. 2019a. Air pollution, particulate matter composition and methylation-based biologic age. *Environ Int* 132:105071, PMID: 31387022, <https://doi.org/10.1016/j.envint.2019.105071>.
- White AJ, Nichols HB, Bradshaw PT, Sandler DP. 2015. Overall and central adiposity and breast cancer risk in the Sister Study. *Cancer* 121(20):3700–3708, PMID: 26193782, <https://doi.org/10.1002/ncr.29552>.
- White AJ, O'Brien KM, Niehoff NM, Carroll R, Sandler DP. 2019b. Metallic air pollutants and breast cancer risk in a nationwide cohort study. *Epidemiology* 30(1):20–28, PMID: 30198937, <https://doi.org/10.1097/EDE.0000000000000917>.
- White AJ, Weinberg CR, O'Meara ES, Sandler DP, Sprague BL. 2019c. Airborne metals and polycyclic aromatic hydrocarbons in relation to mammographic breast density. *Breast Cancer Res* 21(1):24, PMID: 30760301, <https://doi.org/10.1186/s13058-019-1110-7>.
- Yaghiyan L, Arao R, Brokamp C, O'Meara ES, Sprague BL, Ghita G, et al. 2017. Association between air pollution and mammographic breast density in the Breast Cancer Surveillance Consortium. *Breast Cancer Res* 19(1):36, PMID: 28381271, <https://doi.org/10.1186/s13058-017-0828-3>.
- Young MT, Bechle MJ, Sampson PD, Szpiro AA, Marshall JD, Sheppard L, et al. 2016. Satellite-based NO₂ and model validation in a national prediction model based on universal kriging and land-use regression. *Environ Sci Technol* 50(7):3686–3694, PMID: 26927327, <https://doi.org/10.1021/acs.est.5b05099>.