A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP7246.

Fine Particulate Matter Exposure and Cancer Incidence: Analysis of SEER Cancer Registry Data from 1992–2016

Nathan C. Coleman,[1] Richard T. Burnett,[2] Majid Ezzati,[3] Julian D. Marshall,[4] Allen L. Robinson,[5] and C. Arden Pope III[1]

[1]Department of Economics, Brigham Young University, Provo, Utah, USA
[2]Private Consultant, Ottawa, Ontario, Canada
[3]Medical Research Council–Public Health England Centre for Environment and Health, School of Public Health, Imperial College London, London, UK
[4]Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, USA
[5]Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Background: Previous research has identified an association between fine particulate matter ($PM_{2.5}$) air pollution and lung cancer. Most of the evidence for this association, however, is based on research using lung cancer mortality, not incidence. Research that examines potential associations between $PM_{2.5}$ and incidence of non-lung cancers is limited.

Objectives: The primary purpose of this study was to evaluate the association between the incidence of cancer and exposure to $PM_{2.5}$ using >8.5 million cases of cancer incidences from U.S. registries. Secondary objectives include evaluating the sensitivity of the associations to model selection, spatial control, and latency period as well as estimating the exposure–response relationship for several cancer types.

Methods: Surveillance, Epidemiology, and End Results (SEER) program data were used to calculate incidence rates for various cancer types in 607 U.S. counties. County-level $PM_{2.5}$ concentrations were estimated using integrated empirical geographic regression models. Flexible semi-nonparametric regression models were used to estimate associations between $PM_{2.5}$ and cancer incidence for selected cancers while controlling for important county-level covariates. Primary time-independent models using average incidence rates from 1992–2016 and average $PM_{2.5}$ from 1988–2015 were estimated. In addition, time-varying models using annual incidence rates from 2002–2011 and lagged moving averages of annual estimates for $PM_{2.5}$ were also estimated.

Results: The incidences of all cancer and lung cancer were consistently associated with $PM_{2.5}$. The incident rate ratios (IRRs), per 10-$\mu g/m^3$ increase in $PM_{2.5}$, for all and lung cancer were 1.09 (95% CI: 1.03, 1.14) and 1.19 (95% CI: 1.09, 1.30), respectively. Less robust associations were observed with oral, rectal, liver, skin, breast, and kidney cancers.

Discussion: Exposure to $PM_{2.5}$ air pollution contributes to lung cancer incidence and is potentially associated with non-lung cancer incidence. https://doi.org/10.1289/EHP7246

Introduction

Toxicology research indicates that the carcinogenic compounds contained in fine particulate matter ($PM_{2.5}$; particles <2.5 μm in aerodynamic diameter) contribute to chronic systemic inflammation (Loomis et al. 2013), oxidative stress (Risom et al. 2005), and DNA damage (Newby et al. 2015) in the lungs. Furthermore, extensive epidemiological evidence indicates that $PM_{2.5}$ is associated with lung cancer mortality (Crouse et al. 2015; Yin et al. 2017; Lepeule et al. 2012; Turner et al. 2011; Pope et al. 2019). For example, a recent meta-analysis estimated the hazard ratio (HR) for the association between $PM_{2.5}$ and lung cancer to be 1.14 [95% confidence interval (CI): 1.08, 1.21] (Pope et al. 2020). Much of the epidemiological evidence to support this association, however, is based on prospective cohort studies that examined lung cancer mortality, not lung cancer incidence. Although several recent studies have used incidence data to estimate the association between $PM_{2.5}$ and lung cancer (IARC 2013; Bai et al. 2020; Zhang et al. 2020), further research is needed to confirm the association and examine the sensitivity of the results to modeling choices and exposure windows.

In addition to lung cancer, several cohort studies have found limited evidence of an association between mortality and incidence of various non-lung cancers and air pollution (Coleman et al. 2020; Turner et al. 2017; Wong et al. 2016; Ancona et al. 2015; Raaschou-Nielsen et al. 2011). However, these studies were inconsistent in their findings and often limited by small sample size. Furthermore, the use of mortality follow-up is insufficient to address the effect of air pollution on burden of disease for cancer because of the difficulty in addressing the problem of latency, accurately analyzing cancers that are highly survivable, and the possible confounding from mortality of other causes. Further evidence using cancer incidence data instead of mortality contributes to evaluating whether non-lung cancer sites are associated with exposure to $PM_{2.5}$.

The primary purpose of the present study was to evaluate the association between the incidence of cancer and exposure to $PM_{2.5}$, using available cancer incidence data from U.S. cancer registries. Secondary objectives included evaluating the sensitivity of the associations to various lag structures and exposure windows, exploring the sensitivity of results to modeling assumptions, and evaluating potential nonlinearities in the exposure–response relationship for various types of cancers.

Methods

Cancer Incidence Data

The U.S. National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) program contains all cancer cases across cancer registries that cover approximately 34.6% of the United States (NCI 2019b). The SEER program contains individual-level cancer incidence from 1975–2016 collected from cancer registries located in California, Connecticut, Detroit, Georgia, Iowa, Kentucky, Louisiana, New Jersey, New Mexico, Seattle (Puget Sound), and Utah (NCI 2019b). A detailed description of the location

Address correspondence to C. Arden Pope III, 142 FOB, Brigham Young University, Provo, UT 84602 USA. Telephone: (801) 422-2157. Email: cap3@byu.edu

Supplemental Material is available online (https://doi.org/10.1289/EHP7246).

M.E. received charitable grant support from the AstraZeneca Young Health Programme and personal fees from Prudential, Scor, and Third Bridge, all outside the submitted work. All other authors declare they have no actual or potential competing financial interests.

Received 14 April 2020; Revised 2 September 2020; Accepted 15 September 2020; Published 9 October 2020.

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.

of registries is contained in Table S1. These data are publicly available but require a signed SEER research data use agreement (NCI 2019a).

County-level incidence rates were calculated from the SEER program's cancer case data to estimate the association between $PM_{2.5}$ and cancer incidence. First, cancer cases were totaled for every county-year and grouped by the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10; WHO 2016) codes as follows: oral and oropharyngeal (defined by ICD-10 Codes C00–C14), esophageal (C15), stomach (C16), small intestine (C17), colon (C18), rectal (C19–C21), liver and biliary tract (C22–C24), pancreatic (C25), nose (C30–C31), laryngeal and trachea (C32–C33), lung and bronchus (C34), bone (C40–C41), skin (C43–C44), connective and soft tissue (C45–C49), breast (C50), cervical (C53), uterine (C54–C55), ovarian (C56), prostate (C61), other male (C60, C62–C63), kidney (C64–C65), bladder (C67), brain (C71), endocrine (C73–C75), and ill-defined cancers (C76–C80). Next, yearly cancer incidence rates per 100,000 for each county were calculated by dividing by yearly population data (provided by the SEER program via the U.S. Census) and multiplying by 100,000 for every cancer type (NCI 2019d).

For the primary analysis, the yearly cancer incidence rates were averaged for each county from 1992–2016 to allow for harmonizing several key variables and for use in a time-independent model. Average incidence rates from 2008–2016 were also calculated for use in a latency sensitivity analysis. After removing counties that were missing $PM_{2.5}$ estimates or other covariates, 607 counties remained. In addition to the time-independent analysis, cancer incidence data was also used to generate annual-average incidence rates at the county-year level in the 607 counties contained in the SEER program data for a time-varying model. Due to covariate limitations, only incidence data from 2002–2011 were available.

Air Pollution Exposure

Regulatory monitoring data for $PM_{2.5}$ was collected nationwide starting in 1999. These regulatory data, within an integrated empirical geographic regression modeling framework, were used to generate county-level annual-average $PM_{2.5}$ concentrations for 1999–2015. Hold-out cross-validation (CV) indicated good model performance (10-fold CV-R^2: 0.78, 0.90). More details describing this approach is found elsewhere (Pope et al. 2019; Kim et al. 2020). All annual estimates for $PM_{2.5}$ are available at the Center for Air, Climate, and Energy Solutions' website (https://www.caces.us/).

In order to better account for the lagged effect of $PM_{2.5}$ on cancer incidence, backcasted $PM_{2.5}$ estimates for 1988–1998 were also calculated. The estimated PM_{10} concentration in each county from 1988–1998 was multiplied by the county's mean $PM_{2.5}$ to PM_{10} ratio from 1999–2003 to generate estimates of the $PM_{2.5}$ concentration in each county from 1988–1998 (Pope et al. 2019). The mean $PM_{2.5}$ concentrations for 1999–2015 and 1988–2015 were highly correlated ($R \geq 0.97$). $PM_{2.5}$ average exposure from 1999–2015 and from 1988–2015 were linked to average cancer incidence rates from 1992–2016 by county for use in the primary time-independent model. For the latency sensitivity analysis average $PM_{2.5}$ exposures from 1988–2007 were linked to cancer incidence rates from 2008–2016 to allow for a lag period. Finally, for the time-varying model, 1-, 5-, 10-, and 15-y lagged moving averages of $PM_{2.5}$ were estimated and linked to annual incidence rates in each of the counties.

Additional Covariates

The SEER program provides additional county covariate information collected from the U.S. Census and American Community Survey, including the following: percentage male; percentage white, black, Hispanic, and other race/ethnicity; percentage of the population in each 5-y age group from 0 through 85; educational attainment (percentage to not graduate high school, percentage to graduate high school, and percentage to have some college education); median income (adjusted to 2017 U.S. dollars); median home value and rent; percentage below 150% poverty; percentage unemployed; percentage working class; and percentage of the population of the county living in rural regions of the county [NCI 2019c, 2019d]. The Behavioral Risk Factor Surveillance System and the National Health and Nutrition Examination Survey were used to obtain additional county-level information including percentage smoking (available from 1996–2012) (Dwyer-Lindgren et al. 2014), percentage alcohol consumption (available from 2002–2012) (Dwyer-Lindgren et al. 2015), and percentage physically active and obese (available from 2001–2011) (Dwyer-Lindgren et al. 2013). For the primary time-independent analysis, covariate data were averaged over the available time and linked by county to create a cross-sectional data set. For the latency analysis that used 2008–2016 incidence rate data, only covariate data for years before 2008 were averaged and linked. For the time-varying model, covariate information from 2002–2011 were linked by county-year. In addition, spatial indicator variables for urban vs. rural (classified as urban if more than 50% of a county's population lived in an urbanized area of ≥50,000 people or an urban cluster of between 2,500 and 50,000 people), state, and region (Pacific, West, Midwest, Northeast, or South) were constructed.

Statistical Methods

Flexible semi-nonparametric regression models were used to estimate associations between $PM_{2.5}$ and cancer incidence for selected

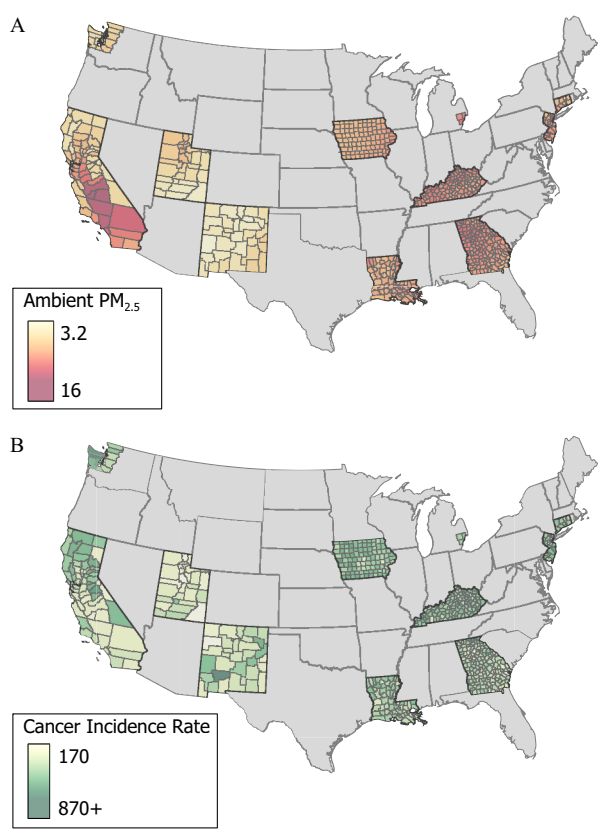

Figure 1. Estimated (A) population-weighted mean (1988–2015) $PM_{2.5}$ concentrations ($\mu g/m^3$) and (B) average incidents rate of all cancer for counties in the SEER database. Note: $PM_{2.5}$, particles <2.5 μm in aerodynamic diameter; SEER, Surveillance, Epidemiology, and End Results (SEER) program.

Table 1. Summary of the total number of cancer cases in counties covered by the SEER program from 1992–2016 and 2008–2016 as well as the mean and standard deviation of incidence rates per 100,000 across counties.

		Cancer cases (n)		Yearly incidence rate (mean ± SD)	
Cancer	ICD-10 code(s)	1992–2016	2008–2016	1992–2016	2008–2016
All cancers	C00–C80	8,658,955	4,130,604	588.83 ± 119.07	636.60 ± 130.65
Digestive tract					
Oral	C00–C14	214,295	105,500	15.34 ± 4.18	17.21 ± 5.49
Esophagus	C15	77,996	36,654	5.74 ± 2.04	6.26 ± 2.90
Stomach	C16	150,349	67,087	8.27 ± 2.52	8.54 ± 3.35
Small intestine	C17	42,103	22,324	2.90 ± 1.10	3.45 ± 1.68
Colon	C18	599,263	249,664	43.68 ± 13.10	41.08 ± 12.59
Rectum	C19–C21	282,683	129,418	19.50 ± 5.26	19.91 ± 6.14
Liver	C22–C24	185,012	102,183	10.05 ± 2.98	12.79 ± 4.68
Pancreas	C25	208,078	106,060	13.78 ± 3.72	15.75 ± 4.82
Respiratory					
Nose	C30–C31	15,186	7,302	0.97 ± 0.55	1.04 ± 0.87
Larynx and trachea	C32–C33	67,281	29,083	5.98 ± 2.53	6.15 ± 3.39
Lung	C34	1,043,065	469,176	85.95 ± 30.06	88.84 ± 33.16
Bone/tissue					
Bone	C40–C41	484,403	256,882	33.13 ± 8.67	39.38 ± 11.65
Skin	C43–C44	680,627	372,095	39.23 ± 16.28	50.56 ± 23.17
Soft tissue	C45–C49	80,524	39,761	4.86 ± 1.42	5.25 ± 2.19
Sex-specific[a]					
Breast	C50	1,473,349	705,738	84.50 ± 17.45	89.80 ± 20.38
Cervix	C53	74,991	31,013	4.70 ± 1.60	4.20 ± 2.10
Uterine	C54–C55	237,965	120,334	7.60 ± 2.30	7.00 ± 3.00
Ovary	C56	126,294	53,269	14.60 ± 4.60	16.50 ± 5.90
Prostate	C61	1,151,454	490,964	74.63 ± 19.08	70.87 ± 19.22
Other male specific	C60, C62–C63	63,570	29,972	3.57 ± 1.29	3.84 ± 2.08
Urinary tract					
Kidney	C64–C65	254,706	136,978	18.51 ± 4.92	22.15 ± 6.59
Bladder	C67	346,681	162,991	23.40 ± 7.51	24.92 ± 8.83
Other					
Brain	C71	127,898	63,270	8.15 ± 2.14	9.09 ± 3.14
Endocrine	C73–C75	253,243	156,581	14.47 ± 4.07	19.67 ± 6.56
Ill defined	C76–C80	417,939	186,305	27.79 ± 6.78	28.08 ± 8.38

Note: ICD-10, *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*; SEER, Surveillance, Epidemiology, and End Results (SEER) program.
[a]Sex-specific cancer incidence rates are calculated using the entire population, not just one sex.

cancers while controlling for important county-level covariates [generalized additive model procedure in SAS (version 9.4; SAS Institute, Inc.)]. In the primary analysis, incident rate ratios (IRRs) and 95% CIs (per 10-$\mu g/m^3$ increase of $PM_{2.5}$) were estimated by regressing the natural logarithm of the average incidence rate for selected cancer types in 607 counties from 1992–2016 on county-level mean $PM_{2.5}$ concentrations from 1988–2015. Specifically, locally weighted smoothing (LOESS) models with three degrees of freedom (df) were used to flexibly control for possible confounders including percentage of the county in various age buckets; percentage male; percentage white, black, Hispanic, and other; percentage who did not graduate high school, graduated high school, or obtained more education than high school; median income, rent, and home value; percentage below 150% poverty; percentage working class; percentage unemployed; percentage living in a rural area; percentage smokers; percentage alcohol consumption; percentage who are physically active; and percentage of individuals in a county who are obese. Indicator variables for urban vs. rural and state were also included in the model.

After estimating the IRRs and nominal p, two approaches were used to adjust p to account for multiple testing. The first approach was the Holm's method—which is a common modification of the Bonferroni approach because it adjusts for multiple comparisons by controlling the family-wise error rate and providing a somewhat more powerful approach to multiple significance testing (Hochberg and Benjamini 1990). The nominal p for all hypotheses tested are ordered from smallest to largest and given a rank, based on their order (the smallest p is given a rank of 1). The Holm-adjusted p are the nominal p multiplied by the total number of tests minus the p rank plus one. The second approach, the false discovery rate (FDR) method, controls for the false discovery rate and is an alternative modification of the Bonferroni approach with more power than the Holm's method (Benjamini and Hochberg 1995). FDR p are obtained by multiplying the nominal p by the total number of texts divided by rank order.

In addition to the primary analysis, time-varying linear regression models that accounted for changes in air pollution and cancer incidence over time were estimated using county-year–level cancer incidence data from 2002–2011. IRRs (per 10-$\mu g/m^3$ increase of $PM_{2.5}$) were estimated by regressing the natural logarithm of the yearly incidence rate on mean $PM_{2.5}$ concentrations for 1-, 5-, 10-, and 15-y lagged moving averages (to explore alternative cancer latency periods). To account for potential correlations within the same counties over time, 95% CIs were based on robust covariance estimators [Taylor series linearization, using SURVEYREG in SAS (version 9.4; SAS Institute, Inc.)]. To flexibly control for general changes in cancer incidence over time, annual indicator variables for each year (2002–2011) were included. Annual values of all other covariables (percentage of the county in various age buckets; percentage male; percentage white, black, Hispanic, and other; percentage who did not graduate high school, graduated high school, or obtained more education than high school; median income, rent, and home value; percentage below 150% poverty; percentage working class; percentage unemployed; percentage living in a rural area; percentage smokers; percentage alcohol consumption; percentage who are physically active; and percentage of individuals in a county who are obese) were also included. In addition, indicator variables for urban/rural and state were included.

To determine whether the results were sensitive to modeling choices, the following additional models using the primary model

Table 2. Summary of baseline county characteristics (mean ± SD for continuous variables and percentages for indicator variables) from 1992–2016, 1992–2007 and 2002–2011.

Variable	1992–2016 counties	1992–2007 counties	2002–2011 counties
$PM_{2.5}$ exposure (y)			
1988–2015	11.50 ± 2.60	—	—
1999–2015	10.00 ± 2.20	—	—
1988–2007	—	12.70 ± 3.10	—
$PM_{2.5}$ moving average (y)			
1	—	—	10.16 ± 2.58
5	—	—	10.77 ± 2.69
10	—	—	11.35 ± 2.87
15	—	—	12.00 ± 3.06
Age buckets [y (%)]			
0	1.29 ± 0.23	1.22 ± 0.23	1.31 ± 0.26
1–4	5.22 ± 0.81	5.06 ± 0.84	5.27 ± 0.90
5–9	6.76 ± 0.91	6.56 ± 0.99	6.66 ± 0.98
10–14	7.13 ± 0.89	6.66 ± 0.92	7.11 ± 0.93
15–19	7.17 ± 1.02	6.74 ± 1.00	7.28 ± 1.08
20–24	6.47 ± 2.19	6.50 ± 2.17	6.47 ± 2.27
25–29	6.06 ± 1.13	6.07 ± 1.21	5.99 ± 1.20
30–34	6.27 ± 0.88	6.06 ± 0.92	6.10 ± 1.00
35–39	6.62 ± 0.73	5.95 ± 0.77	6.49 ± 0.98
40–44	6.96 ± 0.68	6.22 ± 0.79	7.08 ± 0.95
45–49	7.07 ± 0.66	6.75 ± 0.75	7.43 ± 0.77
50–54	6.85 ± 0.75	7.23 ± 0.77	7.10 ± 0.88
55–59	6.21 ± 0.86	6.97 ± 0.93	6.35 ± 1.08
60–64	5.35 ± 0.98	6.25 ± 1.16	5.29 ± 1.20
65–69	4.46 ± 0.98	5.12 ± 1.19	4.21 ± 1.03
70–74	3.56 ± 0.87	3.81 ± 0.96	3.37 ± 0.88
75–79	2.77 ± 0.76	2.81 ± 0.75	2.70 ± 0.78
80–84	1.99 ± 0.64	2.03 ± 0.63	2.00 ± 0.68
85	1.81 ± 0.74	1.98 ± 0.81	1.77 ± 0.80
Race (%)			
White	76.07 ± 20.47	74.25 ± 20.91	75.92 ± 20.47
Black	12.76 ± 16.51	12.99 ± 16.55	12.79 ± 16.55
Hispanic	8.46 ± 13.43	9.71 ± 14.10	8.57 ± 13.49
Other	2.71 ± 5.56	3.04 ± 5.82	2.72 ± 5.60
Sex (%)			
Male	49.68 ± 2.01	49.88 ± 2.20	49.72 ± 2.06
Education (%)			
No high school	23.59 ± 9.24	26.89 ± 10.37	21.03 ± 9.40
Graduate of high school	34.23 ± 6.30	34.03 ± 6.19	34.68 ± 7.03
More than high school	42.17 ± 12.47	39.08 ± 12.80	44.29 ± 12.94
Income			
Median income (2017 adjusted)	37,319 ± 10,992	32,777 ± 9,669	41,196 ± 13,111
Median home value	107,347 ± 73,877	86,868 ± 58,429	124,913 ± 97,211
Median rent	528 ± 171	435 ± 145	585 ± 220
Below 150% poverty (%)	28.21 ± 9.93	27.74 ± 10.33	27.66 ± 9.95
Unemployed (%)	7.36 ± 2.51	6.79 ± 2.60	7.78 ± 3.34
Working class (%)	68.89 ± 5.74	69.82 ± 5.82	68.97 ± 6.48
Health (%)			
Smokers	26.13 ± 4.75	26.80 ± 4.65	25.78 ± 5.13
Consume alcohol	44.98 ± 13.69	44.06 ± 14.29	44.83 ± 13.93
Obese (BMI > 29)	34.90 ± 4.62	33.31 ± 4.49	35.33 ± 5.25
Physically active	71.45 ± 6.51	71.15 ± 6.77	71.60 ± 6.55
Urban vs. rural (%)			
Rural counties	44.06	44.06	44.06
Individuals in rural	54.61 ± 31.99	54.61 ± 31.99	54.59 ± 31.99
Region (%)			
Northeast	4.78	4.78	4.79
Midwest	16.80	16.80	16.67
South	56.50	56.50	56.60
Pacific West	11.70	11.70	11.72
Mountain West	10.22	10.22	10.23
State (%)[a]			
California	9.56	9.56	9.57
Connecticut	1.32	1.32	1.32
Georgia	26.19	26.19	26.24
Iowa	16.31	16.31	16.17
Kentucky	19.77	19.77	19.80
Louisiana	10.54	10.54	10.56
Michigan	0.49	0.49	0.50
New Jersey	3.46	3.46	3.47

Table 2. (*Continued.*)

Variable	1992–2016 counties	1992–2007 counties	2002–2011 counties
New Mexico	5.44	5.44	5.45
Utah	4.78	4.78	4.79
Washington	2.14	2.14	2.15

Note: —, not applicable; BMI, body mass index; $PM_{2.5}$, particles <2.5 μm in aerodynamic diameter; SD, standard deviation; SEER, Surveillance, Epidemiology, and End Results (SEER) program.
[a]SEER registries cover all cancer cases in each state excluding Michigan and Washington, which are limited to cases in the Detroit and Puget Sound area, respectively. See Table S1 for more detail.

(time-independent) were estimated: *a*) a LOESS model that used a cross-validated approach to select the number of degrees of freedom; *b*) a natural smoothing spline model that used a cross-validated approach to select the number of degrees of freedom; *c*) a LOESS model with 3 df, but without state indicator variables; *d*) Model 3, but with regional rather than state indicator variables; *e*) Model 3, but with SEER registry rather than state indicator variables; *f*) a linear regression assuming a Poisson distribution; *g*) a linear regression with estimated standard errors using the sandwich method (White 1980) [using the ROBUST option of the Regress Command in STATA (release 16; StataCorp.)]; *h*) a LOESS model with 3 df that measured $PM_{2.5}$ exposure as the average exposure from 1999–2015 instead of 1988–2015; and *i*) a LOESS model with 3 df that used the average county incident rates from 2008–2016 instead of 1992–2016 and $PM_{2.5}$ exposure from 1988–2007 as well as county averages for all covariates from 1992–2007. Finally, to determine whether the results of the primary model were sensitive to the inclusion of specific covariates, sensitivity analysis was performed by progressively adding control variables into the primary model for selected cancer types.

The shapes of the exposure–response curve for $PM_{2.5}$ and several cancer types were estimated using a LOESS model with 3 df. In addition, the exposure–response curves estimating the association for percentage of a county who identify as smokers and several cancer types were also created using a LOESS model with 3 df. The effect of an increase in the prevalence of smoking in a county on cancer incidence was then compared with the effect of an increase of $PM_{2.5}$ in a county on cancer incidence.

Table 3. Incident rate ratios and 95% confidence interval [IRR (95% CI)] estimates for the association between cancer incidence from 1992–2016 and an increase of 10-$\mu g/m^3$ $PM_{2.5}$ exposure from 1988–2015.

Cancer	LOESS (3 df) [IRR (95% CI)]	Unadjusted *p*-value	Holm's method *p*-value	FDR *p*-value
All cancer	1.09 (1.03, 1.14)	<0.01	0.04	0.02
Digestive tract				
Oral	1.18 (1.03, 1.36)	0.03	0.42	0.09
Esophagus	1.08 (0.88, 1.32)	0.48	1.00	0.69
Stomach	0.96 (0.79, 1.16)	0.68	1.00	0.83
Small intestine	1.13 (0.87, 1.47)	0.35	1.00	0.59
Colon	1.05 (0.96, 1.15)	0.29	1.00	0.54
Rectal	1.15 (1.01, 1.30)	0.03	0.60	0.10
Liver	1.32 (1.11, 1.57)	<0.01	0.04	0.02
Pancreas	0.98 (0.85, 1.12)	0.73	1.00	0.83
Respiratory				
Nose	0.57 (0.35, 0.93)	0.03	0.60	0.10
Larynx	1.19 (0.97, 1.46)	0.09	1.00	0.21
Lung	1.19 (1.09, 1.30)	<0.01	<0.01	<0.01
Bone/tissue				
Bone	1.03 (0.91, 1.16)	0.67	1.00	0.83
Skin	1.22 (1.06, 1.41)	<0.01	0.15	0.04
Soft tissue	1.06 (0.86, 1.29)	0.60	1.00	0.82
Sex-specific				
Breast	1.07 (1.00, 1.16)	0.06	1.00	0.17
Cervix	1.16 (0.93, 1.45)	0.20	1.00	0.43
Uterine	0.99 (0.85, 1.15)	0.87	1.00	0.87
Ovarian	0.98 (0.82, 1.17)	0.81	1.00	0.84
Prostate	0.96 (0.87, 1.06)	0.42	1.00	0.64
Other male	1.12 (0.88, 1.43)	0.36	1.00	0.59
Urinary tract				
Kidney	1.21 (1.06, 1.39)	<0.01	0.13	0.04
Bladder	1.05 (0.93, 1.19)	0.77	1.00	0.83
Other				
Brain	1.10 (0.93, 1.29)	0.27	1.00	0.54
Endocrine	1.19 (0.98, 1.44)	0.07	1.00	0.18
Ill defined	1.04 (0.94, 1.17)	0.77	1.00	0.83

Note: Adjusted for percentage of the county in various age buckets; percentage male; percentage white, black, Hispanic, and other; percentage who did not graduate high school, graduated high school, or obtained more education than high school; median income, rent, and home value; percentage below 150% poverty; percentage working class; percentage unemployed; percentage living in a rural area; percentage smokers; percentage who consume alcohol; percentage who are physically active; and percentage of individuals in a county who are obese using LOESS models with 3 df. A *p* of 1 indicates a value >0.99. df, degrees of freedom; FDR, false discovery rate; LOESS, locally weighted smoothing model.

Results

Figure 1 illustrates the average $PM_{2.5}$ concentration from 1988–2015 and average cancer incidence from 1992–2016 for counties contained in the SEER database. Additional information regarding counties in the SEER registries (Table S1). The average $PM_{2.5}$ concentration across the counties contained in the SEER program was 11.5 $\mu g/m^3$ and the average incidence rate for all cancer was 588.8/100,000. Table 1 contains the total number of cases for each cancer site in the SEER program database for the primary analysis (SEER program counties from 1992–2016) and for a sensitivity analysis (SEER program counties from 2008–2016). In addition, the average yearly incidence rate is provided for both the primary analysis and the sensitivity analysis for all cancer sites. Table 2 contains the mean and standard deviation for county characteristics that were included in the time-independent (SEER program counties from 1992–2016), latency sensitivity analysis (SEER program counties 1992–2007), and for the time-varying model (SEER program counties 2002–2011).

Table 3 contains IRRs and 95% CI estimates for the association between a 10-$\mu g/m^3$ increase of $PM_{2.5}$ from 1988–2015 and selected cancer sites. Statistically significant positive associations were found for oral, rectal, liver, lung, skin, and kidney cancers as well as all cancer in aggregate. A borderline statistically significant effect was also found for breast cancer. However, after multiple comparisons adjustments using the Holm's method, only lung [IRR = 1.19 (95% CI: 1.09, 1.30)], liver [IRR = 1.32 (95% CI: 1.11, 1.57)], and all cancer [IRR = 1.09 (95% CI: 1.03, 1.14)] remained significant at a 0.05 level. Using the less conservative FDR method, significant adverse associations were also observed with skin and kidney cancers.

Figure 2 compares the IRR estimates for the base model with estimates from time-varying models using various lagged moving average estimates (1-, 5-, 10-, and 15-y) of exposure for all cancers that were nominally significant at a 0.05 level in the primary analysis (all, lung, oral, rectal, liver, skin, breast, and kidney cancers). Numeric results for all cancer types are provided in Table S2. The associations for all, lung, oral, rectal, skin, and breast cancers and $PM_{2.5}$ were similar for the primary time-independent model and the time-varying model—especially the time-varying models that

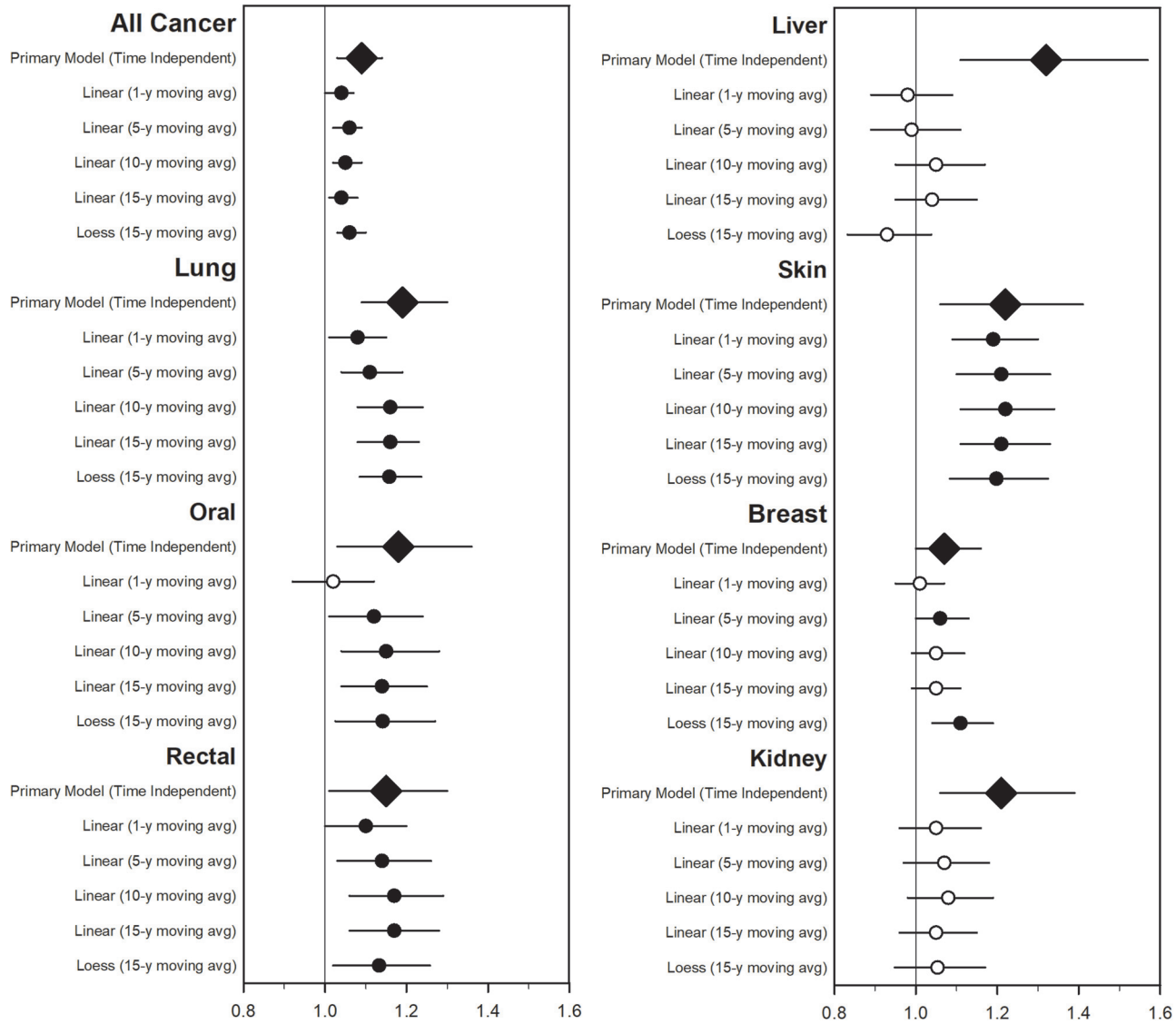

Figure 2. Estimated incident rate ratios (95% CIs) associated with a 10-μg/m^3 increase of $PM_{2.5}$ and selected cancer type incidence from 2002–2011 using time-varying models and compared with the base (time-independent) model. Numerical estimates are included in Table S2. Open circles represent that estimates were not statistically significant at a 0.05 level. Diamonds represent the base (time-independent) model. Models adjusted for percentage of the county in various age buckets; percentage male; percentage white, black, Hispanic, and other; percentage who did not graduate high school, graduated high school, or obtained more education than high school; median income, rent, and home value; percentage below 150% poverty; percentage working class; percentage unemployed; percentage living in a rural area; percentage smokers; percentage alcohol consumption; percentage who are physically active; and percentage of individuals in a county who are obese as well as indicator variables for urban/rural, state, and year. The primary (time-independent) model used a LOESS model with 3 df for all covariates. The linear models used linear yearly estimates for all covariates and 1-, 5-, 10-, and 15-y moving average estimates for $PM_{2.5}$ exposure. The LOESS model was a locally weighted smoothing model with 3 df for all covariates with a 15-y moving average lagged estimate for $PM_{2.5}$ exposure. Note: CI, confidence interval; df, degrees of freedom; $PM_{2.5}$, particles <2.5 μm in aerodynamic diameter; LOESS, locally weighted smoothing model.

used the relatively longer-lagged moving average exposure periods (10 or 15 y). However, for liver and kidney cancers, associations were substantially sensitive to these modeling choices.

Figure 3 contains a forest plot that illustrates the sensitivity analysis performed on those cancer sites that were statistically significant based on the nominal p in the primary model. Numerical results for all cancer sites are provided in Table S3. The results were most statistically robust across modeling choices for lung cancer. All, oral, and skin cancers were largely statistically significant across modeling choices, whereas rectal, liver, breast, and kidney cancers varied substantially across modeling choices. Figure S1 illustrates the sensitivity analysis where covariates were progressively added to the model for the selected cancer types. The estimated IRRs were sensitive to the inclusion of the various levels of covariates. The adverse $PM_{2.5}$–lung cancer association was observed in all models and was most strongly affected by controlling for smoking.

Figure 4 illustrates the lung cancer exposure–response curves for county smoking prevalence and county-level $PM_{2.5}$ concentrations. The relationships between lung cancer and smoking prevalence and the concentration of $PM_{2.5}$ in a county are near linear. County-level smoking prevalence was more strongly associated with lung cancer incidence than $PM_{2.5}$. Figure S2 presents a panel of exposure–response curves for all, oral, rectal, liver, skin, breast, and kidney cancers. Unlike lung cancer, the relationship between various cancer types and $PM_{2.5}$ are not clearly linear, and occasionally $PM_{2.5}$ has a larger effect on cancer incidence than smoking.

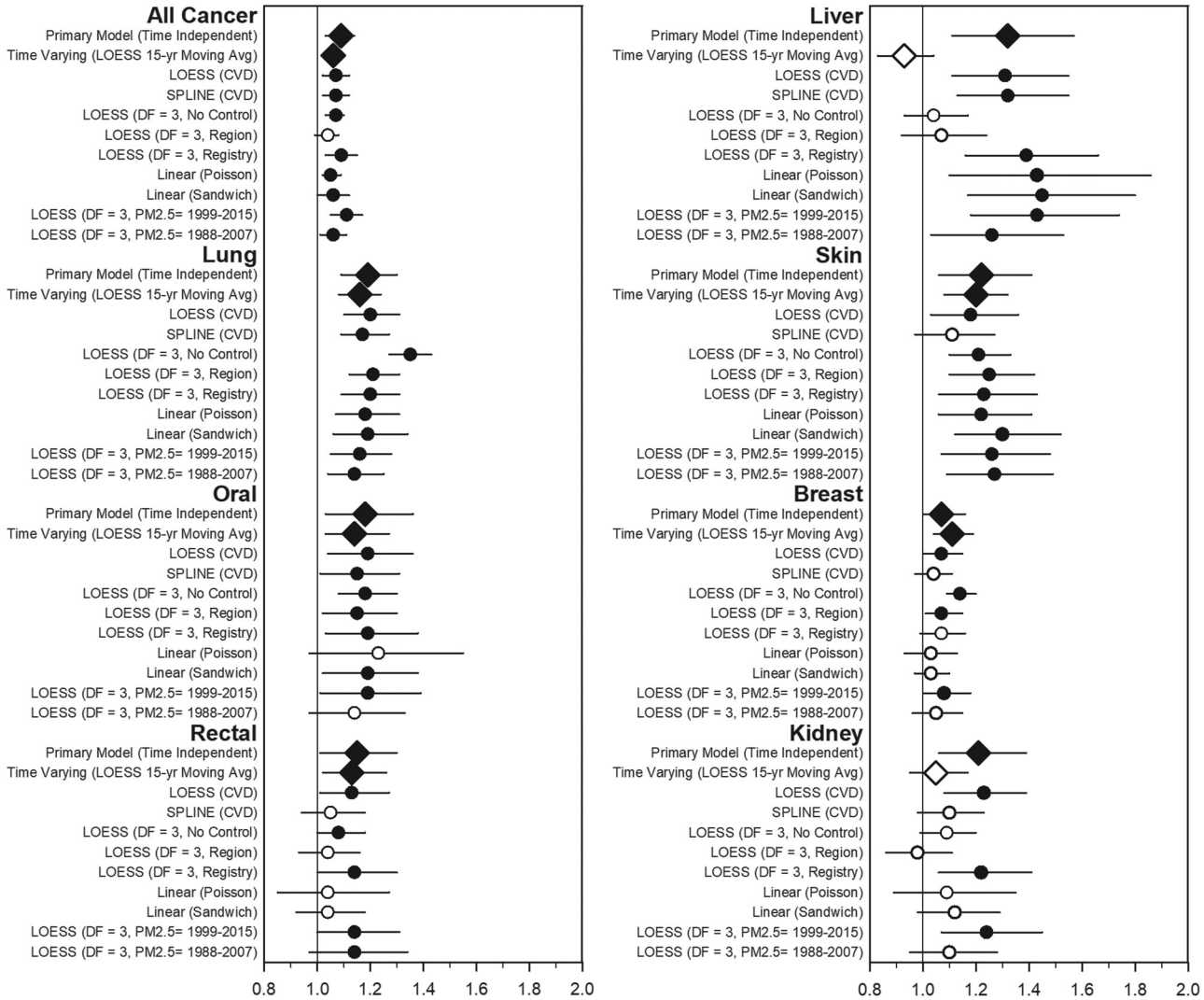

Figure 3. Estimated incident rate ratios and 95% CIs associated with a 10-$\mu g/m^3$ increase of $PM_{2.5}$ from 1988–2015 and average selected cancer type incidence in SEER counties from 1992–2016 across various models. Numerical estimates are included in Table S3. Open circles represent that estimates were not statistically significant at a 0.05 level. Diamonds represent the primary time-independent and time-varying models. Models are adjusted for percentage of the county in various age buckets; percentage male; percentage white, black, Hispanic, and other; percentage who did not graduate high school, graduated high school, or obtained more education than high school; median income, rent, and home value; percentage below 150% poverty; percentage working class; percentage unemployed; percentage living in a rural area; percentage smokers; percentage who consume alcohol; percentage who are physically active; and percentage of individuals in a county who are obese as well as indicator variables for urban/rural and state. All models use the average incidence rate from 1992–2016 (primary time-independent model) unless indicated otherwise. The models include the following: the primary (time-independent) model, a LOESS model with 3 df was used for all covariates; a time-varying mode LOESS model with 3 df for all covariates with an additional indicator variable for year that included a 15-y moving average lagged estimate for $PM_{2.5}$ to estimate exposure for individuals living in SEER counties from 2002–2011; a cross-validated LOESS model for all covariates; a cross-validated spline model for all covariates; a LOESS model with 3 df for all covariates, with the state removed from the model; a LOESS model with 3 df for all covariates, with the state removed from the model and replaced with a region control; a LOESS model with 3 df for all covariates, with the state removed from the model and replaced with a SEER registry control; a linear regression model with only linear terms for the covariates, assuming a Poisson distribution; a linear regression model with only linear terms for the covariates and with the sandwich method used to estimate standard errors; a LOESS model with 3 df for all covariates, with mean $PM_{2.5}$ exposure from 1999–2015; and a LOESS model with 3 df for all covariates, with mean $PM_{2.5}$ exposure from 1988–2007 on SEER counties from 2008–2016. Note: CI, confidence interval; df, degrees of freedom; LOESS, locally weighted smoothing model; $PM_{2.5}$, particles <2.5 μm in aerodynamic diameter; SEER, Surveillance, Epidemiology, and End Results (SEER) program.

Discussion

A growing body of evidence indicates that lung cancer incidence is associated with exposure to $PM_{2.5}$ (IARC 2013; Bai et al. 2020; Zhang et al. 2020). The present study supports this evidence, with a statistically significant IRR of 1.19 (95% CI: 1.09, 1.30), even after conservatively adjusting for multiple comparisons ($p = 0.01$). Furthermore, the lung cancer IRR is remarkably robust across modeling choices, spatial controls, and various exposure windows. Although the present study estimates an IRR that is somewhat higher than the estimate in a recent meta-analysis that examined the association between $PM_{2.5}$ exposure and lung cancer incidence [risk ratio = 1.07 (95% CI: 1.03, 1.12)] (Huang et al. 2017), the IRR from the present study is comparable to the meta-analysis mentioned previously for the association between exposure to $PM_{2.5}$ and lung cancer incidence or mortality [HR = 1.14 (95% CI: 1.08, 1.21)] (Pope et al. 2020). Finally, the exposure–response curve provides evidence that although smoking is a much larger risk factor for lung cancer incidence, $PM_{2.5}$ also contributes to the risk of lung cancer.

The results for non-lung cancers are less conclusive. Although statistically significant associations were found for oral, rectal,

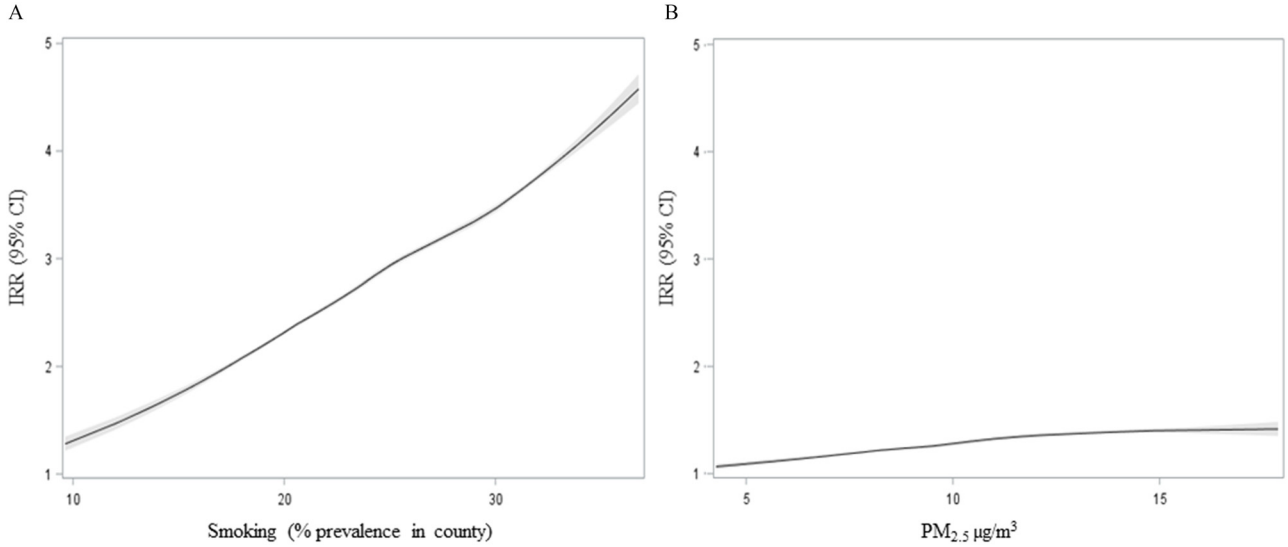

Figure 4. Estimated response relationship between lung cancer incidence and (A) smoking and (B) $PM_{2.5}$. Smoking is estimated as the average percentage of the county's population that identified as smokers from 1996–2012. $PM_{2.5}$ is measured as the population-weighted average concentration in a county from 1988–2015. A locally weighted smoothing (LOESS) model with 3 df to estimate nonlinearity is used. Note: df, degrees of freedom; $PM_{2.5}$, particles <2.5 μm in aerodynamic diameter.

liver, skin, and kidney cancers in the base model, none of these cancer associations were highly robust across sensitivity analysis. Furthermore, no association was found for $PM_{2.5}$ and liver and kidney cancers when time-varying models were used. Previous studies have found statistically significant associations for $PM_{2.5}$ and mortality or incidence from oral and oropharyngeal (Chu et al. 2019), colorectal (Coleman et al. 2020; Turner et al. 2017; Ancona et al. 2015), liver (Coleman et al. 2020; Ancona et al. 2015; Deng et al. 2017; Pan et al. 2016; VoPham et al. 2018), skin (Datzmann et al. 2018) (used PM_{10} instead of $PM_{2.5}$), breast (Coleman et al. 2020; Ancona et al. 2015; Wong et al. 2016; Hu et al. 2013; White et al. 2019; DuPré et al. 2019), and kidney cancers (Turner et al. 2017; Raaschou-Nielsen et al. 2017). Furthermore, the association between all cancer incidence and $PM_{2.5}$ was statistically significant IRR = 1.09 (95% CI: 1.03, 1.14), even after adjusting for multiple comparisons ($p = 0.04$), indicating that the effect of exposure to $PM_{2.5}$ on cancer sites may not be limited to the lungs.

The present study has several strengths. First, the analysis is based on well-documented cancer registry data that contains >8.5 million cases of cancer. Second, this study was able to flexibly control for many relevant county-level risk factors, including smoking, obesity, alcohol consumption, physical activity, income, and education. Third, this study used incidence data instead of mortality data, which avoids the risk of confounding from other causes of death. Finally, the cancer incidence, covariate, and air pollution exposure data are all publicly available.

This study has several limitations. First, this ecological study was unable to control for individual-level risk factors or pollution exposure; therefore, the association between cancer incidence and $PM_{2.5}$ exposure found in this study may not reflect the individual-level association between $PM_{2.5}$ and cancer incidence. However, other studies that have used individual-level data and controlled for a greater variety of risk factors have found comparable associations for cancer mortality and $PM_{2.5}$. Further, this study was unable to control for all potential risk factors of cancer incidence. Several potential confounders include occupational exposures, dietary patterns, diabetes status, or chronic hepatitis B and C virus infection status. Furthermore, the present study found that progressively adding covariates to the model had an impact on the association between $PM_{2.5}$ and cancer incidence, which suggests a possible risk of residual confounding. Finally, the present study does not estimate cancer incidence rates for various age, sex, and race/ethnicity categories. Future studies should examine these associations to determine whether differences in exposure across various substrata, especially race/ethnicity, lead to a substantial difference in $PM_{2.5}$–cancer incidence associations (Zou et al. 2014).

The present study is also limited in its ability to directly measure $PM_{2.5}$ exposures. County-level $PM_{2.5}$ concentrations are generated using population-weighted averages of U.S. Census block-level–modeled estimates that cannot account for the full range of spatial variability. Sensitivity analyses suggest that most cancer associations are not highly sensitive to regional, state, or SEER cancer registry spatial control. It is unclear, however, how the estimates would be affected if the analysis could be conducted at the U.S. Census tract or block level. In addition, the present study had a limited ability to identify the most relevant $PM_{2.5}$ exposure window for cancer incidence. The present study found that the associations between $PM_{2.5}$ and cancer incidence are not sensitive to changes in the exposure windows from 1988–2015, 1999–2015, 1988–2007. Especially for lung cancer, stronger associations were observed for 10- or 15-y lagged moving averages vs. 1- or 5-y lagged moving averages—indicative of a relatively long latency period. This study was unable to generate reliable exposure estimates before 1988. Finally, the primary index of air pollution used in this analysis is $PM_{2.5}$, which does not account for spatial differences in the constituents or characteristics of $PM_{2.5}$ or of various co-pollutants.

The present study supports the growing body of evidence that increased $PM_{2.5}$ exposure is associated with lung cancer incidence. Furthermore, it provides moderate evidence that $PM_{2.5}$ exposure may be associated with the incidence of cancer at other sites, such as oral and oropharyngeal, rectal, liver, skin, breast, and kidney. Although $PM_{2.5}$ is likely not a primary risk factor for cancer incidence, the pervasive nature of air pollution exposure makes further study essential to public health.

Acknowledgments

This publication was developed as part of the Center for Air, Climate, and Energy Solutions, which was supported under assistance agreement no. R835873 awarded by the U.S. Environmental Protection Agency (EPA). It has not been formally reviewed by the U.S. EPA. The views expressed in this

document are solely those of authors and do not necessarily reflect those of the agency. The U.S. EPA does not endorse any products or commercial services mentioned in this publication.

References

Ancona C, Badaloni C, Mataloni F, Bolignano A, Bucci S, Cesaroni G, et al. 2015. Mortality and morbidity in a population exposed to multiple sources of air pollution: a retrospective cohort study using air dispersion models. Environ Res 137:467–474, PMID: 25701728, https://doi.org/10.1016/j.envres.2014.10.036.

Bai L, Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, et al. 2020. Exposure to ambient air pollution and the incidence of lung cancer and breast cancer in the Ontario Population Health and Environment Cohort. Int J Cancer 146(9):2450–2459, PMID: 31304979, https://doi.org/10.1002/ijc.32575.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57(1):289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Chu Y-H, Kao S-W, Tantoh DM, Ko P-C, Lan S-J, Liaw Y-P. 2019. Association between fine particulate matter and oral cancer among Taiwanese men. J Investig Med 67(1):34–38, PMID: 30301867, https://doi.org/10.1136/jim-2016-000263.

Coleman NC, Burnett RT, Higbee JD, Lefler JS, Merrill RM, Ezzati M, et al. 2020. Cancer mortality risk, fine particulate air pollution, and smoking in a large representative cohort of US adults. Cancer Causes Control 31(8):767–776, PMID: 32462559, https://doi.org/10.1007/s10552-020-01317-w.

Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. 2015. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect 123(11):1180–1186, PMID: 26528712, https://doi.org/10.1289/ehp.1409276.

Datzmann T, Markevych I, Trautmann F, Heinrich J, Schmitt J, Tesch F. 2018. Outdoor air pollution, green space, and cancer incidence in Saxony: a semi-individual cohort study. BMC Public Health 18(1):715, PMID: 29884153, https://doi.org/10.1186/s12889-018-5615-2.

Deng H, Eckel SP, Liu L, Lurmann FW, Cockburn MG, Gilliland FD. 2017. Particulate matter air pollution and liver cancer survival. Int J Cancer 141(4):744–749, PMID: 28589567, https://doi.org/10.1002/ijc.30779.

DuPré NC, Hart JE, Holmes MD, Poole EM, James P, Kraft P, et al. 2019. Particulate matter and traffic-related exposures in relation to breast cancer survival. Cancer Epidemiol Biomarkers Prev 28(4):751–759, PMID: 30647065, https://doi.org/10.1158/1055-9965.EPI-18-0803.

Dwyer-Lindgren L, Flaxman AD, Ng M, Hansen GM, Murray CJL, Mokdad AH. 2015. Drinking patterns in US counties from 2002 to 2012. Am J Public Health 105(6):1120–1127, PMID: 25905846, https://doi.org/10.2105/AJPH.2014.302313.

Dwyer-Lindgren L, Freedman G, Engell RE, Fleming TD, Lim SS, Murray CJI, et al. 2013. Prevalence of physical activity and obesity in US counties, 2001–2011: a road map for action. Popul Health Metr 11(1):7, PMID: 23842197, https://doi.org/10.1186/1478-7954-11-7.

Dwyer-Lindgren L, Mokdad AH, Srebotnjak T, Flaxman AD, Hansen GM, Murray CJI. 2014. Cigarette smoking prevalence in US counties: 1996–2012. Popul Health Metr 12(1):5, PMID: 24661401, https://doi.org/10.1186/1478-7954-12-5.

Hochberg Y, Benjamini Y. 1990. More powerful procedures for multiple significance testing. Stat Med 9(7):811–818, PMID: 2218183, https://doi.org/10.1002/sim.4780090710.

Hu H, Dailey AB, Kan H, Xu X. 2013. The effect of atmospheric particulate matter on survival of breast cancer among US females. Breast Cancer Res Treat 139(1):217–226, PMID: 23592372, https://doi.org/10.1007/s10549-013-2527-9.

Huang F, Pan B, Wu J, Chen E, Chen L. 2017. Relationship between exposure to PM2.5 and lung cancer incidence and mortality: a meta-analysis. Oncotarget 8(26):43322–43331, PMID: 28487493, https://doi.org/10.18632/oncotarget.17313.

IARC (International Agency for Research on Cancer). 2013. Outdoor air pollution. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Vol 109. https://monographs.iarc.fr/ENG/Monographs/vol109/mono109-F01.pdf [accessed 19 March 2019].

Kim S-Y, Bechle M, Hankey S, Sheppard L, Szpiro AA, Marshall JD. 2020. Concentrations of criteria pollutants in the contiguous U.S., 1979–2015: role of prediction model parsimony in integrated empirical geographic regression. PLoS One 15(2):e0228535, PMID: 32069301, https://doi.org/10.1371/journal.pone.0228535.

Lepeule J, Laden F, Dockery D, Schwartz J. 2012. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. Environ Health Perspect 120(7):965–970, PMID: 22456598, https://doi.org/10.1289/ehp.1104660.

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.

NCI (National Cancer Institute). 2019a. Surveillance, Epidemiology, and End Results program. Sample SEER research data use agreement. NCI, DCCPS, Surveillance Research Program, released December 2018. https://seer.cancer.gov/data/sample-dua.html [accessed 1 October 2019].

NCI. 2019b. Surveillance, Epidemiology, and End Results program. SEER incidence data, 1975–2017. NCI, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission. https://seer.cancer.gov/data/ [accessed 1 October 2019].

NCI. 2019c. Surveillance, Epidemiology, and End Results program. Time-dependent county attributes. https://seer.cancer.gov/seerstat/variables/countyattribs/time-dependent.html [accessed 1 October 2019].

NCI. 2019d. Surveillance, Epidemiology, and End Results program. U.S. population data—1969–2018. www.seer.cancer.gov/popdata [accessed 1 October 2019].

Newby DE, Mannucci PM, Tell GS, Baccarelli AA, Brook RD, Donaldson K, et al. 2015. Expert position paper on air pollution and cardiovascular disease. Eur Heart J 36(2):83–93, PMID: 25492627, https://doi.org/10.1093/eurheartj/ehu458.

Pan W-C, Wu C-D, Chen M-J, Huang Y-T, Chen C-J, Su H-J, et al. 2016. Fine particle pollution, alanine transaminase, and liver cancer: a Taiwanese prospective cohort study (REVEAL-HBV). J Natl Cancer Inst 108(3):djv341, PMID: 26561636, https://doi.org/10.1093/jnci/djv341.

Pope CA III, Coleman N, Pond ZA, Burnett RT. 2020. Fine particulate air pollution and human mortality: 25+ years of cohort studies. Environ Res 183:108924, PMID: 31831155, https://doi.org/10.1016/j.envres.2019.108924.

Pope CA III, Lefler JS, Ezzati M, Higbee JD, Marshall JD, Kim S-Y, et al. 2019. Mortality risk and fine particulate air pollution in a large, representative cohort of U.S. adults. Environ Health Perspect 127(7):77007, PMID: 31339350, https://doi.org/10.1289/EHP4438.

Raaschou-Nielsen O, Andersen ZJ, Hvidberg M, Jensen SS, Ketzel M, Sørensen M, et al. 2011. Air pollution from traffic and cancer incidence: a Danish cohort study. Environ Health 10(1):67, PMID: 21771295, https://doi.org/10.1186/1476-069X-10-67.

Raaschou-Nielsen O, Pedersen M, Stafoggia M, Weinmayr G, Andersen ZJ, Galassi C, et al. 2017. Outdoor air pollution and risk for kidney parenchyma cancer in 14 European cohorts. Int J Cancer 140(7):1528–1537, PMID: 28006861, https://doi.org/10.1002/ijc.30587.

Risom L, Møller P, Loft S. 2005. Oxidative stress-induced DNA damage by particulate air pollution. Mutat Res 592(1–2):119–137, PMID: 16085126, https://doi.org/10.1016/j.mrfmmm.2005.06.012.

Turner MC, Krewski D, Diver WR, Pope CA III, Burnett RT, Jerrett M, et al. 2017. Ambient air pollution and cancer mortality in the Cancer Prevention Study II. Environ Health Perspect 125(8):087013, PMID: 28886601, https://doi.org/10.1289/EHP1249.

Turner MC, Krewski D, Pope CA III, Chen Y, Gapstur SM, Thun MJ. 2011. Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. Am J Respir Crit Care Med 184(12):1374–1381, PMID: 21980033, https://doi.org/10.1164/rccm.201106-1011OC.

VoPham T, Bertrand KA, Tamimi RM, Laden F, Hart JE. 2018. Ambient $PM_{2.5}$ air pollution exposure and hepatocellular carcinoma incidence in the United States. Cancer Causes Control 29(6):563–572, PMID: 29696510, https://doi.org/10.1007/s10552-018-1036-x.

White AJ, Keller JP, Zhao S, Carroll R, Kaufman JD, Sandler DP. 2019. Air pollution, clustering of particulate matter components, and breast cancer in the Sister Study: a U.S.-wide cohort. Environ Health Perspect 127(10):107002, PMID: 31596602, https://doi.org/10.1289/EHP5131.

White H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 817(4):817–838, https://doi.org/10.2307/1912934.

WHO (World Health Organization). 2016. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*. http://apps.who.int/classifications/icd10/browse/2016/en [accessed 28 April 2020].

Wong CM, Tsang H, Lai HK, Thomas GN, Lam KB, Chan KP, et al. 2016. Cancer mortality risks from long-term exposure to ambient fine particle. Cancer Epidemiol Biomarkers Prev 25(5):839–845, PMID: 27197138, https://doi.org/10.1158/1055-9965.EPI-15-0626.

Yin P, Brauer M, Cohen A, Burnett RT, Liu J, Liu Y, et al. 2017. Long-term fine particulate matter exposure and nonaccidental and cause-specific mortality in a large national cohort of Chinese men. Environ Health Perspect 25(11):117002, PMID: 29116930, https://doi.org/10.1289/EHP1673.

Zhang Z, Zhu D, Cui B, Ding R, Shi X, He P. 2020. Association between particulate matter air pollution and lung cancer. Thorax 75(1):85–87, PMID: 31727788, https://doi.org/10.1136/thoraxjnl-2019-213722.

Zou B, Peng F, Wan N, Mamady K, Wilson GJ. 2014. Spatial cluster detection of air pollution exposure inequities across the United States. PLoS One 9(3):e91917, PMID: 24647354, https://doi.org/10.1371/journal.pone.0091917.