2020-02-26

Cancer Mortality Risk, Fine Particulate Air Pollution, and Smoking in a Large, Representative Cohort of US Adults

Nathan Coleman

Honors Thesis

CANCER MORTALITY RISK, FINE PARTICULATE AIR POLLUTION, AND SMOKING IN A LARGE, REPRESENTATIVE COHORT OF US ADULTS

by
Nathan C. Coleman

Submitted to Brigham Young University in partial fulfillment of graduation requirements for University Honors

Economics Department
Brigham Young University
April 2020

Advisor: C. Arden Pope III

Honors Coordinator: John Stovall

Abstract

CANCER MORTALITY RISK, FINE PARTICULATE AIR POLLUTION, AND SMOKING IN A LARGE, REPRESENTATIVE COHORT OF US ADULTS

Purpose: Studies have indicated that air pollution and smoking are associated with various types of mortality, including cancer. The current study utilizes a nationally representative cohort to explore relationships between fine particulate matter ($PM_{2.5}$) exposure, smoking, and cancer mortality.

Methods: National Health Interview Survey and mortality follow-up data were combined to create a study population of 635,539 individuals surveyed from 1987 to 2014. A sub-cohort of 341,665 never-smokers from the full cohort was also evaluated. Individuals were assigned modeled $PM_{2.5}$ exposure. Cox proportional hazard models were utilized to estimate hazard ratios for cancer-specific mortality controlling for age, sex, race, and other important characteristics.

Results: The risk of all cancer mortality was positively associated with $PM_{2.5}$ (per 10 μg/m3 increase) in the full cohort (hazard ratio [HR] 1.15, 95% confidence interval [CI] 1.08–1.22) and the never-smokers' cohort (HR 1.19, 95% CI 1.06–1.33). $PM_{2.5}$ - morality associations were also observed for stomach, colorectal, liver, breast, cervix, and

Nathan C. Coleman

Economics Department

Bachelor of Arts

bladder, as well as Hodgkin lymphoma, non-Hodgkin lymphoma, and leukemia. After adjusting for multiple comparisons, however, only the $PM_{2.5}$ morality association with lung cancer in the non-smoking cohort was statistically significant. Cigarette smoking was statistically associated with mortality from lung, oral and oropharyngeal, esophageal, colorectal, liver, bladder, laryngeal, leukemia and unspecified cancers, even with adjustment for multiple testing.

Conclusions: Exposure to $PM_{2.5}$ air pollution likely contributes to lung cancer mortality and may be a risk factor for other cancer sites. Cigarette smoking has a much larger and is associated with similar cancer-sites.

ACKNOWLEDGMENTS

I would like to thank Richard T. Burnett, Joshua D. Higbee, Jacob S. Lefler, Ray M. Merrill, Majid Ezzati, Julian D. Marshall, Sun-Young Kim, Matthew Bechle, Allen L. Robinson, and C. Arden Pope III for their vital contributions and assistance as coauthors of this work. Furthermore, this publication was developed as part of the Center for Air, Climate, and Energy Solutions (CACES), which was supported under Assistance Agreement No. R835873 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by EPA. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.

I would also like to thank my wonderful family, specifically Brinlee and Oliver, for their continued support and feedback.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

Empirical evidence indicates that exposure to fine particulate matter air pollution ($PM_{2.5}$, particles < 2.5 μm in aerodynamic diameter) contributes to burden of disease for all-cause mortality [1], cardiopulmonary disease [2], and cancer [3]. Previous studies have focused on the association between lung cancer and $PM_{2.5}$ [4-6]; however, evidence indicates that exposure to air pollution could contribute to chronic systemic inflammation [7], oxidative stress [8], and DNA damage [9] in tissues other than lung. For example, recent studies have found associations with $PM_{2.5}$ and non-lung cancer-sites including bladder, colorectal, and kidney [10], female organ cancers and breast [11], liver [12], and stomach [13]. Unfortunately, these studies are limited in scope and number, and not fully consistent in their findings.

The objective of the current study was to explore $PM_{2.5}$-mortality associations with cancer-site specific mortality among a large, nationally representative cohort of adults residing in the United States. Additionally, this study compared the association of cigarette smoking status and cancer-site specific mortality to $PM_{2.5}$ and cancer-site specific mortality. Cigarette smoke contributes to the development of lung, oral and oropharyngeal, esophageal, stomach, colorectal, liver, pancreatic, laryngeal, cervical, bladder, kidney cancers, and leukemia through inhalation of a complex mix of carcinogenic particles [14]. This analysis explored if the $PM_{2.5}$-mortality associations and the smoking mortality associations are observed for similar cancer sites.

METHODS

Study Subjects

This analysis used public National Health Interview Survey (NHIS) and National Death Index data to construct a cohort of individuals aged 18-84 at the time of survey living in the continental U.S, who completed the NHIS survey between 1987 and 2014 as documented elsewhere [2]. Participants represented the civilian noninstitutionalized US adult population. Participant responses were linked to the National Death Index for mortality follow-up through 2015. In addition, restricted-use geographic data allowed for the assignment of ambient pollution estimates at the census tract level.

Analyses were performed on two cohorts. The first cohort consisted of the 635,539 individuals (age range 18-84 yrs, mean age 45.3) and the second was a subset of this group of 341,665 participants who self-reported as never-smokers (age range 18-84, mean age 43.4). Both cohorts contained information on age, sex, race-ethnicity (Non-Hispanic white, Hispanic, Non-Hispanic black, or other), income buckets ($0-35,000, $35,000-50,000, $50,000-75,000, or over $75,000), marital status (married, divorced, separated, never married, or widowed), educational attainment (less than high school grad, high school grad, some college, college grad, more than college grad), BMI, smoking status (self-identified as current, former, or never smoker), census tract, ambient pollution exposure, interview date, mortality status, and date of death.

Further information about the composition of the cohorts, including details regarding the merging and harmonization of key variables, is provided elsewhere [2]. Procedures for informed consent and data collection and linkage of the NHIS files were approved by the NCHS Ethics Review Board. Findings and conclusions of this research

are those of the authors and do not necessarily represent the views of the RDC, the NCHS, the Environmental Protection Agency, or the Centers for Disease Control and Prevention.

Pollution Concentration

Individuals were assigned air pollution exposure estimates based on their resident census tract at the time of the survey, using year-2000 Census tracts for individuals surveyed from 1987 through 2010 and year-2010 Census tracts for those surveyed from 2011 through 2014. In the baseline analysis, each study subject was assigned the average estimated $PM_{2.5}$ concentration from 1999 through 2015. To obtain a longer exposure window from 1988-1999, mean $PM_{2.5}$ / PM_{10} ratios for 1999-2003 were computed and multiplied by the PM_{10} estimate for each census tract from 1988-1998 [2]. Documentation of air pollution estimates utilized in this study is located elsewhere [15]. The modeled air pollution data are publicly accessible at the Center for Air, Climate, & Energy Solutions website (https://www.caces.us/).

Statistical Methods

Hazards ratios and 95% confidence intervals for cancer mortality risk associated with a 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentrations were estimated using Cox proportional hazards models that accounted for the complex, stratified, multistage NHIS sample design [16]. Estimates were computed using the SURVEYPHREG procedure in SAS version 9.3 (SAS Institute, Cary, North Carolina). Analysis was performed on cancer-site specific cases ICD-10 codes for lung (C33-C34), oral and oropharyngeal (C00-C14), esophageal (C15), stomach (C16), colorectal (C18-C21), liver (C22), pancreatic (C25), laryngeal (C32), melanoma (C43), breast (C50), cervical (C53), ovarian (C54-C55),

uterine (C56), prostate (C61), kidney (C64-C65), bladder (C67), and brain cancer (C70-C72) as well as Hodgkin lymphoma (C81), non-Hodgkin lymphoma (NHL) (C82-C85), leukemia (C91-C95), multiple myeloma (C88, C90), and other unspecified cancers (C17, C23-24, C26-C49, C51-52, C57-60, C62-63, C66, C68-C69, C73-C80, C97). All models were adjusted for age-sex-race interactions (using indicators for 5-year age buckets) and categorical variables for BMI, income, education, marital status, rural versus urban, region, and survey year. In the full cohort, models were also adjusted for smoking status. Hazards ratios and 95% confidence intervals for cancer mortality risk associated with smoking status were also estimated. To account for multiple testing, adaptive Holms adjusted p-values [17] were calculated.

Model sensitivity analysis was performed by estimating six additional models: 1) A model using traditional basic Cox Proportional Hazards model (using the Proc PHREG procedure in SAS version 9.3) and controlling for all combinations of 1-yr age groups, sex and race-ethnicity by allowing them to have their own baseline hazard (by including them in the STRATA statement). 2) Model 1 but with indicator variables for education, income, marital status, BMI, and smoking status also added as covariates in the model. 3) Model 2 with indicator variables for urban/rural, census region, and survey year also added as covariates. 4) Model 3 with mean $PM_{2.5}$ data back casted to 1988 (i.e. exposure window of 1988-2015 rather than 1999-2015). 5) Model 3 using only survey years from 1999-2014. 6) Model 3 using the expanded cohort (all 1,599,329 NHIS participants from 1986-2014, including those without smoking or BMI data) and not controlling for smoking status or BMI.

RESULTS

Figure 1 illustrates the average PM2.5 concentrations from 1999-2015 across census tracts in the United States. Table 1 presents summary statistics for both the full and never-smokers' cohort groups. Individual mean estimated ambient $PM_{2.5}$ exposure was 10.7 $\mu g/m^3$ (standard deviation 2.4) in both the full cohort and never-smokers' cohort. The table also contains the average estimated $PM_{2.5}$ exposure for the levels of the selected variables. Individual mean exposure is relatively consistent across varying factor levels aside from race/ethnicity (greater in non-Hispanic Blacks), urban versus rural (greater in urban areas), and census region (greater in the Midwest).

Figure 1. Average $PM_{2.5}$ concentrations from 1999-2015 for census tracts in the US

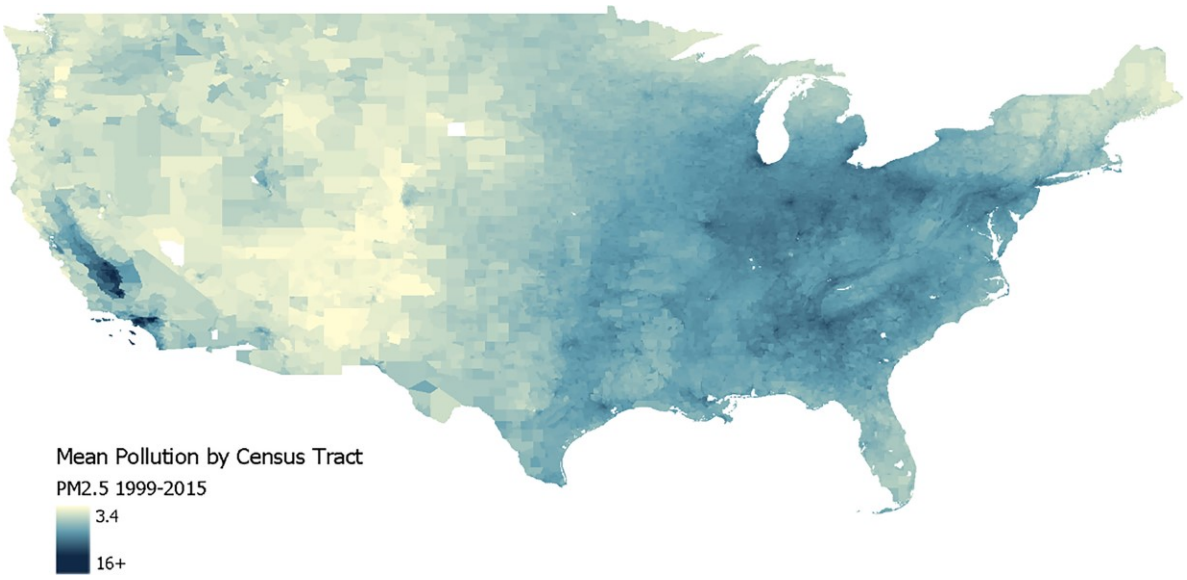


Table 1. Summary of baseline characteristics in the full and never-smoker's cohort.

Variable	Full Cohort (No. = 635,539)		Never-Smokers' Cohort (No. = 341,665)	
	%	Mean (SD) $PM_{2.5}$	%	Mean (SD) $PM_{2.5}$
Sex				
Male	44.5	10.7 (2.4)	38.6	10.7 (2.4)
Female	55.5	10.6 (2.4)	61.4	10.8 (2.4)
Race/Ethnicity				
Non-Hispanic White	67.5	10.3 (2.2)	61.4	10.3 (2.2)
Hispanic	14.1	11.2 (3.0)	17.6	11.3 (3.0)
Non-Hispanic Black	14.0	11.7 (1.9)	15.4	11.7 (1.9)
All other/unknown	4.4	11.0 (2.6)	5.6	11.1 (2.6)
Income (inflation adjusted to 2015)				
$ 0-35,000	38.0	10.8 (2.4)	36.6	10.9 (2.4)
$ 35-50,000	15.5	10.6 (2.4)	14.9	10.7 (2.4)
$ 50-75,000	18.8	10.6 (2.4)	18.7	10.7 (2.4)
$ 75,000+	27.7	10.5 (2.3)	29.9	10.6 (2.3)
Marital Status				
Married	49.6	10.5 (2.4)	49.9	10.6 (2.4)
Divorced	14.1	10.6 (2.4)	10.9	10.7 (2.4)
Separated	3.6	11.1 (2.4)	3.1	11.2 (2.4)
Never Married	24.3	11.0 (2.3)	27.8	11.0 (2.4)
Widowed	8.5	10.7 (2.3)	8.3	10.8 (2.3)
Education				
< High School grad	18.6	11.1 (2.5)	16.8	11.2 (2.5)
High School grad	30.4	10.6 (2.3)	27.1	10.7 (2.4)
Some College	27.1	10.5 (2.4)	27.2	10.6 (2.3)
College grad	15.0	10.6 (2.3)	18.1	10.6 (2.3)
>College grad	8.9	10.6 (2.3)	10.9	10.6 (2.3)
Urban/Rural				
Urban	77.6	11.0 (2.4)	79.4	11.0 (2.4)
Rural	22.4	9.6 (2.1)	20.6	9.6 (2.1)
Census Region				
Northeast	18.1	10.6 (1.9)	17.5	10.8 (1.9)
Midwest	23.7	11.1 (1.7)	22.5	11.1 (1.9)
South	35.7	10.8 (1.7)	36.3	10.8 (1.7)
West	22.5	10.0 (3.6)	23.8	10.3 (3.7)
BMI				
<20	7.3	10.7 (2.3)	7.3	10.7 (2.3)
20-25	36.4	10.6 (2.4)	36.7	10.7 (2.4)
25-30	33.8	10.7 (2.4)	33.1	10.7 (2.4)
30-35	14.4	10.7 (2.4)	14..5	10.8 (2.4)
>35	8.1	10.8 (2.3)	8.4	10.8 (2.3)
Smoking				
Never	53.8	10.7 (2.4)	100	10.7 (2.4)
Current	23.9	10.7 (2.3)	0	-
Former	22.4	10.5 (2.4)	0	-

Table 2 provides cancer-specific mortality hazard ratios (HRs) and 95% confidence intervals (CIs) associated with 10 μg/m^3 increased $PM_{2.5}$ exposure in both the full and never-smokers' cohorts. Without adjustments for multiple testing, statistically significant associations were observed in the full cohort for lung, stomach, colorectal, breast, cervical, and bladder cancer, as well as Hodgkin lymphoma, NHL, and leukemia. However, after adjusting for multiple comparisons, these associations were not statistically significant.

Table 1. Estimated hazard ratios (95% CIs) associated with 10 μg/m3 increase of PM2.5 adjusted for age, sex, race/ethnicity, income, education, marital status, BMI, smoking (for the full cohort), urban versus rural, census regions, and survey year. P-values adjusted using the Holm's method are also included for individual cancer types

Cancer Types	Full Cohort			Never-Smokers' Cohort		
	No. of Deaths	Hazard Ratio (95% CI)	Holm's p-value	No. of Deaths	Hazard Ratio (95% CI)	Holm's p-value
All Cancer	26,453	1.15 (1.08 – 1.22) *	-	17,743	1.19 (1.06 – 1.33) *	-
Lung	7,420	1.13 (1.00 – 1.26) *	0.58	6,710	1.73 (1.20 - 2.49) *†	0.04
Non-Lung	19,033	1.15 (1.07 – 1.24) *	-	11,033	1.15 (1.02 – 1.30) *	-
Digestive and Accessory						
Oral and oropharyngeal	374	1.19 (0.74 – 1.91)	1	291	1.90 (0.65 – 5.54)	1
Esophageal	599	0.59 (0.38 – 0.90)	0.19	460	0.79 (0.32 – 1.96)	1
Stomach	525	1.87 (1.20 – 2.91) *	0.07	301	2.01 (1.01 – 3.98) *	0.51
Colorectal	2,572	1.29 (1.05 – 1.58) *	0.18	1,441	1.26 (0.93 – 1.70)	1
Liver	761	1.32 (0.94 – 1.85)	1	489	2.18 (1.25 – 3.81) *	0.06
Pancreas	1,607	1.09 (0.83 – 1.44)	1	956	0.94 (0.63 – 1.38)	1
Sex Specific Organs						
Breast	2,099	1.33 (1.08 – 1.64) *	0.09	949	1.32 (1.00 – 1.75) *	0.60
Cervical	237	1.77 (1.00 – 3.16) *	0.62	115	2.41 (1.19 – 4.89) *	0.17
Ovarian	392	1.03 (0.69 – 1.53)	1	121	1.06 (0.60 – 1.86)	1
Uterine	750	1.20 (0.73 – 1.96)	1	317	1.64 (0.94 – 2.88)	0.91
Prostate	1,215	0.91 (0.68 – 1.22)	1	802	0.60 (0.39 – 0.93)	0.26
Urinary						
Kidney	603	0.98 (0.66 – 1.46)	1	359	0.94 (0.48 – 1.84)	1
Bladder	589	1.48 (1.00 – 2.29) *	0.63	451	2.00 (0.83 – 4.84)	1
Lymphoid						
Hodgkin lymphoma	59	4.18 (1.20 – 14.60) *	0.30	31	6.21 (1.15 – 33.46) *	0.37
NHL	1,016	1.48 (1.10 – 1.98) *	0.11	558	1.27 (0.81 – 2.01)	1
Leukemia	970	1.43 (1.05 – 1.97) *	0.31	564	1.34 (0.76 – 2.33)	1
Multiple Myeloma	541	0.99 (0.64 – 1.53)	1	270	0.83 (0.45 – 1.54)	1
Other Cancers						
Laryngeal	157	0.82 (0.34 – 1.96)	1	142	0.74 (0.02 – 25.03)	1
Melanoma	392	0.72 (0.39 – 1.33)	1	213	0.54 (0.19 – 1.58)	1
Brain	622	1.48 (0.96 – 2.29)	0.89	344	1.51 (0.84 – 2.70)	1
Unspecified Cancers	2,952	0.89 (0.74 – 1.07)	1	1,858	0.80 (0.60 – 1.07)	1

Note that a p-value of 1 indicates a value greater than 0.9999 as reported by SAS PROC MULTTEST.
* Significant at 95% confidence level using the unadjusted p-values
† Significant at 95% confidence level using Holm's adjusted p-values

In the never-smokers' cohort, statistically significant associations between $PM_{2.5}$ and mortality were found for Hodgkin lymphoma and lung, stomach, liver, breast, and cervical cancers. Only lung cancer was statistically significant after adjusting for multiple comparisons. Table S1 shows sensitivity analysis performed on the full cohort for lung, stomach, colorectal, liver, cervical, breast, and bladder cancers as well as Hodgkin's Lymphoma, Non-Hodgkin's Lymphoma, and Leukemia. The $PM_{2.5}$-moratality associations across the different cancer sites were reasonably insensitive to various modeling choices, different exposure windows, and even using the expanded NHIS cohort; this was especially true for lung, colorectal, liver, and breast cancers.

Table 3 provides HRs and 95% CIs associated with identifying as a current smoker or former smoker and cancer-site specific mortality in the full cohort. Statistically significant smoking-cancer mortality HRs for current smokers were found for lung, oral and oropharyngeal, esophageal, stomach, colorectal, liver, pancreatic, cervical, prostate, kidney, bladder, laryngeal, brain, and unspecified cancers as well as leukemia. For former smokers, statistically significant associations were found for lung, oral and oropharyngeal, esophageal, colorectal, liver, breast, bladder, laryngeal, and unspecified cancers as well as NHL and leukemia. Multiple comparison adjusted p-values were also calculated, which resulted in statistically significant associations for lung, oral and oropharyngeal, esophageal, stomach, colorectal, liver, pancreatic, cervical, bladder, laryngeal, and unspecified cancers in current smokers and lung, oral and oropharyngeal, esophageal, colorectal, liver, pancreatic, bladder, laryngeal, and unspecified cancers in former smokers.

Table 3. Estimated hazard ratios (95% CIs) associated with current or former smoker in comparison to never-smoker. All models were adjusted for age, sex, race, income, education, marital status, BMI, urban versus rural, census regions, survey year, and a 10 $\mu g/m^3$ increase of $PM_{2.5}$. P-values adjusted using the Holm's method are also included for individual cancer types.

Cancer Types	Current Smoker			Former Smoker		
	No. of Deaths	Hazard Ratio (95% CI)	Holm's p-value	No. of Deaths	Hazard Ratio (95% CI)	Holm's p-value
All Cancer	26,453	2.73 (2.64 – 2.83) *	-	17,743	1.48 (1.43 – 1.53) *	-
Lung	7,420	15.11 (13.70 – 16.66) *†	<0.01	6,710	4.90 (4.44 – 5.42) *†	<0.01
Non-Lung	19,033	1.59 (1.53 – 1.66) *	-	11,033	1.18 (1.13 – 1.22) *	-
Digestive and Accessory						
Oral and oropharyngeal	374	4.84 (3.65 – 6.42) *†	<0.01	291	1.66 (1.21 – 2.27) *†	0.02
Esophageal	599	3.25 (2.58 – 4.10) *†	<0.01	460	1.67 (1.32 – 2.12) *†	<0.01
Stomach	525	1.74 (1.28 – 2.38) *†	<0.01	301	1.15 (0.91 – 1.45)	1
Colorectal	2,572	1.37 (1.22 – 1.55) *†	<0.01	1,441	1.23 (1.12 – 1.35) *†	<0.01
Liver	761	2.09 (1.72 – 2.55) *†	<0.01	489	1.45 (1.20 – 1.75) *†	<0.01
Pancreas	1,607	2.04 (1.78 – 2.35) *†	<0.01	956	1.14 (0.99 – 1.32)	0.74
Sex Specific Organs						
Breast	2,099	1.11 (0.99 – 1.26)	0.40	949	1.12 (1.00 – 1.27)	0.62
Cervical	237	1.53 (1.13 – 2.08) *†	0.04	115	1.04 (0.70 – 1.53)	1
Ovarian	392	0.97 (0.79 – 1.18)	1	121	1.05 (0.87 – 1.27)	1
Uterine	750	0.68 (0.50 – 0.93)	0.02	317	0.65 (0.48 – 0.87)	0.20
Prostate	1,215	1.27 (1.05 – 1.54) *	0.09	802	0.99 (0.86 – 1.14)	1
Urinary						
Kidney	603	1.34 (1.06 – 1.69) *	0.09	359	1.10 (0.90 – 1.35)	1
Bladder	589	4.08 (3.20 – 5.20) *†	<0.01	451	2.39 (1.89 – 3.01) *†	<0.01
Lymphoid						
Hodgkin lymphoma	59	0.94 (0.45 – 1.96)	1	31	1.04 (0.51 – 2.09)	1
NHL	1,016	1.13 (0.93 – 1.37)	0.82	558	1.18 (1.01 – 1.38) *	0.48
Leukemia	970	1.23 (1.01 – 1.52) *	0.27	564	1.24 (1.04 – 1.47) *	0.19
Multiple Myeloma	541	0.76 (0.58 – 0.99)	0.27	270	0.95 (0.76 – 1.18)	1
Other Cancers						
Laryngeal	157	10.27 (5.45 – 19.36) *†	<0.01	142	3.04 (1.46 – 6.32) *†	0.04
Melanoma	392	1.01 (0.76 – 1.33)	1	213	0.99 (0.77 – 1.28)	1
Brain	622	1.37 (1.07 – 1.74) *	0.07	344	1.03 (0.83 – 1.27)	1
Unspecified Cancers	2,952	2.14 (1.93 – 2.38) *†	<0.01	1,858	1.27 (1.16 – 1.39) *†	<0.01

Note that a p-value of 1 indicates a value greater than 0.9999 as reported by SAS PROC MULTTEST.
* Significant at 95% confidence level using the unadjusted p-values
† Significant at 95% confidence level using Holm's adjusted p-values

DISCUSSION

Consistent with a growing body of literature, this study provides evidence that cancer mortality is associated with $PM_{2.5}$ exposure in both smokers and never-smokers. Analysis of the full cohort resulted in a hazard ratio of 1.15 (95% confidence interval of 1.08–1.22), which was comparable to that of the never-smokers' cohort (HR 1.19, 95% CI: 1.06–1.33). The result was comparable to a cohort that used 18.9 million Medicare beneficiaries (HR 1.11, 95% CI: 1.09-1.12) [32]. Analysis of the full cohort for non-lung cancers resulted in a hazard ratio of 1.15(95% CI: 1.07-1.24) which was also comparable to the cohort of never-smokers' (HR 1.15, 95% CI: 1.02-1.30). The results for non-lung cancer are much larger than other cohort studies like the Harvard Six Cities Study (HR 1.04, 95% CI: 0.86-1.26) [33], the ACS study (HR 1.06, 95% CI: 1.00-1.129) [33], but not statistically different.

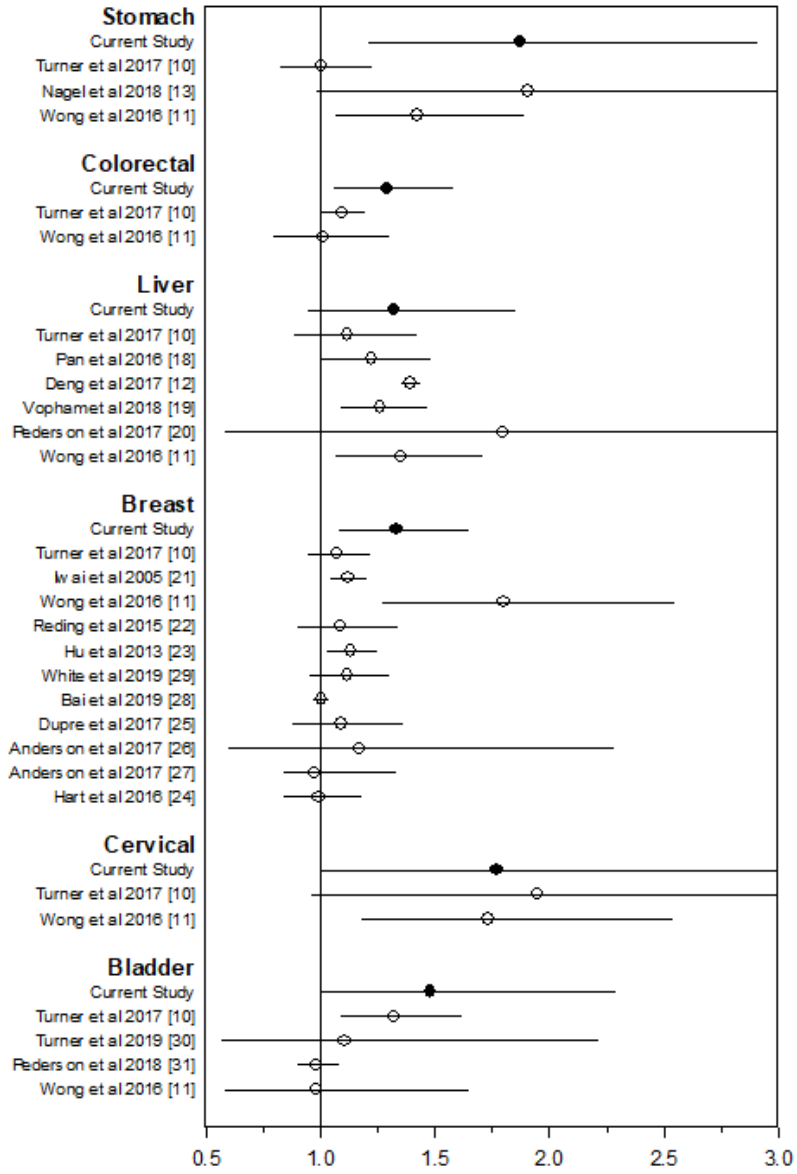
Furthermore, this study provides strong evidence that lung cancer is likely the primary driver of the association between cancer mortality and $PM_{2.5}$. The study found a hazard ratio of 1.13 (95% CI: 1.00-2.60) in the full cohort and a hazard ratio of 1.73 (95% CI: 1.20-2.49) in the never-smokers cohort, which was significant even after multiple comparison adjustment. The $PM_{2.5}$-lung cancer mortality hazard ratio was higher in the never-smokers' cohort than the in full cohort. The larger hazard ratio may be due to a lower baseline mortality risk for lung cancer among never-smokers or a limited sample size. The results from this study are comparable to a recent meta-analysis of cohorts examining $PM_{2.5}$-lung cancer mortality (HR 1.13, 95% CI: 1.07-1.20) [34].

The association between $PM_{2.5}$ and mortality due to non-lung cancers is less clear. Although the study did identify several cancer types (stomach, colorectal, liver, breast,

cervical, and bladder cancers and Hodgkin's lymphoma, NHL, and leukemia) that were associated with $PM_{2.5}$ exposure, none were statistically significant after adjusting for multiple comparisons. Other studies have reported $PM_{2.5}$-mortaltiy associations with other cancers, including multiple studies for stomach cancer [10-11,13], colorectal cancer [10-11], liver cancer [10-12,18-20], breast cancer [10-11,21-29], cervical cancer [10-11], and bladder cancer [10-11,30-31]. Comparisons of the estimated hazard ratios, risk ratios, incident rate ratios, and odds ratios (with their associated confidence intervals) for these cancers are succinctly illustrated in Figure 2. Although there is substantial heterogeneity across study estimates, the results of this study provide additional evidence to the growing body of literature that $PM_{2.5}$ exposure is associated with cancer mortality for lung and some non-lung cancers.

The results are also consistent with existing literature on the relationship between smoking and cancer [14], finding statistically significant associations after multiple testing adjustment for lung, oral and oropharyngeal, esophageal, stomach, colorectal, liver, pancreatic, laryngeal, cervical, kidney, bladder, and unspecified cancers. With the exception of Hodgkin's and non-Hodgkin's lymphoma, cancer sites that were statistically associated with $PM_{2.5}$ in either cohort were also associated with smoking status. This study also provides moderate evidence for the formal establishment of prostate, breast, and unspecified cancers as caused by smoking [35]., cigarette smoking and $PM_{2.5}$ exposure may both be risk factors, with cigarette smoking having a larger impact. Further research is needed to determine the relationship between $PM_{2.5}$, smoking, and cancer-site mortality.

Figure 2. Illustration of the comparison between the Hazard Ratio and 95% CI of full cohort of the current study and hazard ratios [10-13,18,20,22-29,31], risk ratios [21], incident rate ratios [19], and odds ratios [30] of other similar studies that estimated the association between a 10 μg/m^3 increase of $PM_{2.5}$ and various cancer sites. The Wong et al 2016 study includes esophageal cancer in its evaluation of stomach cancer, pancreatic with liver cancer, ovarian and uterine with cervical cancer, and kidney with bladder cancer.



A limitation of this study is the inability to directly measure exposure to ambient air pollution over a lifetime. With extensive follow-up and advanced ground-based monitoring and related modeling, this study used direct exposure estimates from 1999-

2015. However, it does not directly account for exposure before this period. Although back casted estimates of $PM_{2.5}$ exposure and only including individuals surveyed after 1999 are similar to the original model, the estimates of the hazard ratios may still be biased. Another limitation is the inability to control for migration. The migration problem is further exacerbated by the long latency period of some cancer types. In future studies, cancer incidence data could be used to reduce the latency concern. Additionally, this study did not control for other pollutants such as NO_2, SO_2, or CO. Other studies have found associations between pollutants other than $PM_{2.5}$ and incidence and mortality from various cancers [4, 36-37]. Future studies should control for these pollutants.

Another limitation of the study is the potential of residual confounding. The study was unable to control for several important variables such as secondhand smoke, HPV status, occupational exposure, hormonal therapy, oral contraceptive use, menopausal status, alcohol consumption, dietary patterns, and genetic variables that are associated with some cancer types. However, most cancer types were not sensitive to individual risk factors such as age, sex, race and ethnicity, education, income, geographic variables, and survey years, which suggests negligible risk of residual confounding. Furthermore, average air pollution was generally consistent across the factor levels for the individual risk factors, which suggests air pollution is less likely to be correlated with other omitted variables.

A final limitation is the lack of follow-up and quantitative measurements in the smoking data. The lack of follow-up would likely bias the estimates for smoking downwards because the number of smokers is decreasing in America. Future studies should also include quantitative measurements for smoking such as packs per day or

number of years smoking. Although these weaknesses may call the results of the current smoking status-mortality analysis into question, many of the cancer types that were associated with current smokers are also associated with former smokers, so the lack of follow-up and number of years smoked is less concerning. Furthermore, $PM_{2.5}$-cancer associations were similar in the never-smokers' cohort, which suggests little risk of bias.

This study has several important strengths. First, the study uses a cohort that is a representative sample of US adults with high quality survey information. Second, the cohort is large and contains many deaths for most cancer types. Third, the analysis can control for individual risk factors for cancer such as smoking and BMI. Fourth, results were generally not sensitive to cohort selection or modeling approaches. Fifth, the results for the association between cancer and lung cancer mortality and $PM_{2.5}$ were generally comparable to previous literature. Sixth, air pollution estimates and most other analysis variables are publicly available.

Exposure to $PM_{2.5}$ air pollution is a risk factor for lung cancer mortality and a possible risk factor in mortality for various other cancer types. This analysis confirms previous literature that cigarette smoking is associated with many cancer types. The results from the current study and comparable studies suggest that $PM_{2.5}$ may be associated with stomach, colorectal, liver, breast and cervical cancer. All these cancers were associated with smoking in the analysis. Although this exploratory study does not provide definitive conclusions, the strength of the research design and the consistency of results across modeling choices suggest further research is needed into the additional biological pathways by which cancer in humans may be affected by $PM_{2.5.}$ The universal

nature of pollution exposure, and its consequences, makes further study essential to public health.

WORKS CITED

1. Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., et al (2017). Air pollution and mortality in the Medicare population. *New England Journal of Medicine*, 376(26), 2513-2522.
2. Pope III, C. A., Lefler, J. S., Ezzati, M., Higbee, J. D., Marshall, J. D., Kim, S. Y., et al (2019). Mortality Risk and Fine Particulate Air Pollution in a Large, Representative Cohort of US Adults. *Environmental health perspectives*, 127(7), 077007.
3. IARC. 2013. Outdoor air pollution. IARC Monogr Eval Carcinog Risk Hum 109: 1.–. https://monographs.iarc.fr/ENG/Monographs/vol109/mono109-F01.pdf [accessed 19 March 2016].
4. Turner, M. C., Krewski, D., Pope III, C. A., Chen, Y., Gapstur, S. M., & Thun, M. J. (2011). Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. *American journal of respiratory and critical care medicine*, *184*(12), 1374-1381.
5. Hamra, G. B., Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, J. M., et al (2014). Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. *Environmental health perspectives*.
6. Cui, P., Huang, Y., Han, J., Song, F., & Chen, K. (2014). Ambient particulate matter and lung cancer incidence and mortality: a meta-analysis of prospective studies. *The European Journal of Public Health*, 25(2), 324-329.
7. Loomis, D., Grosse, Y., Lauby-Secretan, B., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., et al (2013). The carcinogenicity of outdoor air pollution. *The lancet oncology*, *14*(13), 1262-1263.
8. Risom, L., Møller, P., & Loft, S. (2005). Oxidative stress-induced DNA damage by particulate air pollution. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *592*(1), 119-137.
9. Newby, D. E., Mannucci, P. M., Tell, G. S., Baccarelli, A. A., Brook, R. D., Donaldson, K., et al (2014). Expert position paper on air pollution and cardiovascular disease. *European heart journal*, *36*(2), 83-93.
10. Turner, M. C., Krewski, D., Diver, W. R., Pope III, C. A., Burnett, R. T., Jerrett, M., et al (2017). Ambient air pollution and cancer mortality in the Cancer Prevention Study II. *Environmental health perspectives*, *125*(8).
11. Wong, C. M., Tsang, H., Lai, H. K., Thomas, G. N., Lam, K. B., Chan, K. P., et al (2016). Cancer mortality risks from long-term exposure to ambient fine particle. *Cancer Epidemiology and Prevention Biomarkers,* 25(5), 839-845.
12. Deng, H., Eckel, S. P., Liu, L., Lurmann, F. W., Cockburn, M. G., & Gilliland, F. D. (2017). Particulate matter air pollution and liver cancer survival. *International journal of cancer*, *141*(4), 744-749.
13. Nagel, G., Stafoggia, M., Pedersen, M., Andersen, Z. J., Galassi, C., Munkenast, J., et al (2018). Air pollution and incidence of cancers of the stomach and the upper aerodigestive tract in the European Study of Cohorts for Air Pollution Effects (ESCAPE). *International journal of cancer*.
14. U.S. Department of Health and Human Services. The health consequences of smoking: 50 years of progress. a report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2014. Printed with corrections, January 2014.
15. Kim S-Y, Bechle M, Hankey S, Sheppard EA, Szpiro AA, Marshall JD. 2018. Concentrations of criteria pollutants in the contiguous U.S., 1979 – 2015: Role of model parsimony in integrated empirical geographic regression. (November 2018). UW Biostatistics Working Paper Series. Working Paper 425. https://biostats.bepress.com/uwbiostat/paper425.

16. National Center for Health Statistics (NCHS). 2015. 2014 Survey Description, National Health Interview Survey (NHIS). Division of Health Interview Statistics. Hyattsville, Maryland. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2014/srvydesc.pdf. (accessed 31 August 2018).
17. Hochberg, Y. and Benjamini, Y. (1990), More powerful procedures for multiple significance testing. *Statistics in Medicine*, 9, 811–818.
18. Pan, W. C., Wu, C. D., Chen, M. J., Huang, Y. T., Chen, C. J., Su, H. J., & Yang, H. I. (2016). Fine particle pollution, alanine transaminase, and liver cancer: a Taiwanese prospective cohort study (REVEAL-HBV). *JNCI: Journal of the National Cancer Institute*, *108*(3).
19. VoPham, T., Bertrand, K. A., Tamimi, R. M., Laden, F., & Hart, J. E. (2018). Ambient PM 2.5 air pollution exposure and hepatocellular carcinoma incidence in the United States. *Cancer Causes & Control*, *29*(6), 563-572.
20. Pedersen, M., Andersen, Z. J., Stafoggia, M., Weinmayr, G., Galassi, C., Sørensen, M., et al (2017). Ambient air pollution and primary liver cancer incidence in four European cohorts within the ESCAPE project. *Environmental research*, *154*, 226-233.
21. Iwai, K., Mizuno, S., Miyasaka, Y., & Mori, T. (2005). Correlation between suspended particles in the environmental air and causes of disease among inhabitants: cross-sectional studies using the vital statistics and air pollution data in Japan. *Environmental Research*, *99*(1), 106-117.
22. Reding, K. W., Young, M. T., Szpiro, A. A., Han, C. J., DeRoo, L. A., Weinberg, C., et al (2015). Breast cancer risk in relation to ambient air pollution exposure at residences in the Sister Study Cohort. *Cancer Epidemiology and Prevention Biomarkers*, cebp-0787.
23. Hu, H., Dailey, A. B., Kan, H., & Xu, X. (2013). The effect of atmospheric particulate matter on survival of breast cancer among US females. *Breast cancer research and treatment*, *139*(1), 217-226.
24. Hart, J. E., Bertrand, K. A., DuPre, N., James, P., Vieira, V. M., Tamimi, R. M., 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 Epidemiology and Prevention Biomarkers, cebp-0246.
25. DuPre, N., Poole, E. M., Holmes, M. D., Hart, J. E., James, P., Beck, A., et al (2017). Particulate matter and traffic-related exposures in relation to breast cancer survival.
26. Andersen, Z. J., Stafoggia, M., Weinmayr, G., Pedersen, M., Galassi, C., Jørgensen, J. T.,et al(2017). Long-term exposure to ambient air pollution and incidence of postmenopausal breast cancer in 15 European cohorts within the ESCAPE project. *Environmental health perspectives*, *125*(10), 107005.
27. Andersen, Z. J., Ravnskjær, L., Andersen, K. K., Loft, S., Brandt, J., Becker, T., ... & Bräuner, E. V. (2017). Long-term exposure to fine particulate matter and breast cancer incidence in the Danish Nurse Cohort Study. *Cancer epidemiology and prevention biomarkers*, *26*(3), 428-430.
28. Bai, L., Shin, S., Burnett, R. T., Kwong, J. C., Hystad, P., van Donkelaar, A., et al (2019). Exposure to ambient air pollution and the incidence of lung cancer and breast cancer in the Ontario Population Health and Environment Cohort (ONPHEC). *International journal of cancer*.
29. White, A. J., Keller, J. P., Zhao, S., Carroll, R., Kaufman, J. D., & Sandler, D. P. (2019). Air Pollution, Clustering of Particulate Matter Components, and Breast Cancer in the Sister Study: A US-Wide Cohort. *Environmental health perspectives*, *127*(10), 107002.
30. Turner, M. C., Gracia-Lavedan, E., Cirac, M., Castaño-Vinyals, G., Malats, N., Tardon, A., et al (2019). Ambient air pollution and incident bladder cancer risk: Updated analysis of the Spanish Bladder Cancer Study. *International journal of cancer*, *145*(4), 894-900.
31. Pedersen, M., Stafoggia, M., Weinmayr, G., Andersen, Z. J., Galassi, C., Sommar, J., et al (2018). Is there an association between ambient air pollution and bladder cancer incidence? Analysis of 15 European Cohorts. *European urology focus*, 4(1), 113-120.

32. Pun, V. C., Kazemiparkouhi, F., Manjourides, J., & Suh, H. H. (2017). Long-term PM2. 5 exposure and respiratory, cancer, and cardiovascular mortality in older US adults. *American journal of epidemiology*, *186*(8), 961-969.
33. Krewski, D., Burnett, R., Goldberg, M., et al. (2003). Overview of the reanalysis of the Harvard six cities study and American Cancer Society study of particulate air pollution and mortality. *Journal of Toxicology and Environmental Health Part A*, *66*(16-19), 1507-1552.
34. Pope III, C. A., Coleman, N., Pond, Z. A., & Burnett, R. T. (2019). Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environmental Research*, 108924.
35. Carter, B. D., Abnet, C. C., Feskanich, D., et al. (2015). Smoking and mortality—beyond established causes. *New England journal of medicine*, *372*(7), 631-640.
36. Raaschou-Nielsen, O., Andersen, Z. J., Hvidberg, M., Jensen, S. S., Ketzel, M., Sørensen, M., et al (2011). Air pollution from traffic and cancer incidence: a Danish cohort study. *Environmental Health*, *10*(1), 67.
37. 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. *Environmental research*, *137*, 467-474.

APPENDIX 1

Table S1. Estimated hazard ratios (95% CIs) from sensitivity analysis performed on the full cohort for statistically significant cancers in Table 2. The base model is a model using traditional basic Cox Proportional Hazards model (using the Proc PHREG procedure in SAS version 9.3) and controlling for all combinations of 1-yr age groups, sex and race-ethnicity by allowing them to have their own baseline hazard (by including them in the STRATA statement). The individual model is the base model, but with indicator variables for education, income, marital status, BMI, and smoking status also added as covariates in the model. The full model is the individual model, but with indicator variables for urban/rural, census region, and survey year also added as covariates. The back casted model is the full model, with mean $PM_{2.5}$ data back casted to 1988 (i.e. exposure window of 1988-2015 rather than 1999-2015). The ≥1999 Survey Years model is the full model, but only with survey years from 1999-2014. The expanded cohort is the full model, but it uses the expanded cohort (all 1,599,329 NHIS participants from 1986-2014, including those without smoking or BMI data) and not controlling for smoking status or BMI.

Cancer	Traditional Model					
	Base	Individual	Full	Back casted	≥1999 Survey Years	Expanded Cohort
Lung	1.22 (1.10-1.35)	1.14 (1.03-1.26)	1.12 (1.00-1.25)	1.10 (1.00-1.20)	1.16 (0.95-1.41)	1.09 (1.02-1.17)
Stomach	1.81 (1.24-2.63)	1.78 (1.22-2.58)	1.82 (1.24-2.69)	1.63 (1.20-2.22)	2.18 (1.10-4.33)	1.80 (1.45-2.24)
Colorectal	1.36 (1.14-1.61)	1.32 (1.11-1.57)	1.23 (1.02-1.47)	1.18 (1.02-1.36)	1.28 (0.92-1.77)	1.26 (1.12-1.41)
Liver	1.38 (1.02-1.87)	1.37 (1.01-1.85)	1.35 (0.99-1.84)	1.22 (0.95-1.56)	1.39 (0.89-2.17)	1.33 (1.10-1.60)
Cervix	2.45 (1.42-4.22)	2.18 (1.26-3.78)	2.22 (1.27-3.88)	1.69 (1.09-2.63)	2.04 (0.80-5.24)	1.71 (1.20-2.44)
Breast	1.28 (1.06-1.55)	1.28 (1.06-1.55)	1.26 (1.03-1.54)	1.22 (1.04-1.43)	1.38 (0.97-1.96)	1.28 (1.13-1.45)
Bladder	1.29 (0.90-1.84)	1.26 (0.88-1.80)	1.26 (0.86-1.84)	1.22 (0.91-1.65)	1.49 (0.80-2.78)	1.00 (0.79-1.27)
Hodgkin	3.59 (1.22-10.56)	3.45 (1.18-10.08)	3.22 (1.05-9.86)	2.33 (0.96-5.67)	11.27 (1.45-87.79)	1.57 (0.78-3.14)
NHL	1.59 (1.21-2.08)	1.61 (1.22-2.11)	1.49 (1.12-1.98)	1.34 (1.07-1.68)	1.94 (1.18-3.18)	1.35 (1.14-1.61)
Leukemia	1.30 (0.98-1.72)	1.34 (1.01-1.77)	1.27 (0.94-1.71)	1.25 (0.99-1.58)	1.92 (1.17-3.15)	1.22 (1.02-1.46)