Air pollution and lung cancer survival in Pennsylvania

Thomas P. McKeon[a,b,c,*], **Anil Vachani**[a,d,e], **Trevor M. Penning**[a,b,d], **Wei-Ting Hwang**[a,d,f]

[a]Center of Excellence in Environmental Toxicology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[b]Departments of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[c]Department of Geography and Urban Studies, Temple University, Philadelphia, PA 19122, USA

[d]Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[e]Department of Medicine, Pulmonary, Allergy, and Critical Care Division, Hospital of University of Pennsylvania, Philadelphia, PA 19104, USA

[f]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

Objective: Lung cancer is a leading cause of cancer death in the United States. Exposure to outdoor air pollution (OAP) is associated with increased lung cancer incidence, however little is known about the association of OAP and survival after diagnosis.

Methods: We investigated the effects of OAP and lung cancer survival in Pennsylvania using data from Pennsylvania Cancer Registry. The study population consisted of 252,123 patients diagnosed between 1990 and 2017. The Environmental Protection Agency's ambient air monitoring network provided information on OAP exposure of NO_2, O_3, $PM_{2.5}$, and PM_{10}. Mean OAP exposures were calculated by interpolating exposure concentrations from the five nearest monitors within a 50-kilometer radius of each patient's residential address from date of diagnosis to date of death or last contact. Cox proportional-hazards models were used to estimate the hazard ratios (HR) for OAP exposures for overall and lung cancer-specific survival. Statistical analyses

*Corresponding author at: Department of Systems Pharmacology & Translational Therapeutics 1315, BRBII/III, 421 Curie Blvd, Philadelphia, PA 19104, USA. mckeont@pennmedicine.upenn.edu (T.P. McKeon).

were stratified by SEER cancer stage groupings (localized, regional, and distant) and adjusted for individual-level and area-level covariates.

Results: Median survival time was 0.76 [CIs: 0.75, 0.77] years for the study population and for localized, regional, and distant site diagnosis were 2.2 [CIs: 2.17, 2.23], 1.13 [CIs: 1.12, 1.15], and 0.42 [CIs: 0.41, 0.43] years, respectively. NO_2 indicated the greatest HR which increased with increasing magnitude of exposure across all cancer staging groups for deaths before 2-years post-diagnosis. HRs varied by stage and magnitude of OAP exposure with greatest overall effects shown in NO_2 followed by $PM_{2.5}$, O_3, and PM_{10}. A subgroup analysis of patients with treatment status information (2010–2017) showed similar associations of increasing HRs with increasing exposure.

Conclusion: These findings supported the hypotheses that OAP can influence the carcinogenic process, impairing chemotherapy treatment, and provide important public health implications since environmental factors are not often considered in prognosis of survival after diagnosis.

Keywords

Lung cancer; Air pollution; Survival analysis; Geospatial; Exposomics; Public health

1. Introduction

Lung cancer is the second most commonly diagnosed cancer (11.4%) and was the leading cause of death (18.0%) of all cancers worldwide in 2020[1]. In men, lung cancer ranks first in incidence and mortality, and in females, it ranks third in incidence and second for mortality[1,2]. In the United States, lung cancer stands as the third most common cancer, and the leading cause of cancer death in both men and women[3]. Nationwide, new lung cancer cases have decreased from 1999 to 2018, possibly impacted by changes in risk exposures such as reduced tobacco use and improved protection to occupational hazards[4,5]. In the Commonwealth of Pennsylvania, lung cancer mortality has decreased from 54.5 per 100,000 persons in 1999 to 37.4 per 100,000 persons in 2018[3]. However, Pennsylvania's rate of lung cancer deaths is stubbornly higher than the national average of 34.8 per 100,000 persons in 2018[3].

Air pollution, especially year-long outdoor particle pollution and other components of air pollution with evidence of modifying or influencing the carcinogenic process, are known risk factors for lung cancer [6,7,8]. In 2013, the World Health Organization (WHO)'s International Agency for Research on Cancer (IARC) concluded that there is indisputable evidence that air pollution is a human carcinogen and accounts for 230,000 lung cancer deaths per year globally[9,10]. Research investigating residential proximity to major roadways, ozone, and NO_x emissions suggest an increased risk of lung cancer attributable to air pollution exposure[11,12,13]. Considering that air pollution is classified as a group 1 carcinogen, signifying that poor outdoor air quality has clear evidence of causing cancer, it is important to know how levels of measurable air pollution exposures differ across the country and to identify vulnerable populations most at risk[14]. The American Lung Association (ALA) ranks the air quality of U.S. cities annually based on ozone and particle pollution. The ALA's "State of the Air 2021″ report using air quality data

during 2017, 2018, and 2019, signaled that Pennsylvania's major metropolitan areas around Pittsburgh and Philadelphia currently rank in the top 20 worst cities for air quality[15]. This is worrisome because exposure to outdoor air pollution is associated with increased lung cancer occurrence in both smokers and non-smokers [16,17].

Not much is known about the association between air pollution exposure and length of survival after a lung cancer diagnosis. Following the rationale that components of air pollution increase the risk of developing lung cancer, we hypothesized that exposure to air pollution would worsen survival after lung cancer diagnosis. Air pollution exposure causes oxidative stress and inflammation which are important in tumor promotion[18]. Of particular importance is the activation of Nuclear factor-erythroid factor 2-related factor 2 (NRF2), an important transcription pathway that regulates the expression of over 1,000 genes in the cell under normal and stressed conditions. When activated, NRF2 can result in resistance to cancer chemotherapeutic agents used to treat lung cancer patients[19,20].

A recent study that investigated the effects of air pollution on lung cancer survival in the state of California supports the hypothesis that air pollution exposures after lung cancer diagnosis shortens survival[21]. However, to the best of our knowledge, no one has investigated how exposure to outdoor air pollution impacts lung cancer survival after diagnosis in other US states including Pennsylvania. This is particularly important given that along with California, Pennsylvania has had historically some of the worst cities for air pollution seen across the nation, home to metropolitan areas ranking in the top 25 for ozone, year-round particle pollution, and short-term particle pollution[22]. The five-year survival rate for lung cancer in Pennsylvania is 23.7%, although this has improved by 14.5% over the past five years[23]. Improvements in Pennsylvania's five-year survival rate may be the result of improved cancer treatment options and improvements in the state's average overall air quality. Still, the five-year survival rate for lung in Pennsylvania cancer remains one of the worst of all cancers[24].

New approaches to improve survival after diagnosis are needed. An investigation into differences in lung cancer survival based on air pollution exposure may elucidate the contribution of poor outdoor air quality on lung cancer mortality at varying exposure levels. In order to determine the contribution of air pollution on lung cancer survival, we conducted a survival analysis using a cohort of geocoded Pennsylvania residents that have been diagnosed with lung cancer, with an assessment of their air pollution exposure using air monitoring network data from the Environmental Protection Agency (EPA).

2. Methods

2.1. Study population and individual-level data

The Pennsylvania Cancer Registry (PCR), a statewide data system responsible for collecting information on all new cases of cancer diagnosed or treated in Pennsylvania, provided the data used in this study including patient- and tumor-specific characteristics for individual cases. PCR has earned Gold Certification from the North American Association of Central Cancer Registries (NAACCR), the highest level of data quality achieving at least 95% completeness, for all years of data available except 2001 (not certified) and 2012 (Silver

Certification with at least 85% completeness). Patients diagnosed with lung cancer, coded as C34 in the ICD-O-3 behavior code, and diagnosed during the 1990–2017 timeframe made up our study population. The PCR gathers individual-level data on demographics (age, sex, race/ethnicity, date of diagnosis, date of last contact, tumor characteristics, vital status, cause of death, and residential address at time of diagnosis).

The following exclusion criteria were applied: Patients with incomplete or missing diagnosis/last contact dates (N = 33,552); patients with negative and zero survival times (N = 7,851); patients with diagnosis of in-situ and non-carcinoma histology (N = 13,351); patients with nonbinary sex (N = 5); and patients without nearby air pollution exposure assignments (N = 728). We chose 1990 as the start date for this study because adequate monitoring information for air pollution exposures was not available for most patients before that time. After excluding patients who did not meet our inclusion criteria, a total of 252,123 patients were included in the analyses.

2.2. Geocoding

Environmental Systems Research Institute (ESRI)'s ArcGIS, a commercial geocoding system commonly used in health research[25,26] was used to geocode patient residential addresses reported at the date of diagnosis to point-level coordinates (latitude and longitude). Incomplete addresses were geocoded to the centroid of the smallest resolvable area based on the address completeness.

2.3. Air pollution exposure

The exposure variable of interest is the air pollution associated with four air pollutant categories: nitrogen dioxide (NO_2, parts per billion) ozone (O_3, parts per billion), fine particulate matter with diameter < 10 μm (PM_{10}, μg/m^3), and non-Federal Reference Methods / Federal Equivalent Methods (FEM/FRM) fine particulate matter with diameter < 2.5 μm ($PM_{2.5}$, μg/m^3). The air pollution exposure data were obtained from The US-Environmental Protection Agency's Air Quality System (AQS) database. Non-FEM/FRM $PM_{2.5}$ values were used because data on FEM/FRM were not available until 1998. Air pollution information from the EPA was presented as hourly measurements summarized as 24-hour averages for NO_2, PM_{10}, $PM_{2.5}$, and average 8-hour daily maximum for O_3. For each patient, the monthly residential ambient air pollution exposure from the month of diagnosis to the month of last follow-up or death was interpolated using the five closest air quality monitoring stations within a 50-kilometer (km) radius of the patient's geocoded residence based on the inverse distance weighting method within R package *gstat*. The mean exposure for the patient was calculated as the sum of the interpolated monthly concentrations divided by the number of months the patient was followed. Exposure status for a patient to an air pollutant for statistical analyses was assigned into one of the four categories (Q1, Q2, Q3, Q4) according to the quartiles of the mean exposure distribution for the whole study samples. Because there is an overall decreasing trend for all pollutants over time and the decreasing trends were, in general, the same for all areas [27,28], we believe the average annual exposure over the follow-up/survival duration preserves the high/low exposure differences between patients while avoiding minor fluctuations in the exposures, and can be more easily interpreted. We used the quartiles based on the average annual

exposure over the follow-up/survival duration to represent the average exposure for a given patient following diagnosis instead of treating the exposure as a time-varying covariate. Further examining the functional form of the continuous version of the average annual exposure variables, as described in the statistical section below, promoted the use of the categorized version of the exposures.

2.4. Area-level covariates

Area-level covariates were assigned to each patient at the census tract level. The Area Deprivation Index (ADI), a well-established measure of neighborhood disadvantage that includes factors of income, education, employment, and housing quality was considered[29,30]. The ADI values range from 1 to 100 with a ranking of 1 indicating the lowest level of "disadvantage" within the nation and an ADI with a ranking of 100 indicating the highest level of "disadvantage". The ADI index values were categorized into quintiles assigned into the following groupings: "Lowest", "Lower-middle", "Middle", "Higher-middle", "Highest", and "Unknown". Rural-urban community area (RUCA) codes based on the 2010 decennial census from the U.S. Department of Agriculture Economic Research Service were used to classify patient census tracts into ordinal ranks (1 to 10) of metropolitan (rank 1) to rural (rank 10)[31]. Each patient's estimated educational attainment was derived from the 2000 U.S. Census tract information on the percentage of adults 25 years and older with at least a bachelor's degree[32]. Educational attainment was categorized into quartiles of "low", "low-medium", "medium-high", and "high".

2.5. Survival outcome

Overall survival was calculated as the time between the date of lung cancer diagnosis to the date of death from any cause or censored at the date of the last contact if alive. As a sensitivity analysis, lung cancer-specific survival was calculated as the time between the date of lung cancer diagnosis to the date of death due to lung cancer (ICD-9 code 1629 for years before 1998 and ICD-10 code C349 for years after 1998). Deaths due to other causes were treated as a competing risk. All survival times were converted from survival in days into years by dividing the days by 365.25.

2.6. Statistical analysis

Descriptive statistics were calculated for the patient characteristics, air pollution exposures, and covariates. Patients were stratified into four cancer stages at diagnosis based on metrics from the National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER). The following SEER cancer stages at diagnosis were used: Localized (N = 52,792); Regional (N = 62,952); Distant (N = 121,888), and Unknown (N = 14,491).

Kaplan-Meier estimates were used to estimate the overall survival function of the total patient population and each SEER cancer stage grouping. Median survival and 5-year survival probabilities were calculated for the whole group and each stage group separately. We used Cox proportional hazard models to estimate the hazard ratios (HR) between overall survival and categorized air pollution exposures and cause-specific Cox proportional hazard model for lung cancer-specific survival while adjusting for potential confounders. Separate models were generated for each SEER cancer stage. Potential individual and

area-level confounding variables were included in the model for adjustment including age (continuous); sex; race/ethnicity (Hispanic, Non-Hispanic-Black, Non-Hispanic White, or other); categorized percent of those 25 years of age or greater with Bachelor's; categorized Area Deprivation Index; categorized RUCA codes (Urban, Metropolitan, Other); categorized histology codes (Squamous cell carcinoma, adenocarcinoma, small cell carcinoma, large cell carcinoma and other carcinomas); categorized year of diagnosis (1990–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2017); and month of diagnosis. To account for potential confounding of pollution caused by local traffic, we adjusted for the distance from residential address to a primary interstate highway, categorized into (<300 m, 300–1500 m, > 1500 m). The ICD-O-3 morphology codes for carcinoma allowed for grouping histology into the following five groupings: Squamous cell carcinoma (codes: 8050–8078,8083–8084); Adenocarcinoma (codes: 8140, 8211, 8230–8231, 8250–8260, 8323, 8480–8490, 8550–8551, 8570–8574, 8576); Small cell carcinoma (codes: 8041–8045, 8246); Large cell carcinoma (codes: 8010–8012, 8014–8031, 8035, 8310); and other carcinomas (remaining codes). The functional form of continuously measured covariates were examined by visually checking the smoothing plot of the martingale residuals for any nonlinear patterns. For each air pollutant exposure, we fitted a Cox proportional hazard models with all covariates and a polynomial spline for the continuously measured averaged annual exposure with two degrees to evaluate departures from a linear pattern or possiblly a dose–response relationships.

The proportional hazards assumption was evaluated by utilizing the $\log(-\log(S))$ plot based on the Kaplan-Meier (KM) estimates. This approach presents a graph of the log–log survival function versus the survival time, resulting in parallel lines if the categorized exposure assignments are proportional[32]. O_3 and PM_{10} satisfied the proportional hazard assumption, whereas NO_2 and $PM_{2.5}$ violated the proportional hazards assumptions and were adjusted by including a time-dependent covariate to allow for estimating two HRs for before and after the time when the survival curves cross: 2-year post-diagnosis for NO_2, and 5- year post-diagnosis for $PM_{2.5}$ [33].

A sub-group analysis was performed by restricting to patients with known treatment status. The rationale for this sub-group analysis is to determine whether increased exposure to air pollution worsens survival time post-diagnosis in the overall patient population and sub-group with treatment status. The PCR is only required to provide treatment status for patients diagnosed after 2010 and if treatment was given, information about the type of the treatment was not collected. For the current analysis, the PCR coded treatment status into the following four categories: "No treatment given", "Treatment given", "Active surveillance (watchful waiting)", or "unknown if treatment was given": 60,463 patients received treatment, and 13,019 patients did not receive treatment or active surveillance were included in the sub-group analysis. Patients with unknown treatment status were excluded. The model was adjusted for the same covariates, same time split, and an additional covariate indicating treatment status.

Statistical analysis was performed using R statistical software Version 4.0.0. Hypothesis tests were two-sided with statistical significance defined at $p < 0.05$.

3. Results

Descriptive statistics and characteristics of the total study population, stratified by cancer stage are presented in Table 1. 252,123 patients diagnosed with lung cancer between 1990 and 2017 were analyzed. On average, patients were majority male (55.44%), predominately non-Hispanic White (89.01%), and 68.89 years of age at the time of diagnosis. Patients diagnosed at an advanced stage (distant) lung cancer, made up the largest group (48.3%). The number of patients with "unknown" stage at diagnosis decreased from 1990 to 1995 to 2011–2017 in these time intervals. During the study period, 88% of the patients died of any cause, with 68.2% (N = 171,935) dying from lung cancer, and 19.8% (N = 49,970) dying from other causes. Median survival time was 0.76 [CIs: 0.75, 0.77] years for the entire study population and median survival times for localized, regional, and distant site diagnosis were 2.2 [CIs: 2.17, 2.23], 1.13 [CIs: 1.12, 1.15], and 0.42 [CIs: 0.41, 0.43] years, respectively. For patients with a localized stage at diagnosis, median survival was shortest for patients with small and large cell carcinomas (1.12 and 1.08 years, respectively), and longest for patients with adenocarcinoma (3.27 years) (Table 1).

Mean and standard deviation (SD) exposure to the four air pollution variables investigated in this study are shown in Table 1 for the study population stratified by cancer stage. Mean and SD exposure for covariates by cancer stage is shown in in Supplemental Table 1. Across all patients, air pollution exposure means and SDs were 8.14 ± 2.24 ppb for NO_2, 22.95 ± 3.72 ppb for O_3, 26.53 ± 5.26 μg/m^3 for PM_{10}, and 6.67 ± 1.72 μg/m^3 for $PM_{2.5}$. Exposure outcomes were similar across the staging groups.

The estimated median survival time and 5-year survival probability for each air pollution exposure quartile are shown in Table 2. In patients with a local or regional stage at diagnosis, higher NO_2, O_3, $PM_{2.5}$ exposures categories revealed shorter median survival and lower 5-year survival. For example, median survival for patients with a local stage at diagnosis was 1.27 years for those with high NO_2 exposure (>17.25 ppb) and 2.77 years for those with low NO_2 exposure (<6.22 ppb). Survival for patients with a distant stage at diagnosis was poor and showed little variation according to the level of air pollution exposure.

After adjusting for the beforementioned covariates, the estimated HRs from the Cox proportional hazards models in this study for local staging were associated with an increased risk of death and the greatest magnitudes of exposure. When treating air pollutants as continuous variable, increasing averaged annual exposure were associated with an increased risk of death for all air pollutants at increasing magnitudes of exposure. Associations varied by stage at diagnosis but similar HRs were estimated for lung-cancer specific survival (Table 3, Supplemental Table 2). We noted that NO_2 showed the strongest positive association for worsened overall survival, strengthening with increasing magnitude of exposure across all cancer staging groups before 2-years postdiagnosis, this was especially evident for the localized cancer staging group which presented the greatest HR for all pollutants at the greatest magnitude (Q4) of NO_2 exposure: < 2 years HR = 5.07 [CIs: 4.80, 5.35]. In contrast, the HRs for NO_2 showed a steady negative association for all cancer stagings after 2-years post-diagnosis.

Within each stage, the associations between overall survival and air-pollutant exposures varied, with NO_2 and $PM_{2.5}$ showing the strongest positive association before the 2- and 5-years cutoffs. (Fig. 1). The shape of the observed dose–response curves varied by air pollutant exposure category, yet remained similar within the same pollutant type across cancer staging, with the strongest positive associations presenting in the localized stage group for all pollutants, although this was minimal for O_3 and PM_{10}. NO_2 expressed a J-shaped curve for all stages < 2-years postdiagnosis. After 2-years post-diagnosis, a J-shaped curve was not observed for NO_2, and the hazard ratios consistently remained under 1.00 across the localized, regional, and distant cancer stage groupings. The O_3 and PM_{10} exposures presented strong "U-shaped" dose–response curves with similar HRs across cancer stages, with a slight decrease in HRs for the highest exposure quartile from localized to distant staging. $PM_{2.5}$ expressed a steady linear curve with increasing HRs for all staging groups within 5-years post-diagnosis. After 5-years post-diagnosis, the dose–response curves for $PM_{2.5}$ were "U-shaped" and showed little variation across the localized, regional, and distant staging groups and such patterns were also supported by the spline modeling. The adjusted HRs for all air-pollutant exposures were generally larger in small cell and large cell carcinoma when compared to other histology.

The HRs for overall survival and lung cancer-specific survival for the sub-group analysis are shown in Supplemetental Table 3. For overall survival, the estimated HRs for NO_2 exposure were similar to the HRs for the whole cohort for all cancer stages, before and after the 2-year time split, although significance was lost > 2 years post diagnosis. The HRs for O_3 revealed a stronger positive association across all cancer staging groups. PM_{10} presented a similar "U-shaped" curve with HRs with a stronger positive association for Q2 across all cancer staging groups. $PM_{2.5}$ showed a stronger positive association before the 5-year cut off, however after the 5-year cut off was not significant across all cancer stages. Similar patterns were observed for lung-cancer specific survival.

4. Discussion

In recent years the World Health Organization has classified air pollution as a group 1 carcinogen, and it is well known that air pollution is associated with lung cancer morbidity and mortality[9,10]. In recent years there has been interest to understand how ambient outdoor air pollution affects lung cancer survival after diagnosis, and how this differs by SEER cancer staging. As far as we know, no one has assessed the survival time of diagnosed lung cancer patients by incorporating data on concentrations of air pollutants around their residential locations in Pennsylvania. Using patient data from the PCR, we conducted a population-level study using 252,123 patients diagnosed with lung cancer between 1990 and 2017. We investigated the following four air pollutant exposure variables: NO_2, O_3, $PM_{2.5}$, and PM_{10} and their association with survival after diagnosis because these air pollutants have strong evidence of causing cancer directly, or influencing the carcinogenesis process, or by potentially activating NRF2 which may result in chemotherapeutic drug resistance[11,18,6,34].

Reduced survival time was observed for all air pollutant exposures at the highest average exposure quartile for the localized cancer staging group. We observed a lower survival time

from the highest mean exposure quartile for NO_2, O_3, and $PM_{2.5}$, for the regional cancer staging group. Although hazard ratios were larger for the highest average exposure in the early-stage cancer groupings, overall survival was poor for the late-stage cancer groupings at all exposure levels. The associations between air pollution and lung cancer survival after diagnosis were similar for both all-cause mortality and lung-cancer specific mortality. The adjusted HRs for all air-pollutant exposures were generally larger in small cell and large cell carcinoma when compared to other histology.

Following lung cancer screening uptake, interventions to reduce poor outdoor air quality exposure for those with early stage lung cancer may improve survival time. Our findings supported our hypothesis that air pollution is not only causing cancer but is also associated with reduced lung cancer survival. One possible explanation could be the role played by NRF2 activation. Fine particulates, especially particles from diesel exhaust produce oxidative stress which could activate the NRF2 pathway[35]. Activation of the NRF2 pathway leads to the elimination of cytotoxic species which are required for the efficacy of cancer chemotherapeutic agents. J-shaped dose response curves were observed for NO_2 exposure and poor lung cancer survival and "U-shaped" dose response curves were observed for O_3 and PM_{10} exposures and perhaps $PM_{2.5}$ as well. The J-shaped dose response curve for NO_2 may be attributed to its harmful effects once they exceed a threshold level. NO_2 has a wide range of harmful effects on the lungs including reduced lung function and inflammation[36], as well as the formation of nitric oxide (NO) which can give rise to peroxynitrite and nitrosoperoxycarbonate resulting in protein damage[36,37,38]. High levels of NO_2 also contain NO[39,40]. Insufficient evidence exists to link NO_2 to NRF2 activation, however, NO is linked with NRF2 activation[18].

The "U-shaped" dose–response curves from the low to high quartile categories for O_3 and PM_{10} exposures are reminiscent of the hormesis principle. In this principle, both low exposures and high exposures can have the same deleterious effect and is often seen with an essential nutrient. In the experimental setting, electrophilic and redox-active compounds, especially lung carcinogens, may induce a hermetic dose–response in the NRF2 system. This appears important in a real-world setting where populations may have chronic low exposure to these pollutants and basal levels of NRF2 activator may be insufficient to deal with the toxic insult. In addition, over activation of the NRF2 pathway can lead to changes in redox state to favor oncogenesis and promote cancer chemotherapeutic drug resistance[34]. A 2016 study investigating air pollution's effects on lung cancer survival in California also presented similar "U-shaped" curves in their findings, although they did not use timesplit covariates for NO_2 and $PM_{2.5}$, and cannot act as a direct comparison, their measurements were in the ballpark of our outcomes [21]. Other research investigating empirical data on $PM_{2.5}$ and mortality risks have supported the hypothesis that concentration–response relationships may not be linear [41]. A 2013 study highlighted the need to further investigate the plausibility of biology effecting length of survival due to air pollution exposure [42]. This American study, investigated the deleterious effects of air pollution exposure (O_3, $PM_{2.5}$, and PM_{10}) on length of survival from respiratory cancer patients by investigating two metropolitan areas with differing air pollution exposure levels. The research revealed that magnitude of exposure plays a role in survival, bringing to attention the need for preventive efforts to protect cancer patients from air pollution, and which is still something

unemphasized in cancer patient care [42]. If air pollution exposure can have nonmontonic dose–response relationships, then the hypothesis that cancer survival has a non-linear association with increasing magnitude of air pollution exposure is a phenonomen that may be worth further investigation in future studies.

Information about patient treatment status, or whether or not they received treatment for their lung cancer diagnosis is limited in the PCR. The PCR did not require reporting treatment status before 2010. Given that the interest of this study is to better understand the association of air pollution on lung cancer survival, the effect is apparent that increased exposure to air pollution worsens survival time post-diagnosis in the overall patient population and sub-group with treatment status. Future research investigating survival post lung cancer diagnosis in Pennsylvania would benefit from exploring treatment status effects on survival as more data becomes available.

Our study focused on Pennsylvania which has significantly higher rates of new lung cancer cases higher than the national average. Although improvements to air quality standards have occurred during the study period both nationally and statewide, Pennsylvania is still home to some of the worst metropolitan regions for air quality. Access to a large sample size from the PCR is a strength of this study.

Our study does come with limitations. This is an observational study so no definite conclusions can be drawn about cause and effect. EPA's air quality monitoring network is limited for the early years especially before 1990, creating more uncertainty in interpolating granular exposures in regions of the state with limited outdoor air monitoring. Additionally, the lack of individual covariate data such as, complete information on tumor staging (e.g., AJCC), cigarette smoking history, and other comorbidities (heart disease, diabetes, obesity, etc) are large predictors of survival and may also be associated with air pollution. Constraints in treatment status before 2010 make investigating the impact of treatment limited to more recent years. Patients with missing address information were geocoded to the centroid of the smallest resolvable area which provided less residential location certainty than the patients with complete street address information. Residential addresses are based on date of diagnosis and updated residential information is not provided for patients who have moved.

In summary, this study found evidence that exposure to air pollution after diagnosis affects lung cancer survival in Pennsylvania. This association was greatest for NO_2 and for those with localized cancer staging. This has important public health implications in patient care since environmental factors for lung cancer survival are minimally considered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Abbreviations:

OAP	Outdoor Air Pollution
EPA	Environmental Protection Agency
NO$_2$	Nitrogen Dioxide
NO$_x$	Nitrogen oxides
O$_3$	Ozone
PM$_{2.5}$	Fine Particulates < 2.5 mieometers
PM$_{10}$	Fine Particulates < 10 micometers
SEER	Surveillance, Epidemiology, and End Results Program
HR	Hazard Ratios
CIs	Confidence Intervals
WHO	World Health Organization
IARC	International Agency for Research on Cancer
FRM	Federal Reference Methods
FEM	Federal Equivalent Methods
ALA	American Lung Association
Nrf2	Nuclear factor-erythroid factor 2-related factor 2
PCR	Pennsylvania Cancer Registry
NAACCR	North American Association of Central Cancer Registries
KM	Kaplan-Meier
ICD	International Classification of Diseases

References

[1]. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 COUNTRIES, CA Cancer J. Clin 71 (3) (2021) 209–249, 10.3322/caac.21660. [PubMed: 33538338]

[2]. Howlader N, Forjaz G, Mooradian MJ, Meza R, Kong CY, Cronin KA, Mariotto AB, Lowy DR, Feuer EJ, The effect of advances in lung-cancer treatment on population mortality, N. Engl. J. Med 383 (7) (2020) 640–649, 10.1056/nejmoa1916623. [PubMed: 32786189]

[3]. Centers for Disease Control and Prevention. Lung cancer statistics. Centers for Disease Control and Prevention. https://www.cdc.gov/cancer/lung/statistics/index.htm, 2021 (accessed 4 January 2022).

[4]. National Cancer Institute (Surveillance, Epidemiology, and End Results Program). Cancer of the lung and bronchus - cancer stat facts. SEER. https://seer.cancer.gov/statfacts/html/lungb.html, 2021 (accessed 4 January 2022).

[5]. Burns J, Boogaard H, Polus S, Pfadenhauer LM, Rohwer AC, van Erp AM, Turley R, Rehfuess EA, Interventions to reduce ambient air pollution and their effects on health: An abridged Cochrane Systematic Review, Environ. Int 135 (2020), 105400, 10.1016/j.envint.2019.105400. [PubMed: 31855800]

[6]. Li R, Zhou R, Zhang J, Function of PM2.5 in the pathogenesis of lung cancer and chronic airway inflammatory diseases (Review), Oncology Letters 15 (2018) 7506–7514, 10.3892/ol.2018.8355. [PubMed: 29725457]

[7]. Consonni D, Carugno M, De Matteis S, Nordio F, Randi G, Bazzano M, Caporaso NE, Tucker MA, Bertazzi PA, Pesatori AC, Lubin JH, Landi MT, Outdoor particulate matter (PM10) exposure and lung cancer risk in the eagle study, PLoS ONE 13 (9) (2018), 10.1371/journal.pone.0203539.

[8]. Kamis A, Cao R, He Y, Tian Y, Wu C, Predicting lung cancer in the United States: A multiple model examination of public health factors, Int. J. Environ. Res. Public Health 18 (11) (2021) 6127, 10.3390/ijerph18116127. [PubMed: 34204140]

[9]. Loomis D, Grosse Y, Lauby-Secretan B, Ghissassi FE, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K, The carcinogenicity of outdoor air pollution, Lancet Oncol. 14 (13) (2013) 1262–1263, 10.1016/s1470-2045(13)70487-x. [PubMed: 25035875]

[10]. World Health Organization. (2013, October 17). Outdoor air pollution a leading environmental cause of cancer deaths. World Health Organization. https://www.euro.who.int/en/health-topics/environment-and-health/air-quality/news/news/2013/10/outdoor-air-pollution-a-leading-environmental-cause-of-cancer-deaths, 2013 (accessed 4 January 2022).

[11]. Hamra GB, Laden F, Cohen AJ, Raaschou-Nielsen O, Brauer M, Loomis D. Lung Cancer and Exposure to Nitrogen Dioxide and Traffic: A Systematic Review and Meta-Analysis. Environ Health Perspect. 2015 Nov;123(11):1107–12. doi: 10.1289/ehp.1408882. Epub 2015 Apr 14. [PubMed: 25870974]

[12]. Zhu Y, McKeon TP, Tam V, Vachani A, Penning TM, Hwang WT, Geographic Differences in Lung Cancer Incidence: A Study of a Major Metropolitan Area within Southeastern Pennsylvania, Int. J. Environ. Res. Public Health 17 (24) (2020) 9498, 10.3390/ijerph17249498. [PubMed: 33352953]

[13]. Kazemiparkouhi F, Eum KD, Wang B, et al. , Long-term ozone exposures and cause-specific mortality in a US Medicare cohort, J. Exp. Sci. Environ. Epidemiol 30 (2020) 650–658, 10.1038/s41370-019-0135-4.

[14]. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans (2016). Outdoor Air Pollution. IARC monographs on the evaluation of carcinogenic risks to humans, 109, 9–444.

[15]. State of the air 2021 - American lung association. https://www.lung.org/getmedia/17c6cb6c-8a38-42a7-a3b0-6744011da370/sota-2021.pdf, 2021 (accessed 4 Janaury 2022).

[16]. Tseng C-H, Tsuang B-J, Chiang C-J, Ku K-C, Tseng J-S, Yang T-Y, Hsu K-H, Chen K-C, Yu S-L, Lee W-C, Liu T-W, Chan C-C, Chang G-C, The relationship between air pollution and lung cancer in nonsmokers in Taiwan, J. Thorac. Oncol 14 (5) (2019) 784–792, 10.1016/j.jtho.2018.12.033. [PubMed: 30664991]

[17]. Myers R, Brauer M, Durnmer T, Atkar-Khattra S, Yee J, Melosky B, Ho C, McGuire AL, Sun S, Grant K, Lee A, Lee M, Yuchi W, Tammemagi M, & Lam S (2021). High Ambient Air Pollution Exposure Among Never Smokers Versus Ever Smokers with Lung Cancer. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer, S1556-0864(21)02256–5. Advance online publication. Doi: 10.1016/j.jtho.2021.06.015.

[18]. Rubio V, Valverde M, Rojas E, Effects of atmospheric pollutants on the Nrf2 survival pathway, Environ. Sci. Pollut. Res 17 (2010)369–382, 10.1007/s11356-009-0140-6.

[19]. Penning TM, Jonnalagadda S, Trippier PC, Rižner TL, Aldo-Keto reductases and cancer drug resistance, Pharmacol. Rev 73 (3) (2021) 1150–1171, 10.1124/pharmrev.120.000122. [PubMed: 34312303]

[20]. Zang H, Mathew RO, Cui T, The Dark Side of Nrf2 in the Heart, Front. Physiol 11 (2020) 722, 10.3389/fphys.2020.00722. [PubMed: 32733266]

[21]. Eckel SP, Cockburn M, Shu YH, Deng H, Lurmann FW, Liu L, Gilliland FD, Air pollution affects lung cancer survival, Thorax 71 (10) (2016) 891–898, 10.1136/thoraxjnl-2015-207927. [PubMed: 27491839]

[22]. Most Polluted Cities | State of the Air | American Lung Association. https://www.lung.org/research/sota/city-rankings/most-polluted-cities, 2021 (accessed 4 Janaury 2022).

[23]. State of Lung Cancer: Pennsylvania. State Data | Pennsylvania | American Lung Association. https://www.lung.org/research/state-of-lung-cancer/states/pennsylvania , 2020 (accessed 5 January 2022).

[24]. Cancer facts & figures 2019. American Cancer Society. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2019.html, 2019 (accessed 4 Janaury 2022).

[25]. Goldberg DW, Cockburn MG, Improving geocode accuracy with candidate selection criteria, Transactions in GIS 14 (2010) 149–176, 10.1111/j.1467-9671.2010.01211.x.

[26]. McKeon TP, Hwang WT, Ding Z, Tam V, Wileyto P, Glanz K, Penning TM, Environmental exposomics and lung cancer risk assessment in the Philadelphia metropolitan area using ZIP code-level hazard indices, Environ. Sci. Pollut. Res. Int 28 (24) (2021) 31758–31769, 10.1007/s11356-021-12884-z. [PubMed: 33611735]

[27]. Currie J, Walker R, What do economists have to say about the Clean Air Act 50 years after the establishment of the Environmental Protection Agency? J. Econ. Perspect 33 (4) (2019) 3–26, 10.1257/jep.33.4.3.

[28]. Camiña N, McWilliams TL, McKeon TP, et al. , Identification of spatio-temporal clusters of lung cancer cases in Pennsylvania, USA: 2010–2017, BMC Cancer 22 (2022) 555, 10.1186/s12885-022-09652-8. [PubMed: 35581566]

[29]. Kind A, Buckingham WR, Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas, New Engl. J. Med 378 (26) (2018) 2456–2458, 10.1056/NEJMp1802313. [PubMed: 29949490]

[30]. Mora J, Krepline AN, Aldakkak M, Christians KK, George B, Hall WA, Erickson BA, Kulkarni N, Evans DB, Tsai S, Adjuvant therapy rates and overall survival in patients with localized pancreatic cancer from high area deprivation index neighborhoods, Am. J. Surg 222 (1) (2021) 10–17, 10.1016/j.amjsurg.2020.12.001. [PubMed: 33308823]

[31]. "Rural-Urban Commuting Area Codes." USDA ERS - Rural-Urban Commuting Area Codes. https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/, 2020 (accessed 4 January 2022).

[32]. Census Data Educational Attainment. Explore census data. https://data.census.gov/cedsci/table?q=education&tid=ACSST1Y2019.S1501, 2020 (accessed 4 January 2022).

[33]. Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn C, Time-varying covariates and coefficients in Cox regression models, Ann. Transl. Med 6 (7) (2018) 121. 10.21037/atm.2018.02.12. [PubMed: 29955581]

[34]. Penning TM, Aldo-Keto Reductase Regulation by the Nrf2 System: Implications for Stress Response, Chemotherapy Drug Resistance, and Carcinogenesis, Chem. Res. Toxicol 30 (1) (2017) 162–176, 10.1021/acs.chemrestox.6b00319. [PubMed: 27806574]

[35]. Reddy NM, Kleeberger SR, Yamamoto M, Kensler TW, Scollick C, Biswal S, Reddy SP, Genetic dis\ion of the Nrf2-depedent redox signaling regulated transcriptional programs of cell proliferation cytoprotection, Physiol. Genomics 32 (1) (2007) 74–81. [PubMed: 17895394]

[36]. Jiang Y, Niu Y, Xia Y, Liu C, Lin Z, Wang W, Ge Y, Lei X, Wang C, Cai J, Chen R, Kan H, Effects of personal nitrogen dioxide exposure on airway inflammation and lung function, Environ. Res 177 (2019), 108620, 10.1016/j.envres.2019.108620. [PubMed: 31400563]

[37]. Radi R, Cassina A & Hodara R (2002). Nitric Oxide and Peroxynitrite Interactions with Mitochondria, 383(3-4), 401–409. Doi: 10.1515/BC.2002.044.

[38]. Radi R, Oxygen radicals, nitric oxide, and peroxynitrite: Redox pathways in molecular medicine, PNAS 115 (23) (2018) 5839–5848, 10.1073/pnas.1804932115. [PubMed: 29802228]

[39]. Möller MN, Rios N, Trujillo M, Radi R, Denicola A, Alvarez B, Detection and quantification of nitric oxide-derived oxidants in biological systems, J. Biol. Chem 294 (40) (2019) 14776–14802, 10.1074/jbc.REV119.006136. [PubMed: 31409645]

[40]. Atkinson RW, Butland BK, Anderson HR, Maynard RL, Long-term concentrations of nitrogen dioxide and mortality, Epidemiology 29 (4) (2018) 460–472, 10.1097/ede.0000000000000847. [PubMed: 29746370]

[41]. Cox LA Jr. (2012). Hormesis for fine particulate matter (PM 2.5). Dose-response : a publication of International Hormesis Society, 10(2), 209–218. Doi: 10.2203/dose-response.11-040.Cox. [PubMed: 22740783]

[42]. Xu X, Ha S, Kan H, et al. , Health effects of air pollution on length of respiratory cancer survival, BMC Public Health 13 (2013) 800, 10.1186/1471-2458-13-800. [PubMed: 24004483]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
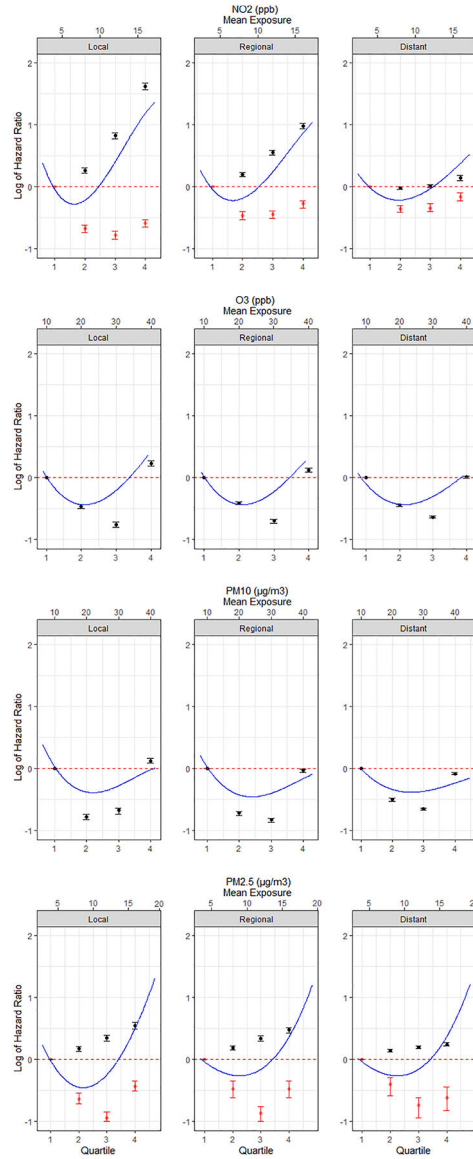
Fig. 1.

Adjusted HRs and 95% confidence intervals for all-cause mortality associated with a 1 unit increase in quartile categorized air pollutant exposures within each SEER cancer stage: Localized—Cancer has shown no sign of spread and is limited to the place where it started; Regional—Cancer has spread to nearby ares of the body; Distant—Cancer has spread to distant parts of the body. HRs of the time-dependent covariates are shown for before (black) and after (red) time splits for NO_2 and $PM_{2.5}$. Polynomial spline curves at 2 degrees of freedom shown for continuous measurements (blue).

Table 1

Demographic, tumor, and characteristics of patients with lung cancer in Pennsylvania by stage of diagnosis.

Characteristics (mean ± SD or %)	Localized (n = 52,792)	Regional (n = 62,952)	Distant (n = 121,888)	Unknown (n = 14,491)	Total (n = 252,123)
Histology at diagnosis, %					
Squamous cell	29.7	30.98	17.86	22.86	23.9
Adenocarcinoma	41.37	33.99	36.52	27.03	36.36
Small cell	7.85	14.72	21.01	13.56	16.25
Large cell	9.94	9.81	12.973	26.98	12.35
Other	11.11	10.47	11.62	9.55	11.11
Treatment Status (2010–2017, n)					
Yes	13,694	15,075	30,373	1,321	60,463
No	1,716	2,293	8,475	535	13,019
Age (years)	69.65 ± 10.25	68.61 ± 10.47	68.34 ± 11.13	72 ± 10.94	68.89 ± 10.81
Male, %	52.05	56.03	56.70	54.88	55.44
Race/ethnicity, %					
Non-Hispanic white	90.02	89.13	88.42	89.73	89.01
Hispanic	0.57	0.63	0.73	0.44	0.65
Non-Hispanic black	8.51	9.36	9.86	8.58	9.38
Other	0.90	0.87	0.98	1.23	0.95
Bachelor's degree, %					
Low (<25)	24.03	24.73	25.52	25.4	25.01
low-medium (26–50)	24.48	24.89	25.18	25.76	24.99
medium–high (51–75)	25.11	25.21	24.88	24.51	24.99
High (76–100)	26.36	25.14	24.39	24.31	24.99
Area Deprivation Index (ADI) (1 to 100)					
Lowest (1–20)	17.7	16.9	16.5	16.9	16.9
Lower-middle (21–40)	15.7	15.6	15.6	16.2	15.7
Middle (41–60)	19.7	19.6	19.9	18.9	19.7
Higher-middle (61–80)	20.5	20.7	21.0	21.1	20.8
Highest (81–100)	23.9	24.7	24.3	24.4	24.3

Characteristics (mean ± SD or %)	Localized (n = 52,792)	Regional (n = 62,952)	Distant (n = 121,888)	Unknown (n = 14,491)	Total (n = 252,123)
Unknown	2.5	2.5	2.7	2.5	2.6
Rural-urban commuting area (RUCA), %					
Non-metropolitan core	8.34	9.08	8.91	8.52	8.81
Metropolitan core	91.65	90.91	91.08	91.47	91.18
Unknown	0	0	0	0	0
Year of Diagnosis, %					
1990–1995	24.95	19.74	15.83	29.41	19.5
1996–2000	18.33	18.42	16.42	26.94	17.92
2001–2005	14.58	19.03	18.94	20.62	18.15
2006–2010	15.94	18.59	20.92	10.92	18.72
2011–2017	26.18	24.19	27.86	12.08	25.68
Geocoding Match Quality, %					
Street address match	89.8	89.3	89.9	86.6	89.5
Area-level match	9.8	10.3	9.7	13.2	10.0
Other or missing	0.4	0.4	0.4	0.2	0.5
Median Survival Time (years)	2.20	1.13	0.42	0.71	0.76
Median Survival Time (years), by histology at diagnosis					
Squamous cell	1.93	1.07	0.44	0.74	0.86
Adenocarcinoma	3.27	1.55	0.47	0.85	0.96
Small cell	1.12	1.01	0.50	0.82	0.67
Large cell	1.08	0.65	0.23	0.44	0.39
Other	2.98	1.13	0.41	0.95	0.76
Air pollution exposures (mean ± SD)					
NO_2 (ppb)	8.14 ± 2.10	8.13 ± 2.15	8.03 ± 2.34	9.06 ± 2.14	8.14 ± 2.24
O_3 (ppb)	22.80 ± 3.28	22.96 ± 3.31	22.92 ± 4.09	23.75 ± 3.60	22.95 ± 3.72
PM_{10} (ug/m^3)	26.50 ± 4.42	26.48 ± 4.61	26.63 ± 5.74	25.89 ± 5.41	26.53 ± 5.20
$PM_{2.5}$ (ug/m^3)	6.61 ± 1.63	6.68 ± 1.66	6.62 ± 1.78	7.31 ± 1.67	6.67 ± 1.72

Table 2

Median overall survival and 5-year overall survival rate, by stage at diagnosis and air pollution exposure.

	Median Survival in years (95% CIs)			Five-year survival rate Localized in percentage (95% CIs)		
	Localized	Regional	Distant	Localized	Regional	Distant
NO$_2$ (ppb)						
Q1 (<6.22)	2.77 (2.67–2.83)	1.49 (1.45–1.54)	0.49 (0.48–0.50)	27.00 (26.00–28.10)	13.30 (12.60–13.90)	2.00 (1.84–2.17)
Q2 (6.22–8.10)	3.31 (3.19–3.46)	1.35 (1.31–1.39)	0.45 (0.44–0.46)	40.60 (39.70–41.60)	19.40 (18.70–20.10)	3.34 (3.11–3.58)
Q3 (8.1–9.93)	2.60 (2.50–2.70)	1.05 (1.02–1.08)	0.43 (0.42–0.44)	33.80 (32.90–34.70)	12.10 (11.60–12.70)	1.64 (1.48–1.80)
Q4 (>9.93)	1.27 (1.24–1.31)	0.87 (0.85–0.89)	0.33 (0.32–0.33)	13.50 (12.90–14.10)	5.10 (4.80–5.50)	0.68 (0.59–0.78)
O$_3$ (ppb)						
Q1 (<20.63)	1.76 (1.67–1.83)	0.98 (0.95–1.02)	0.31 (0.30–0.32)	19.30 (18.20–20.50)	9.43 (8.76–10.2)	1.58 (1.41–1.78)
Q2 (20.63–23.23)	2.21 (2.14–2.28)	1.24 (1.21–1.27)	0.63 (0.62–0.64)	28.40 (27.60–29.30)	13.50 (13.00–14.10)	2.37 (2.19–2.57)
Q3 (23.23–25.17)	3.80 (3.70–3.89)	1.66 (1.63–1.703)	0.67 (0.66–0.69)	41.70 (41.00–42.50)	19.70 (19.10–20.30)	3.67 (3.44–3.92)
Q4 (>25.17)	1.02 (0.98–1.06)	0.68 (0.66–0.70)	0.25 (0.25–0.26)	8.87 (8.25–9.53)	2.83 (2.55–3.13)	0.32 (0.26–0.39)
PM$_{10}$ (µg/m3)						
Q1 (<23.37)	1.11 (1.07–1.15)	0.64 (0.62–0.66)	0.24 (0.23–0.25)	10.90 (10.10–11.70)	3.80 (3.43–4.21)	0.53 (0.43–0.64)
Q2 (23.37–26.40)	3.23 (3.13–3.35)	1.30 (1.28–1.34)	0.60 (0.58–0.60)	38.50 (37.70–39.30)	15.70 (15.10–16.30)	2.44 (2.25–2.65)
Q3 (26.40–29.19)	3.35 (3.27–3.45)	1.90 (1.86–1.94)	0.72 (0.70–0.73)	36.90 (36.10–37.70)	20.50 (19.90–21.20)	4.38 (4.12–4.65)
Q4 (>29.19)	1.14 (1.10–1.18)	0.81 (0.79–0.84)	0.33 (0.32–0.33)	8.40 (7.73–9.13)	3.89 (3.53–4.27)	0.54 (0.46–0.63)
PM$_{2.5}$ (µg/m3)						
Q1 (<5.20)	2.43 (2.35–2.51)	1.45 (1.42–1.50)	0.58 (0.56–0.59)	25.90 (24.8–27.0)	13.50 (12.8–14.2)	2.21 (2.00–2.44)
Q2 (5.20–6.47)	2.38 (2.29–2.48)	1.20 (1.17–1.24)	0.45 (0.44–0.46)	31.30 (30.40–32.20)	15.10 (14.50–15.70)	2.44 (2.26–2.63)
Q3 (6.47–7.96)	2.32 (2.23–2.40)	1.09 (1.07–1.11)	0.35 (0.34–0.35)	31.20 (30.40–32.10)	11.90 (11.40–12.40)	1.59 (1.45–1.74)
Q4 (>7.69)	1.74 (1.68–1.80)	0.91 (0.89–0.93)	0.34 (0.34–0.35)	23.30 (22.60–24.10)	8.55 (8.12–9.01)	1.07 (0.95–1.20)

Table 3

Adjusted overall hazard ratios.

Model adjusted with age; sex; race/ethnicity (Hispanic, Non-Hispanic-Black, Non-Hispanic White, or other); categorized percent of those 25 years of age or greater with Bachelor's ("low", "low-medium", "medium-high", and "high"); categorized Area Deprivation Index ("Lowest", "Lower-middle", "Middle", "Higher-middle", "Highest", and "Unknown"); categorized RUCA codes (Urban, Metropolitan, Other); categorized histology codes (Squamous cell carcinoma, adenocarcinoma, small cell carcinoma, large cell carcinoma and other carcinomas); categorized year of diagnosis (1990–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2017); month of diagnosis; and distance from residential address to a primary interstate highway, categorized into (<300 m, 300–1500 m, > 1500 m). NO_2 adjusted with 2-year post-diagnosis timesplit covariate and $PM_{2.5}$ adjusted with a 5-year post-diagnosis timesplit covariate.

Overall Pollutant	All Cause Death HRs (95%CIs)	Person-years	Number of deaths	Lung Cancer Specific HRs (95%CIs)	Person-years	Number of deaths
NO_2 Q1	(ref)	84,572	51,515	(ref)	46,083	39,989
NO_2 Q2	<2y 1.06 (1.04–1.07)>2y 0.60 (0.58–0.62)	118,849	52,091	<2y 1.00 (0.98–1.02)>2y 0.64 (0.61–0.66)	53,789	38,884
NO_2 Q3	<2y 1.35 (1.33–1.37)>2y 0.58 (0.56–0.60)	112,194	57,155	<2y 1.23 (1.21–1.25)>2y 0.64 (0.61–0.66)	59,257	43,942
NO_2 Q4	<2y 1.89 (1.86–1.93)>2y 0.78 (0.76–0.80)	74,487	61,144	<2y 1.64 (1.61–1.68)>2y 0.83 (0.80–0.86)	49,584	49,120
NO_2 Continuous	1.17 (1.16–1.17)			1.14 (1.14–1.14)		
O_3 Q1	(ref)	54,065	47,988	(ref)	25,642	36,995
O_3 Q2	0.63 (0.63–0.64)	111,682	55,239	0.64 (0.63–0.65)	57,906	42,852
O_3 Q3	0.47 (0.46–0.48)	174,740	59,477	0.45 (0.45–0.46)	89,948	44,065
O_3 Q4	1.12 (1.11–1.14)	49,615	59,201	1.09 (1.07–1.11)	35,217	48,023
O_3 Continuous	1.04 (1.04–1.04)			1.04 (1.04–1.04)		
PM_{10} Q1	(ref)	44,585	53,674	(ref)	27,555	42,472
PM_{10} Q2	0.47 (0.46–0.47)	141,076	56,653	0.47 (0.46–0.48)	71,709	42,462
PM_{10} Q3	0.43 (0.42–0.44)	154,003	55,871	0.44 (0.43–0.44)	75,297	41,595
PM_{10} Q4	0.94 (0.92–0.95)	50,437	55,707	0.97 (0.96–0.99)	34,151	45,406
PM_{10} Continuous	1.01 (1.01– 1.01)			1.02 (1.01–1.02)		
$PM_{2.5}$ Q1	(ref)	75,396	44,907	(ref)	35,574	33,409
$PM_{2.5}$ Q2	<5y 1.22 (1.21–1.24)>5y 0.56 (0.53–0.59)	101,657	55,185	<5y 1.25 (1.24–1.28)>5y 0.66 (0.61–0.71)	53,178	43,057
$PM_{2.5}$ Q3	<5y 1.41 (1.39–1.43)>5y 0.35 (0.33–0.37)	121,374	60,263	<5y 1.39 (1.36–1.41)>5y 0.57 (0.52–0.61)	63,775	46,934
$PM_{2.5}$ Q4	<5y 1.59 (1.56–1.62)>5y 0.67 (0.64–0.71)	91,675	61,550	<5y 1.54 (1.50–1.57)>5y 0.38 (0.35–0.42)	56,186	48,535

Overall Pollutant	All Cause Death HRs (95%CIs)	Person-years	Number of deaths	Lung Cancer Specific HRs (95%CIs)	Person-years	Number of deaths
$PM_{2.5}$ Continuous	1.15 (1.15– 1.16)			1.14 (1.13– 1.14)		