Pollution and regional variations of lung cancer mortality in the United States

Justin Xavier Moore, MPH[1,2,3], **Tomi Akinyemiju, PhD**[1,2], and **Henry E. Wang, MD, MS**[3]

[1]Departments of Epidemiology, University of Alabama at Birmingham, Birmingham AL, USA

[2]Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, USA

[3]Department of Emergency Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

Abstract

Introduction—The aims of this study were to identify counties in the United States (US) with high rates of lung cancer mortality, and to characterize the associated community-level factors while focusing on particulate-matter pollution.

Methods—We performed a descriptive analysis of lung cancer deaths in the US from 2004 through 2014. We categorized counties as "clustered" or "non-clustered" – based on whether or not they had high lung cancer mortality rates – using novel geospatial autocorrelation methods. We contrasted community characteristics between cluster categories. We performed logistic regression for the association between cluster category and particulate-matter pollution.

Results—Among 362 counties (11.6%) categorized as clustered, the age-adjusted lung cancer mortality rate was 99.70 deaths per 100,000 (95%CI: 99.1–100.3). Compared with non-clustered counties, clustered counties were more likely in the south (72.9% versus 42.1%, $P < 0.01$) and in non-urban communities (73.2% versus 57.4, $P < 0.01$). Clustered counties had greater particulate-matter pollution, lower education and income, higher rates of obesity and physical inactivity, less access to healthcare, and greater unemployment rates ($P < 0.01$). Higher levels of particulate-matter pollution (4th quartile versus 1st quartile) were associated with two-fold greater odds of being a clustered county (adjusted OR: 2.10; 95%CI: 1.23–3.59).

Conclusion—We observed a belt of counties with high lung mortality ranging from eastern Oklahoma through central Appalachia; these counties were characterized by higher pollution, a more rural population, lower socioeconomic status and poorer access to healthcare. To mitigate the

Corresponding author: Justin Xavier Moore, MPH, Department of Epidemiology, University of Alabama at Birmingham, RPHB 220 1665 University Boulevard, Birmingham, AL 35249, 404-906-8446 (phone), 205-975-4662 (fax), jxmoore@uab.edu.

Conflict of interest statement

The authors declare no conflicts of interest.

burden of lung cancer mortality in the US, both urban and rural areas should consider minimizing air pollution.

Keywords

neoplasms; public health; socioeconomic factors

Introduction

In the United States (US) cancer is the second leading cause of death, being responsible for more than 550,000 deaths annually [1–3]. Among specific cancers, lung cancer is the leading cause of cancer deaths for both men and women [3, 4]. Ambient particulate-matter pollution was recently classified as a group 1 carcinogen in human populations [5–10]. Outdoor air pollution includes a combination of various pollutants derived from natural and anthropogenic sources [5–7]. More specifically, particulate matter comprises fine particles <2.5 μg/m^3 (PM2.5) present in ambient air; these are the result of combustion processes derived from sources such as exhaust from motor vehicles, power generators, industrial activity, biomass burning, and domestic heating and cooking [5–7].

To date, few studies have examined regional trends in lung cancer mortality and/or prevalence using geospatial epidemiologic methodology [11–14]. Underwood et al (2011) found that there are regional differences in lung cancer incidence, with the southern census region having the highest incidence rate for lung cancer corresponding to 74.7 cases per 100,000 persons [14]. While prior studies provide information regarding the region affected, there remains a need for a more granular investigation of the geospatial trends in lung cancer, in addition to the assessment of community characteristics associated with these regions. Therefore, it is imperative that we examine and target the geographic areas within the US that are suffering from high lung cancer mortality rates, and deduce what factors drive this increased burden of lung cancer mortality and, specifically, whether exposure to particulate matter is higher in these areas.

Finding geospatial patterns of lung cancer deaths may provide further insight into the underlying socioeconomic and community factors associated with the disease. Because of the lack of literature regarding regional disparities in lung cancer mortality within the US, there are many unanswered questions about the community-level factors that may explain lung cancer mortality clusters. In this study we aimed to define US counties with high lung cancer mortality, and to assess the county-level characteristics associated with lung cancer mortality – focusing specifically on exposure to particulate matter.

Methods

Study design

We performed an ecologic study of deaths from lung cancer in the United States between 2004 and 2014. We complemented county-level lung cancer mortality data from the CDC with county-level data obtained from the CDC, the 2014 County Health Rankings (CHR) and the American Community Survey (ACS)[15, 16]. We performed analysis at the county

level because of our focus on community-level factors (state-level analyses are not too granular) and of the most available granular data within the US (census tract and block group data were not readily available for variables of study).

County-level particulate matter exposure

Our primary exposure was county-level mean particulate matter <2.5 μg/m^3 (PM2.5). We obtained data on PM2.5 for years 2003 through 2011 from the US Environmental Protection Agency (EPA) and the National Aeronautics and Space Administration (NASA) by assessing the CDC "wide-ranging online data for epidemiologic research" (Wonder) system [17, 18]. Al-Hamdan et al. (2009) generated continuous spatial surfaces of daily PM2.5 for the entire contiguous US for 2003 through 2011 using environmental data from the US EPA and NASA [17]. The daily PM2.5 values were estimated using satellite measurements of aerosol optical depth (AOD) from the moderate-resolution-imaging spectroradiometer (MODIS) instrument on the NASA Terra and Aqua satellites [19]. Using the daily PM2.5 values for the contiguous US counties derived by Al-Hamdan et al (2009), we estimated the mean daily PM2.5 exposure for the time period for each county. For statistical analysis, we categorized PM2.5 into quartiles based on distribution at the county level.

County-level demographic characteristics

Demographic variables included the county-level proportions for age, sex, race, household income, population with college education, and rural/urban commuting area (RUCA) codes. We categorized race into five groups: White (non-Hispanic), Black (non-Hispanic), Hispanic, Asian/Pacific Islander (non-Hispanic), and Native American. We classified counties as urban or non-urban using the 2010 RUCA classifications. RUCA codes are based on population size and density, urbanization, and primary and secondary worker commuter flow patterns to larger urban areas and clusters [20]. The 11 RUCA codes were then aggregated into a dichotomized variable: (1) urban (i.e., population centers with 50,000 residents) and (2) non-urban (i.e., towns or small cities with population centers of <50,000 residents) [18, 20]. We obtained county-level demographic statistics from the 2014 ACS via the National Historical Geographic Information System (NHGIS) [15]. The 2014 ACS provided an aggregated estimate of demographic statistics for the preceding 5-year period (2010–2014).

County-level community characteristics

Community variables included the county-level proportions of adults who smoked, who were obese, physically inactive, unemployed or uninsured, those who could not see a doctor because of cost, and the total number of primary-care physicians. We calculated the ratio of primary-care physicians per 10,000 persons by dividing the total number of primary-care physicians by the county population and then multiplying the quotient by 10,000. Geographic region was defined using the census definition (i.e., Midwest, Northeast, South, and West) [21]. This analysis focused on the contiguous US, and therefore we did not include counties in Alaska, Hawaii, and other US territories in this analysis. We obtained county-level community characteristics from the 2014 county health rankings (CHR) [15, 16].

Identification of lung cancer deaths

We defined lung cancer deaths as deaths associated with cancers of the bronchus and lung for the years 2004 through 2014. Lung and bronchus cancer deaths were taken from the compressed mortality file (CMF) produced by the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC) [22]. The CMF data are derived from death certificates and include a record for every death of a US resident recorded in the United States. We identified the county-level total number of deaths related to lung cancer using the following ICD-10 codes: C34 (malignant neoplasm of bronchus and lung), C45.0 (mesothelioma of pleura), C45.9 (mesothelioma, unspecified/malignant neoplasms), and C78.0 (secondary malignant neoplasm of lung/malignant neoplasms), and C78.2 (secondary malignant neoplasm of pleura/malignant neoplasms). We included lung-cancer-related deaths for all persons of aged 15 years during our observation period. We included deaths for persons aged 15–19 because the CMF uses a single reference standard population for ages 15–24, and therefore the inclusion of the 15–19-year age groups is necessary for the estimation of age-adjusted rates.

Defining lung mortality clusters

To date, there is no formal approach for identifying spatial disease clustering [23]. However, disease clustering has been defined as the spatial aggregation of cases in an identifiable subpopulation [24]. Therefore we adopted novel geospatial autocorrelation methods to define counties with highly clustered lung cancer mortality. We combined three separate spatial clustering methods to identify areas that were hotspots for lung cancer mortality [25]. We categorized county-level lung cancer mortality clustering into two groups: "clustered" or "non-clustered". We considered a county to be clustered if it was identified as high-risk or a lung cancer mortality hotspot using two or more of the geospatial analyses: fifth quintile of smoothed empirical Bayes (EB) lung cancer mortality rates, high-high cluster using local indicators of spatial association (LISA), or a lung cancer mortality hotspot as defined by the Getis-Ord Gi* statistic [26, 27]. All other US counties were categorized as non-clustered.

First, we estimated the smoothed EB lung cancer mortality rates using the total number of lung cancer mortality events in each county divided by the county population, with smoothing performed using the EB tools in GeoDa 1.6.7.9 (http://geodacenter.asu.edu)[25]. We further categorized the EB smoothed lung cancer mortality rates into quintiles, and we defined counties as high-risk if the smoothed EB lung cancer mortality rates were in the fifth quintile. Second, we used LISA to measure similarity between counties and calculate values both within and across geographic boundaries, additionally identifying spatial outliers [25, 26, 28]. For each US county we estimated local Moran's I statistic values, using associated z-scores and P-values to assess the magnitude of spatial autocorrelation and statistical significance, respectively [25]. Counties with statistically significant positive z-scores indicate areas surrounded by areas with similar lung cancer mortality rates – either similarly high or similarly low (positive spatial autocorrelation)[25]. Lastly, we used the Gi* statistic to identify areas where lung cancer mortality rates with either high or low values clustered within the context of the neighboring county [25, 27, 29, 30]. In contrast to LISA, the Gi* statistic is not related to the global statistic of spatial association [26]. Moreover, while the LISA statistic includes a diagnostic for outliers with respect to a measure of global

association, the Gi* only examines associations for counties that share borders [26]. Further, the LISA analysis allows for examining local associations while accounting for the global spatial association (that can be observed from the counties identified as either high-low or low-high outliers). Further details of the geospatial autocorrelation analysis were introduced by Nassel et al. (2014) and are described elsewhere [25].

Statistical analysis

We used the mean crude and age-adjusted mortality rates provided by the CMF, which uses the US census population estimates. All standardization was performed relative to the 2000 US standard population. We performed multivariable Poisson regression to estimate both the crude and age-adjusted lung cancer mortality rates for each lung-cancer-clustering category (i.e., clustered or non-clustered). We additionally calculated the crude and age-adjusted lung cancer mortality rates for the entire contiguous US. We contrasted regional (i.e., Midwest versus Northeast versus South versus West) differences in high-risk lung cancer categories (EB, LISA, Gi*, and lung cancer mortality cluster category) using the chi-square test of association. We contrasted differences between clustered and non-clustered counties for demographic and community characteristics using the Wilcoxon test. We additionally examined the correlation of demographic and community characteristics with being a lung cancer mortality cluster county using a Spearmen correlation (ρ).

We performed logistic regression to determine the association between clustered counties and county-level daily mean PM2.5, with clustered as the outcome and categorized PM2.5 as the exposure. We performed models adjusted for: (1) demographics which included age, race, and sex; (2) demographics plus socioeconomic status which included age, race, sex, household income, percentage completed college, proportion of adults smoking, proportion of adults obese, and proportion of adults physically inactive; and (3) demographics plus access to healthcare which included age, race, sex, proportion of adults uninsured, proportion of adults that could not see a doctor because of cost within the past year, and the primary-care physician ratio. We examined whether urban county residence was an effect modifier by performing our logistic regression models stratified by urban and non-urban status. We used ArcGis version 10.4 to produce all map images. We used SAS version 9.4, and GeoDa version 1.6.7.9 for all statistical analyses.

Ethical statement

The Institutional Review Board (IRB) of the University of Alabama at Birmingham deemed this study exempt from IRB review because we used publicly available, unidentifiable, and secondary data sources

Results

LISA, EB, and GI* analyses

Using the LISA analysis, 328 of the 446 (73.54%), $P<0.01$) "high-high" lung cancer mortality cluster counties were located in the southern United States (Table 1, Supplemental Figure 1). Overall, when using the smoothed EB analysis, 622 counties (20.01%) were categorized within the highest quintile for lung cancer mortality. Southern counties were

more likely to have EB rates within the fifth quintile (66.08%, *P* < 0.01) when compared with other US census regions (Table 1, Supplemental Figure 2). The Gi* analysis identified 226 counties as "lung cancer hotspots". Counties in the northeastern counties (30.97%) and southern counties (28.76%) were more likely to be identified as a lung cancer hotspot by Gi* criteria (Table 1, Supplemental Figure 3).

Demographic characteristics

Among a total of 3,108 contiguous US counties included in this study, 362 (11.65%) were classified as clustered counties and 2,746 (88.35%) were classified as non-clustered counties (Figure 1, Table 2). Clustered counties were more likely to have a higher proportion of NH-White residents when compared with non-clustered counties (93.95% versus 84.10%, *P* <0.01, ρ = 0.23). Clustered counties were more likely to have residents aged 45 when compared with non-clustered counties.

Community characteristics

Clustered counties had higher levels of daily mean PM2.5 when compared with non-clustered counties (Figure 2; 13.61 versus 11.70 μg/m^3, *P* < 0.01, ρ = 0.18). Clustered counties had a higher proportion of residents with household income below $20,000 (26.91% versus 19.82, *P* < 0.01, ρ = 0.34) and a lower proportion of adults without a college education (12.87% versus 18.59%, *P* < 0.01, ρ = −0.32) when compared with non-clustered counties. Clustered counties had a larger proportion of adults currently smoking tobacco (25.50% versus 18.70%, *P* < 0.01, ρ = 0.30), obese adults (32.90% versus 30.30%, *P* < 0.01, ρ = 0.21), adults that are physically inactive (33.10% versus 27.70%, *P* < 0.01, ρ = 0.31), and unemployed adults (9.00% versus 7.30%, *P* < 0.01, ρ = 0.24) when compared with non-clustered counties. Clustered counties also had a larger proportion of adults without insurance (23.10% versus 21.05%, *P* < 0.01, ρ = 0.11) and of adults that could not see a doctor because of cost (18.35% versus 12.00%, *P* < 0.01, ρ = 0.23) when compared with non-clustered counties. In addition, clustered counties had a lower number of PCPs per 10,000 persons (4.18 versus 4.99, *P* < 0.01, ρ = −0.09) when compared with non-clustered counties. Clustered counties were more likely to be in non-urban communities when compared with non-clustered counties (73.20% versus 57.39%, *P* < 0.01, ρ = 0.10).

Lung cancer mortality rates

During the observation period (2004–2014), and among persons aged 15 years and older, there was a total of 1,754,907 lung-cancer-related deaths, corresponding to a national crude mortality rate of 65.50 deaths per 100,000 persons (95%CI, 65.40–65.60; Table 3). Among the contiguous US counties, the age-adjusted lung cancer mortality rate was 78.20 (95%CI, 78.00–78.30). Over the observation period, the crude lung cancer mortality rate was 1.85 times higher for the clustered counties (116.80 deaths per 100,000 persons; 95%CI, 116.20–117.40) when compared with the non-clustered counties (63.00 deaths per 100,000 persons; 95%CI, 62.90–63.10). In addition, clustered counties (99.70 deaths per 100,000 persons) had higher age-adjusted lung cancer mortality rates when compared with non-clustered counties (75.60 deaths per 100,000 persons).

Odds of lung cancer mortality clustering

When examining the odds of being considered a clustered lung cancer mortality county, counties with PM2.5 values within the highest quartile (4th quartile: 13.56–14.95 μg/m^3) were more than five times more likely to be a clustered county before adjustment for confounders (crude OR: 5.41, 95%CI: 3.85–7.58; Table 4). After adjustment for sociodemographic characteristics (age, race, and sex), highest PM2.5 exposure was associated with greater odds of being a clustered county (model 1 OR: 3.48, 95%CI: 2.15–5.66). In a model additionally adjusted for sociodemographics and socioeconomic characteristics (i.e., household income, college education, adult smoking, obesity, and physical inactivity) the association between the highest PM2.5 exposure and clustered status attenuated (model 2 OR: 1.59, 95%CI: 0.88–2.88). However, in a model adjusted for sociodemographics and access to healthcare characteristics (i.e., insurance, could not see doctor due to cost, and PCP) greater PM2.5 exposure was associated with increased odds of clustering (model 3 OR: 2.10, 95%CI: 1.23–3.59).

Odds of lung cancer mortality clustering by urban status

Among non-urban counties, a total of 265 counties were categorized as clustered. In the unadjusted analysis, higher levels of PM2.5 were associated with increased odds of lung cancer mortality clustering for the 2nd (crude OR: 1.85, 95%CI: 1.11–3.07), 3rd (crude OR: 2.22, 95%CI: 1.35–3.65), and 4th quartiles (crude OR: 7.95, 95%CI: 5.07–12.45) of PM2.5. After adjustment for sociodemographic characteristics, counties with higher levels (within the fourth quartile) of PM2.5 were more likely to be clustered (model 1 OR: 4.06, 95%CI: 2.23–7.39). However, after adjustment of both sociodemographics and socioeconomic characteristics the association between PM2.5 and clustering attenuated (model 2 OR: 1.86, 95%CI: 0.90–3.87). When adjusted for sociodemographics and access to healthcare characteristics, higher levels (3rd and 4th quartiles, respectively) of PM2.5 were associated with increased odds of being clustered (3rd quartile model 3 OR: 2.35, 95%CI: 1.27–4.34 and 4th quartile model 3 OR: 3.19, 95%CI: 1.70–5.96). When limited to urban counties only, there was a total of 97 counties categorized as clustered. In unadjusted analysis, counties with PM2.5 within the 4th quartile were at increased odds of lung cancer mortality clustering (crude OR: 3.26, 95%CI: 1.77–5.99) when compared to counties within the 1st quartile of PM2.5. However, the association between PM2.5 and lung cancer mortality clustering attenuated and became non-significant after adjustment for confounders.

Discussion

In this study we sought to examine county-level regional variation in lung cancer mortality, associated community characteristics, and the modifiable effects of fine particulate matter. Lung cancer mortality remains a huge public health burden in the US population, responsible for more than 1.7 million deaths in the period 2004 through 2014. During the observation period, 11.6% of the contiguous US counties (362/3,108) were categorized as clustered counties, defined as counties with high levels of lung cancer mortality using novel geospatial methods. PM2.5 levels were higher in clustered counties compared with non-clustered counties. In addition, the age-adjusted lung mortality rate was 1.3 times higher for persons living in clustered counties when compared with citizens living in the non-clustered

counties. We observed that there was a specific cluster of high lung mortality counties ranging from eastern Oklahoma through central Appalachia. This cluster of high lung mortality counties formed a definitive "lung cancer mortality belt." While simultaneously avoiding all major cities, the lung cancer mortality belt included counties within the states of Oklahoma, Arkansas, southern Missouri, southern Illinois, Kentucky, Tennessee, southern Ohio, West Virginia, and western Virginia. Non-urban communities with lower levels of education and income, higher rates of obesity and physical inactivity, poorer access to healthcare, and greater unemployment rates characterized clustered counties, and likewise the lung cancer mortality belt.

To our knowledge, this is the first study to employ multiple geospatial analyses (LISA, EB rates, Gi*) to define an aggregated cluster of counties suffering from high lung cancer mortality rates. In addition, we identified that counties with higher lung cancer mortality rates are in areas with greater community-level disadvantages such as greater poverty, poorer education, higher unemployment, and higher rates of uninsured adults. Research has consistently shown that greater access to healthcare and higher socioeconomic status is associated with lower risk of cancer mortality [31–34]. Tannenbaum et al (2014) found that individuals living in communities with the highest SES had a 13% reduced hazard (HR: 0.87, 95%CI: 0.84–0.91) for lung cancer mortality compared to individuals living in impoverished communities. In addition, Torre et al (2016) found that lung cancer mortality is higher among people with lower SES and in the mid-south region (containing the states of Kentucky, Mississippi, Arkansas, and Tennessee) similar to our lung cancer mortality belt [34]. However, we found that even after adjustment for sociodemographics and access to healthcare, counties with higher PM2.5 levels are associated with a >2-fold increased odds of being in a lung cancer mortality cluster. We also note that, historically, regional disparities in other health outcomes have been found in the geographic area observed in this study [35–46]. Similarly to the counties located in the western region of the lung cancer mortality belt (near the Mississippi river), other studies have shown that the Mississippi valley has higher rates of coronary heart disease [47]. Likewise, Moore et al (2016) found that there were sepsis mortality clusters in the Mississippi valley and Central Appalachian regions similar to the regions affected by high lung cancer mortality in this study [46]. These results further illuminate the fact that the current lung cancer mortality belt is a region within the United States that is representative of poorer health outcomes due to increased poverty risk and poorer access to health. Areas of greater unemployment and lower SES are subsequently victim to greater health disparities that alternatively perpetuate cultural patterns that further reduce access to healthcare and opportunities for sustainable disease prevention [48]. As a result, it is ever more important that US policy focuses on increasing employment opportunities and affordable healthcare for individuals throughout the US, specifically focusing on rural populations [49, 50].

Interestingly, clustered counties were more likely to be in non-urban areas and to have higher levels of environmental fine particulate matter. Historically, US urban areas have been the epicenter for land development, mass motor vehicle congestion, and industrial facilities, and therefore the majority of prior research investigating the relationship between environmental fine particulate pollution and health risks have been among urban populations [51–53]. However, consistent with the results of our study, studies suggest that rural

environments are potentially exposed to a variety of pollutants derived from industrial facilities, animal containment facilities, mining operations, and agricultural activities [52]. Moreover, Hendryx et al (2010) examined the distribution of pollution sources and mortality risk across rural and non-rural areas in the US and found that a greater density of pollution sources (i.e., number of fossil fuel plants and anthropogenic water pollution) was associated with higher adjusted cancer mortality in both rural and non-rural regions [52]. We found that there was no significant difference between the total number of active mining facilities between clustered and non-clustered communities (data not shown). However, it is plausible that other point pollution sources such as land use development and agricultural facilities account for a large proportion of fine particulate matter in rural areas. Furthermore, our results suggest that the improvement of air quality should not focus exclusively on large metropolitan areas, but should also focus on non-urban communities.

There is a biologically plausible relationship between PM2.5 and lung cancer. PM2.5 includes inhalable particles that are small enough to penetrate the thoracic region of the respiratory system, including the airway and alveolar surfaces [54, 55]. Additionally, research has shown that exposure to fine particulate matter can induce inflammatory cytokines and alveolar epithelium–fibroblast interaction essential in the pathogenesis of chronic lung disease [56, 57]. Studies have shown that there are numerous negative health effects of chronic PM2.5 exposure, including an increased risk of physical inactivity [54], cardiovascular mortality and morbidity [58–60], asthma [61], community-acquired pneumonia [62], lung cancer incidence [5, 63], and all-cause mortality [54, 55]. Moreover, Pope et al (2002) observed that for every 10-μm increase in chronic exposure to fine particulate matter, lung cancer mortality increased by 6% [64]. Likewise, a prospective analysis performed in Europe found that a 10-μm increase in fine particulate matter was associated with a 22% increased hazard (HR: 1.22; 95%CI: 1.03–1.45) for lung cancer [10]. Our results further imply that fine particulate matter is associated with higher lung cancer mortality as indicated by the nearly two-fold increased odds of lung cancer mortality clustering among counties with PM2.5 levels within the highest quartile.

We additionally observed that after adjustment for adult smoking and socioeconomic factors (i.e., household income, % completed college, adult obesity, and physical inactivity) the association between PM2.5 and lung cancer mortality clustering attenuated. Therefore, high lung cancer mortality in these clustered areas are in part explained by common contributing factors to lung cancer, mainly smoking prevalence. Prior studies have noted that smoking is an effect modifier on the association between pollution and lung cancer [65, 66]. For example, in a large prospective cohort of 1.2 million participants, Turner et al (2014) found that approximately 14% of lung cancer mortality among cigarette smokers with high PM2.5 may be attributable to interaction of smoking and PM2.5: attributable proportion (AP) = 0.14, 95%CI = 0.00–0.25) [66]. Thus, it is likely that the effect of smoking prevalence provides a synergistic effect on the association between PM2.5 and lung cancer mortality. Furthermore, while smoking is a pertinent risk factor for lung cancer mortality, reductions in both smoking and air pollution seem warranted due to the deleterious effects of both on lung cancer mortality.

The results of our study should be interpreted bearing in mind certain limitations. First, our results are subject to ecologic fallacy. Due to ecologic study design, we assessed the exposure (PM2.5) and outcome (lung cancer mortality) using data from similar dates: we used data from 2003 through 2011 for exposure and 2004 through 2014 for outcome. Moreover, our objective was not to determine causal inference, but rather to examine possible county-level clustering trends and associated socioeconomic characteristics. The lung cancer deaths in this study are derived from publicly available death records ascertained from the CMF, and thus the estimates are subject to misclassification of the listed causes of death. Additionally, the CMF suppressed county-level data representing fewer than ten events, and as a result we were unable to ascertain estimates if counties had less than ten lung cancer deaths over the total 11-year period. Thus, counties with much less than ten events were far more likely to be in areas of low clustering or non-significance during our analysis on geospatial autocorrelation. We identified lung cancer deaths using the International Classification of Diseases, 10th version, codes for lung and bronchus cancers, and consequently the estimates presented may be underestimates of the true number of lung cancer rates in the United States. We used demographic and community-level characteristics derived from US population survey samples, and therefore the estimates for these variables are subject to participation biases. However, the ACS and BRFSS are nationally representative samples that have been used in studies examining sociodemographic and community characteristics [46, 67]. While examining lung cancer incidence would have been more opportune for assessing the risk associated with pollution exposure, we examined lung cancer mortality due to data availability for the entire US population. However, it should be noted that the latest reports from the Surveillance, Epidemiology and End Results (SEER) program estimate an 18% 5-year survival for cancers of the lung and bronchus in the US for years 2007 through 2013 [68]. Therefore, lung cancer mortality estimates used in our analysis are likely reflective of lung cancer incidence patterns in the US population. Another limitation is that indoor air pollution may not be highly correlated with outdoor pollution levels; thus we were unable to account for variance in pollution exposure among US residents [69]. Lastly, we examined characteristics at the community level, and therefore individual factors associated with lung cancer mortality were not attainable.

There are, however, several strengths in the current study. The results of this study have great generalizability as estimates are derived using data from several nationally representative data sources, including the CDC, ACS, and CHR. This study further generates hypotheses regarding the region affected and the population-level characteristics that define the lung cancer mortality. Furthermore, the association found between the lung cancer mortality belt and increased PM2.5 warrants further investigation among a prospective cohort. We found that there was a significantly higher prevalence of adult smoking in clustered counties, thus indicating that smoking remains a key component in the risk of lung cancer mortality [70]. However, while smoking prevalence remains an important contributor, another cause for concern is the association seen between counties with a greater pollution burden and lung cancer mortality clustering.

Conclusion

The results of this study warrant that implementation of lung cancer preventative resources and care should be considered across rural areas in the United States, specifically focusing on the clustered counties described as the lung cancer mortality belt. Particulate matter is a modifiable risk factor that can be reduced by reduction in chemical waste, combustion byproducts from industrial and agricultural facilities, and emissions from motor vehicles. Thus, to reduce the burden of lung cancer mortality in the United States, both urban and rural areas should consider minimizing air pollution.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Health, United States, 2015: With Special Feature On Racial And Ethnic Health Disparities. Hyattsville, Md: National Center For Health Statistics; 2016.

2. Heron M. Deaths: Leading Causes For 2013. National Vital Statistics Reports: From The Centers For Disease Control And Prevention, National Center For Health Statistics, National Vital Statistics System. 2016; 65:1–95.

3. Siegel, Rl, Miller, KD., Jemal, A. Cancer Statistics, 2016. Ca Cancer J Clin. 2016; 66:7–30. [PubMed: 26742998]

4. Ryerson, Ab, Eheman, CR., Altekruse, SF., Ward, JW., Jemal, A., Sherman, RL., Henley, SJ., Holtzman, D., Lake, A., Noone, AM., Anderson, RN., Ma, J., Ly, KN., Cronin, KA., Penberthy, L., Kohler, BA. Annual Report To The Nation On The Status Of Cancer, 1975–2012, Featuring The Increasing Incidence Of Liver Cancer. Cancer. 2016; 122:1312–37. [PubMed: 26959385]

5. Hamra, Gb, Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, JM., Vineis, P., Forastiere, F., Saldiva, P., Yorifuji, T., Loomis, D. Outdoor Particulate Matter Exposure And Lung Cancer: A Systematic Review And Meta-Analysis. Environmental Health Perspectives. 2014; 122:906–11. [PubMed: 24911630]

6. Loomis D, Huang W, Chen G. The International Agency For Research On Cancer (Iarc) Evaluation Of The Carcinogenicity Of Outdoor Air Pollution: Focus On China. Chinese Journal Of Cancer. 2014; 33:189–96. [PubMed: 24694836]

7. Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K. International Agency For Research On Cancer Monograph Working Group I. The Carcinogenicity Of Outdoor Air Pollution. The Lancet Oncology. 2013; 14:1262–3. [PubMed: 25035875]

8. Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, Turner MC, Pope CA 3rd, Thurston G, Calle EE, Thun MJ, Beckerman B, Deluca P, Finkelstein N, Ito K, Moore DK, Newbold KB, Ramsay T, Ross Z, Shin H, Tempalski B. Extended Follow-Up And Spatial Analysis Of The

American Cancer Society Study Linking Particulate Air Pollution And Mortality. Research Report. 2009:5–114. Discussion 5–36.

9. Vineis P, Hoek G, Krzyzanowski M, Vigna-Taglianti F, Veglia F, Airoldi L, Autrup H, Dunning A, Garte S, Hainaut P, Malaveille C, Matullo G, Overvad K, Raaschou-Nielsen O, Clavel-Chapelon F, Linseisen J, Boeing H, Trichopoulou A, Palli D, Peluso M, Krogh V, Tumino R, Panico S, Bueno-De-Mesquita HB, Peeters PH, Lund EE, Gonzalez CA, Martinez C, Dorronsoro M, Barricarte A, Cirera L, Quiros JR, Berglund G, Forsberg B, Day NE, Key TJ, Saracci R, Kaaks R, Riboli E. Air Pollution And Risk Of Lung Cancer In A Prospective Study In Europe. Int J Cancer. 2006; 119:169–74. [PubMed: 16463382]

10. Raaschou-Nielsen O, Andersen Zj, Beelen R, Samoli E, Stafoggia M, Weinmayr G, et al. Air Pollution And Lung Cancer Incidence In 17 European Cohorts: Prospective Analyses From The European Study Of Cohorts For Air Pollution Effects (Escape). The Lancet Oncology. 2013; 14:813–22. [PubMed: 23849838]

11. Blackley D, Zheng S, Ketchum W. Implementing A Weighted Spatial Smoothing Algorithm To Identify A Lung Cancer Belt In The United States. Cancer Epidemiology. 2012; 36:436–8. [PubMed: 22728280]

12. Blot, Wj, Fraumeni, JF, Jr. Geographic Patterns Of Lung Cancer: Industrial Correlations. Am J Epidemiol. 1976; 103:539–50. [PubMed: 937339]

13. Devesa, Ss, Grauman, DJ., Blot, WJ., Fraumeni, JF, Jr. Cancer Surveillance Series: Changing Geographic Patterns Of Lung Cancer Mortality In The United States, 1950 Through 1994. J Natl Cancer Inst. 1999; 91:1040–50. [PubMed: 10379967]

14. Underwood, Jm, Townsend, JS., Tai, E., Davis, SP., Stewart, SL., White, A., Momin, B., Fairley, TL. Racial And Regional Disparities In Lung Cancer Incidence. Cancer. 2012; 118:1910–8. [PubMed: 21918961]

15. Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, Mn: 2011. [Cited 2016 May 15]; Available From: Https://Www.Nhgis.Org/

16. County Health Rankings & Roadmaps. 2016. [Cited 2016 May 15]; Available From: Http://Www.Countyhealthrankings.Org

17. Al-Hamdan, Mz, Crosson, WL., Economou, SA., Estes, MG., Jr, Estes, SM., Hemmings, SN., Kent, ST., Puckett, M., Quattrochi, DA., Rickman, DL., Wade, GM., Mcclure, LA. Environmental Public Health Applications Using Remotely Sensed Data. Geocarto International. 2014; 29:85–98. [PubMed: 24910505]

18. Gruca, Ts, Pyo, TH., Nelson, GC. Improving Rural Access To Orthopaedic Care Through Visiting Consultant Clinics. The Journal Of Bone And Joint Surgery American Volume. 2016; 98:768–74. [PubMed: 27147690]

19. Loop, Ms, Kent, ST., Al-Hamdan, MZ., Crosson, WL., Estes, SM., Estes, MG., Jr, Quattrochi, DA., Hemmings, SN., Wadley, VG., Mcclure, LA. Fine Particulate Matter And Incident Cognitive Impairment In The Reasons For Geographic And Racial Differences In Stroke (Regards) Cohort. Plos One. 2013; 8:E75001. [PubMed: 24086422]

20. Kirchhoff, Ac, Hart, G., Campbell, EG. Rural And Urban Primary Care Physician Professional Beliefs And Quality Improvement Behaviors. The Journal Of Rural Health: Official Journal Of The American Rural Health Association And The National Rural Health Care Association. 2014; 30:235–43.

21. Census. Geographic Terms And Concepts - Census Divisions And Census Regions. 2016. [Cited 2016 May 15]; Available From: Https://Www.Census.Gov/Geo/Reference/Gtc/Gtc_Census_Divreg.Html

22. Cdc. Wide-Ranging Online Data For Epidemiologic Research (Cdc-Wonder). 2016. [Cited 2016 May 15]; Available From: Http://Wonder.Cdc.Gov

23. Moore, Jx, Donnelly, JP., Griffin, R., Howard, G., Safford, MM., Wang, HE. Defining Sepsis Mortality Clusters In The United States. Crit Care Med. 2016; 44:1380–7. [PubMed: 27105174]

24. Wartenberg D. Investigating Disease Clusters: Why, When, And How? Jr Statist Soc A. 2001; 164:13–22.

25. Nassel, Af, Root, ED., Haukoos, JS., Mcvaney, K., Colwell, C., Robinson, J., Eigel, B., Magid, DJ., Sasson, C. Multiple Cluster Analysis For The Identification Of High-Risk Census Tracts For

Out-Of-Hospital Cardiac Arrest (Ohca) In Denver, Colorado. Resuscitation. 2014; 85:1667–73. [PubMed: 25263511]

26. Anselin L. Local Indicators Of Spatial Association - Lisa. Geogr Anal. 1995; 27:93–115.

27. Getis A, Ord K. The Analysis Of Spatial Association By Use Of Distance Statistics. Geogr Anal. 1992; 24:189–206.

28. Anselin, L. Exploring Spatial Data With Geoda: A Workbook. Urbana-Champaign: University Of Illinois; 2004.

29. Li H, Calder CA, Cressie NAC. Beyond Moran's I: Testing For Spatial Dependence Based On The Spatial Autoregressive Model. Geogr Anal. 2007; 39:357–75.

30. Li H, Calder CA, Cressie NAC. One-Step Estimation Of Spatial Dependence Parameters: Properties And Extensions Of The Aple Statistic. J Multivariate Anal. 2012; 105:68–84.

31. Akinyemiju T, Moore JX, Ojesina AI, Waterbor JW, Altekruse SF. Racial Disparities In Individual Breast Cancer Outcomes By Hormone-Receptor Subtype, Area-Level Socio-Economic Status And Healthcare Resources. Breast Cancer Res Treat. 2016; 157:575–86. [PubMed: 27255533]

32. Akinyemiju T, Waterbor JW, Pisu M, Moore JX, Altekruse SF. Availability Of Healthcare Resources And Colorectal Cancer Outcomes Among Non-Hispanic White And Non-Hispanic Black Adults. Journal Of Community Health. 2016; 41:296–304. [PubMed: 26446012]

33. Tannenbaum, Sl, Koru-Sengul, T., Zhao, W., Miao, F., Byrne, MM. Survival Disparities In Non-Small Cell Lung Cancer By Race, Ethnicity, And Socioeconomic Status. Cancer Journal. 2014; 20:237–45.

34. Torre, La, Siegel, RL., Jemal, A. Lung Cancer Statistics. Advances In Experimental Medicine And Biology. 2016; 893:1–19. [PubMed: 26667336]

35. Howard G, Evans GW, Pearce K, Howard VJ, Bell RA, Mayer EJ, Burke GL. Is The Stroke Belt Disappearing? An Analysis Of Racial, Temporal, And Age Effects. Stroke. 1995; 26:1153–8. [PubMed: 7604406]

36. Howard G. Why Do We Have A Stroke Belt In The Southeastern United States? A Review Of Unlikely And Uninvestigated Potential Causes. American Journal Of The Medical Sciences. 1999; 317:160–7. [PubMed: 10100689]

37. Howard, Vj, Cushman, M., Pulley, L., Gomez, CR., Go, RC., Prineas, RJ., Graham, A., Moy, CS., Howard, G. The Reasons For Geographic And Racial Differences In Stroke Study: Objectives And Design. Neuroepidemiology. 2005; 25:135–43. [PubMed: 15990444]

38. Le A, Judd SE, Allison DB, Oza-Frank R, Affuso O, Safford MM, Howard VJ, Howard G. The Geographic Distribution Of Obesity In The Us And The Potential Regional Differences In Misreporting Of Obesity. Obesity (Silver Spring). 2014; 22:300–6. [PubMed: 23512879]

39. Pickle, Lw, Mungiole, M., Gillum, RF. Geographic Variation In Stroke Mortality In Blacks And Whites In The United States. Stroke. 1997; 28:1639–47. [PubMed: 9259762]

40. Pickle, Lw, Gillum, RF. Geographic Variation In Cardiovascular Disease Mortality In Us Blacks And Whites. J Natl Med Assoc. 1999; 91:545–56. [PubMed: 10599187]

41. Lanska, Dj, Kuller, LH. The Geography Of Stroke Mortality In The United States And The Concept Of A Stroke Belt. Stroke. 1995; 26:1145–9. [PubMed: 7604404]

42. Gillum, Rf, Ingram, DD. Relation Between Residence In The Southeast Region Of The United States And Stroke Incidence: The Nhanes I Epidemiologic Followup Study. American Journal Of Epidemiology. 1996; 144:665–73. [PubMed: 8823063]

43. Casper, Ml, Wing, S., Anda, RF., Knowles, M., Pollard, RA. The Shifting Stroke Belt: Changes In The Geographic Pattern Of Stroke Mortality In The United States, 1962 To 1988. Stroke. 1995; 26:755–60. [PubMed: 7740562]

44. Lanska, Dj, Peterson, PM. Geographic Variation In The Decline Of Stroke Mortality In The United States. Stroke. 1995; 26:1159–65. [PubMed: 7604407]

45. Howard G, Anderson R, Johnson NJ, Sorlie P, Russell G, Howard VJ. Evaluation Of Social Status As A Contributing Factor To The Stroke Belt Region Of The United States. Stroke. 1997; 28:936–40. [PubMed: 9158628]

46. Moore, Jx, Donnelly, Jp, Griffin, R., Howard, G., Safford, Mm, Wang, He. Defining Sepsis Mortality Clusters In The United States. Crit Care Med. 2016; 44:1380–7. [PubMed: 27105174]

47. Casper, Ml, Wing, S., Anda, RF., Knowles, M., Pollard, RA. The Shifting Stroke Belt. Changes In The Geographic Pattern Of Stroke Mortality In The United States, 1962 To 1988. Stroke. 1995; 26:755–60. [PubMed: 7740562]

48. Thomas, Tl, Diclemente, R., Snell, S. Overcoming The Triad Of Rural Health Disparities: How Local Culture, Lack Of Economic Opportunity, And Geographic Location Instigate Health Disparities. Health Education Journal. 2014; 73:285–94. [PubMed: 25242822]

49. Gardner T, Gavaza P, Meade P, Adkins DM. Delivering Free Healthcare To Rural Central Appaclachia Population: The Case Of The Health Wagon. Rural And Remote Health. 2012; 12:2035. [PubMed: 22452285]

50. Lyon, Sm, Douglas, IS., Cooke, CR. Medicaid Expansion Under The Affordable Care Act. Implications For Insurance-Related Disparities In Pulmonary, Critical Care, And Sleep. Ann Am Thorac Soc. 2014; 11:661–7. [PubMed: 24708065]

51. Northridge, Me, Stover, GN., Rosenthal, JE., Sherard, D. Environmental Equity And Health: Understanding Complexity And Moving Forward. Am J Public Health. 2003; 93:209–14. [PubMed: 12554571]

52. Hendryx M, Fedorko E, Halverson J. Pollution Sources And Mortality Rates Across Rural-Urban Areas In The United States. The Journal Of Rural Health: Official Journal Of The American Rural Health Association And The National Rural Health Care Association. 2010; 26:383–91.

53. Ailshire, Ja, Crimmins, EM. Fine Particulate Matter Air Pollution And Cognitive Function Among Older Us Adults. Am J Epidemiol. 2014; 180:359–66. [PubMed: 24966214]

54. An R, Xiang X. Ambient Fine Particulate Matter Air Pollution And Leisure-Time Physical Inactivity Among Us Adults. Public Health. 2015; 129:1637–44. [PubMed: 26277287]

55. Autrup H. Aimbient Air Pollution And Adverse Health Effects. Procedia Social And Behavioral Sciences. 2010; 2:7333–8.

56. Ueng, Th, Hung, CC., Kuo, ML., Chan, PK., Hu, SH., Yang, PC., Chang, LW. Induction Of Fibroblast Growth Factor-9 And Interleukin-1alpha Gene Expression By Motorcycle Exhaust Particulate Extracts And Benzo(A)Pyrene In Human Lung Adenocarcinoma Cells. Toxicological Sciences: An Official Journal Of The Society Of Toxicology. 2005; 87:483–96. [PubMed: 16002475]

57. Sakai N, Tager AM. Fibrosis Of Two: Epithelial Cell-Fibroblast Interactions In Pulmonary Fibrosis. Biochimica Et Biophysica Acta. 2013; 1832:911–21. [PubMed: 23499992]

58. Brook, Rd, Rajagopalan, S., Pope, Ca, 3rd, Brook, Bhatnagar, A., Diez-Roux, Av, et al. Particulate Matter Air Pollution And Cardiovascular Disease: An Update To The Scientific Statement From The American Heart Association. Circulation. 2010; 121:2331–78. [PubMed: 20458016]

59. Johnson, Rl, Jr. Relative Effects Of Air Pollution On Lungs And Heart. Circulation. 2004; 109:5–7. [PubMed: 14707016]

60. Pope, Ca, 3rd, Burnett, RT., Thurston, GD., Thun, MJ., Calle, EE., Krewski, D., Godleski, JJ. Cardiovascular Mortality And Long-Term Exposure To Particulate Air Pollution: Epidemiological Evidence Of General Pathophysiological Pathways Of Disease. Circulation. 2004; 109:71–7. [PubMed: 14676145]

61. Young, Mt, Sandler, DP., Deroo, LA., Vedal, S., Kaufman, JD., London, SJ. Ambient Air Pollution Exposure And Incident Adult Asthma In A Nationwide Cohort Of U.S. Women. Am J Respir Crit Care Med. 2014; 190:914–21. [PubMed: 25172226]

62. Neupane B, Jerrett M, Burnett Rt, Marrie T, Arain A, Loeb M. Long-Term Exposure To Ambient Air Pollution And Risk Of Hospitalization With Community-Acquired Pneumonia In Older Adults. Am J Respir Crit Care Med. 2010; 181:47–53. [PubMed: 19797763]

63. Puett, Rc, Hart, JE., Yanosky, JD., Spiegelman, D., Wang, M., Fisher, JA., Hong, B., Laden, F. Particulate Matter Air Pollution Exposure, Distance To Road, And Incident Lung Cancer In The Nurses' Health Study Cohort. Environmental Health Perspectives. 2014; 122:926–32. [PubMed: 24911062]

64. Pope, Ca, 3rd, Burnett, RT., Thun, MJ., Calle, EE., Krewski, D., Ito, K., Thurston, GD. Lung Cancer, Cardiopulmonary Mortality, And Long-Term Exposure To Fine Particulate Air Pollution. Jama. 2002; 287:1132–41. [PubMed: 11879110]

65. Li H, Li Qd, Wang Ms, Li Fj, Li Qh, Ma Xj, et al. Smoking And Air Pollution Exposure And Lung Cancer Mortality In Zhaoyuan County. International Journal Of Hygiene And Environmental Health. 2013; 216:63–70. [PubMed: 22841879]

66. Turner, Mc, Cohen, A., Jerrett, M., Gapstur, Sm, Diver, Wr, Pope, Ca, 3rd, et al. Interactions Between Cigarette Smoking And Fine Particulate Matter In The Risk Of Lung Cancer Mortality In Cancer Prevention Study Ii. Am J Epidemiol. 2014; 180:1145–9. [PubMed: 25395026]

67. Akinyemiju T, Jha M, Moore Jx, Pisu M. Disparities In The Prevalence Of Comorbidities Among Us Adults By State Medicaid Expansion Status. Preventive Medicine. 2016; 88:196–202. [PubMed: 27095325]

68. Jemal A, Ward Em, Johnson Cj, Cronin Ka, Ma J, Ryerson B, et al. Annual Report To The Nation On The Status Of Cancer, 1975–2014, Featuring Survival. J Natl Cancer Inst. 2017:109.

69. Monn C. Exposure Assessment Of Air Pollutants: A Review On Spatial Heterogeneity And Indoor/Outdoor/Personal Exposure To Suspended Particulate Matter, Nitrogen Dioxide And Ozone. Atmospheric Environment. 2001; 35:1–32.

70. Jacobs, Ej, Newton, CC., Carter, BD., Feskanich, D., Freedman, ND., Prentice, RL., Flanders, WD. What Proportion Of Cancer Deaths In The Contemporary United States Is Attributable To Cigarette Smoking? Ann Epidemiol. 2015; 25:179–82 E1. [PubMed: 25487970]

Highlights

- Clusters of high lung cancer mortality are defined using novel geospatial methods.

- Clustered counties ranged from eastern Oklahoma through central Appalachia.

- Clustered counties had higher pollution, rural population, and lower healthcare access.
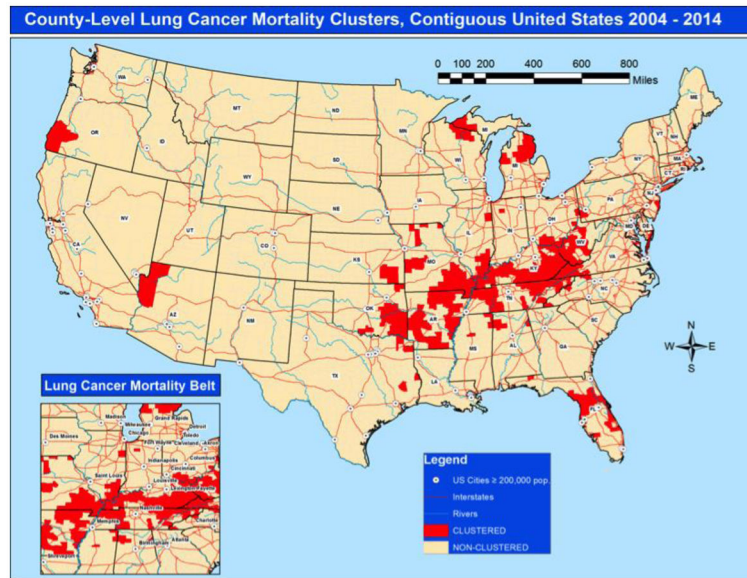
Figure 1.
Lung cancer mortality clusters for county-level regional variation, contiguous United States, 2004–2014.
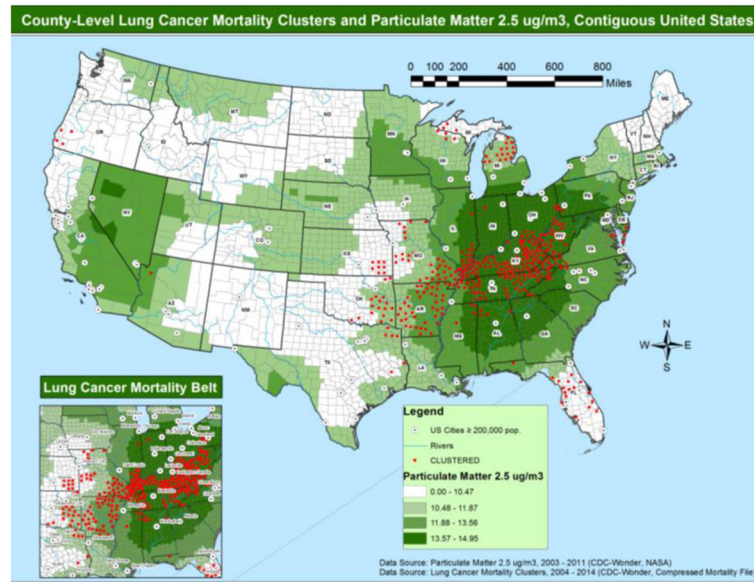
Figure 2.
County-level lung cancer mortality clusters and particulate matter 2.5 μg/m^3, contiguous United States.

Table 1

Comparison of geospatial clustering[a] by counties within United States regions, 2004–2014.

	Midwest counties[b] (N = 1,055)	Northeast counties[b] (N = 217)	Southern counties[b] (N = 1,422)	Western counties[b] (N = 414)	Total (N = 3108)	P value[c]
	Presented as N (%)[d]				N (%)[e]	
LISA						
High-high cluster	112 (25.11)	2 (0.45)	328 (73.54)	4 (0.90)	446 (14.35)	<0.01
High-low outlier	16 (34.78)	1 (2.17)	6 (13.04)	23 (50.00)	46 (1.48)	
Low-high outlier	13 (28.26)	0 (0.00)	32 (69.57)	1 (2.17)	46 (1.48)	
Low-low cluster	94 (20.75)	26 (5.74)	129 (28.48)	204 (45.03)	453 (14.58)	
Non-significant	820 (38.79)	188 (8.89)	927 (43.85)	182 (8.61)	2116 (68.08)	
EB						
1st Quintile	155 (24.92)	34 (5.47)	202 (32.48)	231 (37.14)	622 (20.01)	<0.01
2nd Quintile	250 (40.26)	68 (10.95)	215 (34.62)	88 (14.17)	621 (19.98)	
3rd Quintile	246 (39.55)	69 (11.09)	269 (43.25)	38 (6.11)	622 (20.01)	
4th Quintile	223 (35.91)	40 (6.44)	325 (52.33)	33 (5.31)	621 (19.98)	
5th Quintile	181 (29.10)	6 (0.96)	411 (66.08)	24 (3.86)	622 (20.01)	
Gi* (%)						
Lung cancer hotspot	42 (18.58)	70 (30.97)	65 (28.76)	49 (21.68)	226 (7.27)	<0.01
Lung cancer cold spot	337 (53.58)	0 (0.00)	148 (23.53)	144 (22.89)	629 (20.24)	
Non-significant	676 (30.04)	147 (6.53)	1209 (53.73)	221 (9.82)	2252 (72.46)	
Lung cancer mortality cluster category						
Clustered	91 (25.14)	2 (0.55)	264 (72.93)	5 (1.38)	362 (11.65)	<0.01
Non-clustered	964 (35.11)	215 (7.83)	1158 (42.17)	409 (14.89)	2746 (88.35)	

[a] Defined as counties estimated as clustered by two of three cluster methods (local indicators of spatial autocorrelation, empirical Bayes, and Gi*)

[b] US regions as determined by the US Census Bureau.

[c] Significance determined using chi-square test.

[d] Denotes row percentages.

[e] Denotes column percentages.

Author Manuscript

LISA, local indicators of spatial association; EB, empirical Bayes.

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

County-level demographic and community characteristics by lung cancer mortality cluster category.

	Clustered[a] (*N* = 362)	Non-clustered (*N* = 2,746)	*P* value[b]	ρ[c]
	Presented as median (IQR)			
Race				
NH White	93.95 (86.10–96.20)	84.10 (65.20–93.30)	<0.01	0.23
NH Black	1.60 (0.40–5.40)	2.10 (0.50–11.30)	<0.01	−0.06
Hispanic	0.40 (0.20–0.70)	3.80 (1.90–9.60)	<0.01	−0.25
Asian	0.30 (0.02–0.05)	0.60 (0.40–1.20)	<0.01	−0.25
Native American	0.40 (0.20–0.07)	0.50 (0.30–1.20)	<0.01	−0.13
Sex				
Male	49.31 (48.69–49.99)	49.60 (48.88–50.51)	<0.01	−0.07
Female	50.69 (50.01–51.31)	50.40 (49.49–51.12)		
Age				
<18	21.90 (20.16–23.46)	22.99 (21.01–24.86)	<0.01	−0.14
18–29	13.04 (12.07–14.06)	14.22 (12.40–16.25)	<0.01	−0.16
30–44	17.62 (16.29–18.64)	17.72 (16.29–19.05)	0.06	−0.03
45–64	28.70 (27.50–29.94)	27.93 (26.05–29.52)	<0.01	0.13
65–79	13.74 (12.36–15.58)	11.77 (10.04–13.66)	<0.01	0.25
80+	4.46 (3.83–5.34)	4.17 (3.34–5.16)	<0.01	0.09
Community characteristics				
Particulate matter 2.5 μg/m^3	13.61 (11.54–14.02)	11.70 (10.35–13.30)	<0.01	0.18
Household income <$20,000	26.91 (23.69–30.67)	19.82 (16.05–24.42)	<0.01	0.34
% Completed college	12.87 (10.93–15.72)	18.59 (14.65–24.76)	<0.01	−0.32
% Adult smoking	25.50 (21.80–30.30)	18.70 (13.90–23.20)	<0.01	0.30
% Adult obesity	32.90 (30.90–34.50)	30.30 (28.10–32.80)	<0.01	0.21
% Physical inactivity	33.10 (29.60–35.90)	27.70 (24.20–30.60)	<0.01	0.31
% Unemployment	9.00 (7.80–10.40)	7.30 (5.50–9.10)	<0.01	0.24
% Uninsured	23.10 (20.30–25.60)	21.05 (16.30–26.20)	<0.01	0.11
% Could not see doctor due to cost	18.35 (13.60–22.60)	12.00 (0.00–16.50)	<0.01	0.23
PCP per 10,000 persons	4.18 (2.57–6.02)	4.99 (3.21–7.12)	<0.01	−0.09
Non-urban[d]	265 (73.20)	1576 (57.39)	<0.01	0.10

[a]Defined as counties estimated as clustered by two of three cluster methods (local indicators of spatial autocorrelation, empirical Bayes, and Gi*).

[b]Significance determined using Wilcoxon test.

NH, non-Hispanic; PCP, primary-care physician.

[c]Spearman correlation with county-level lung cancer mortality cluster category.

[d]Presented as N (Column %)

IQR, interquartile range; PCP, primary-care physician.

Table 3

Lung cancer mortality[a] in the United States by clustering category, 2004–2014, excluding Alaska and Hawaii.

	Lung cancer clustering category		Total
	Clustered[a] (N = 362)	**Non-clustered (N = 2,746)**	**All counties (N = 3,108)**
Lung cancer deaths[b]	142,380	1,612,527	1,754,907
Mean county population	336,652	931,526	862,239
Crude Lung cancer mortality rate (95%CI)	116.80 (116.20 – 117.40)	63.00 (62.90 – 63.10)	65.50 (65.40 – 65.60)
Age adjusted (95%CI)	99.70 (99.10 – 100.30)	75.60 (75.40 – 75.70)	78.20 (78.00 – 78.30)

[a]Annual deaths per 100,000 persons.

[b]Total number of deaths, identified using the *International Classification of Diseases,* 10th Version codes for lung cancer.

Table 4

Odds ratios (ORs) and associated 95% confidence intervals (95%CIs) for the association between quartile of particulate matter 2.5 μg/m³ and lung cancer mortality clustering.

	Clustered counties (N = 362) N (%)	OR (95%CI)			
		Crude	Model 1[a]	Model 2[b]	Model 3[c]
Particulate matter[e] 2.5μg/m³					
1st quartile (0.00–10.47)	46 (12.71)	Referent	Referent	Referent	Referent
2nd quartile (10.47–11.87)	58 (16.02)	1.28 (0.86–1.91)	1.36 (0.86–2.15)	1.06 (0.63–1.78)	1.25 (0.76–2.05)
3rd quartile (11.87–13.56)	61 (16.85)	1.35 (0.91–2.01)	1.90 (1.14–3.17)	1.58 (0.85–2.92)	1.74 (1.00–3.04)
4th quartile (13.56–14.95)	197 (54.42)	5.41 (3.85–7.58)	3.48 (2.15–5.66)	1.59 (0.88–2.88)	2.10 (1.23–3.59)
Non-urban	**Clustered counties (N = 265)**				
Particulate matter[e] 2.5μg/m³					
1st quartile (7.95–10.19)	25 (9.43)	Referent	Referent	Referent	Referent
2nd quartile (10.19–11.40)	44 (16.60)	1.85 (1.11–3.07)	1.62 (0.91–2.89)	1.30 (0.68–2.51)	1.27 (0.70–2.29)
3rd quartile (11.40–13.06)	52 (19.62)	2.22 (1.35–3.65)	2.68 (1.48–4.84)	1.80 (0.88–2.68)	2.35 (1.27–4.34)
4th quartile (13.06–14.91)	144 (54.34)	7.95 (5.07–12.45)	4.06 (2.23–7.39)	1.86 (0.90–3.87)	3.19 (1.70–5.96)
Urban	**Clustered Counties (N = 97)**				
Particulate matter[d] 2.5μg/m³					
1st quartile (0.00–10.99)	15 (15.46)	Referent	Referent	Referent	Referent
2nd quartile (10.99–12.62)	11 (11.34)	0.72 (0.33–1.60)	0.96 (0.36–2.55)	1.26 (0.41–3.86)	1.20 (0.42–3.93)
3rd quartile (12.62–13.79)	27 (27.84)	1.87 (0.98–3.60)	1.72 (0.67–4.41)	1.34 (0.40–4.48)	2.26 (0.80–6.37)
4th quartile (13.79–14.95)	44 (45.36)	3.26 (1.77–5.99)	1.57 (0.65–3.78)	1.22 (0.41–3.63)	1.93 (0.74–5.02)

%Column percentage.

[a] Adjusted for age, race, and sex.

[b] Model 1 plus adjustment for household income, % completed college, adult smoking, adult obesity, and physical inactivity.

[c] Model 1 plus adjustment for uninsured, could not see doctor due to cost, and PCP.

[d] Mean particulate matter value per county from 2003–2011.

Urban denotes population centers with 50,000 or more residents

Non-urban denotes towns or small cities with population centers with less than 50,000 residents.