Lung Cancer 82 (2013) 55–62

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Spatial and temporal distributions of lung cancer histopathology in the state of Maine

H. Dean Hosgood III[a,*], Christopher Farah[b,c], Candice C. Black[d], Molly Schwenn[e], Janet M. Hock[b,c]

[a] Division of Epidemiology, Albert Einstein College of Medicine, Bronx, NY, United States
[b] Maine Institute for Human Genetics and Health, Brewer, ME, United States
[c] The Polis Center and the Center for Health Geographics, Indiana University-Purdue University in Indianapolis, Indianapolis, IN, United States
[d] Dartmouth-Hitchcock Medical Center, Pathology, Lebanon, NH, United States
[e] Maine Cancer Registry, Maine Center for Disease Control and Prevention, DHHS, Augusta, ME, United States

ARTICLE INFO

Article history:
Received 3 January 2013
Received in revised form 26 April 2013
Accepted 28 June 2013

Keywords:
Lung cancer
Subtype
New England
Wood
Radon

ABSTRACT

Maine has among the highest rates of lung cancer in the United States (US). Maine serves as a geographical representation of US rural communities, and their associated health disparities. As the key risks of tobacco use decrease and radon abatement increases, previously obscured environmental exposures may measurably contribute to the attributable risk fraction of lung cancer. To generate hypotheses of novel environmental exposures associated with lung cancer, we investigated if there was non-random spatial distribution of lung cancer in Maine. Case data (n = 14,038) between 1995 and 2006 were obtained from the Maine Cancer Registry. Population data were obtained from the 2000 US Census. We assessed the spatial distribution of lung cancers among white cases by histopathology subtype [non-small cell lung carcinoma (NSCLC): adenocarcinoma (n = 3680), squamous cell (n = 2801) and large cell (n = 1195); and small cell lung carcinoma (SCLC) (n = 1994)], using spatial scan statistic, assuming a discrete Poisson distribution adjusted for age and population density. Because of time-dependent trends in lung cancer differential diagnostic criteria, we repeated our analyses, limiting it to 2002–2006. While SCLC rates were equivalent across the state, we identified discrete regions with elevated rates of adenocarcinoma among females and squamous cell carcinoma among males. Independent of gender, the most striking geospatial observation was elevated large cell lung cancer specifically in one of the poorest counties in the US. A selective spatial distribution of large cell lung cancer has not been previously reported. More research is needed to identify factors inducing large cell carcinoma pathology, and to determine if in rural communities health disparities are associated with increased risk for this diagnosis.

1. Introduction

Lung cancer is a leading cause of death among cancers in the United States (US) [1]. Maine has among the highest age-adjusted incidence and mortality rates of lung cancer among US states [2]. Maine reports significant health and cancer disparities related to poverty, geographic isolation and health access [3]. As these challenges are common to most rural areas of the US, Maine with 16 large counties and a population of 1.4 million, provides a unique opportunity to study lung cancer in an under-studied rural population, and to establish a baseline for comparison to other disadvantaged groups in the US and globally.

Exposures to tobacco [4,5] and radon [6] are among the highest risks for lung cancer in the US. Smoking rates in Maine vary from 20 to 30%, and the downward trend in tobacco use over the past decade mirrors that of the rest of the nation [7,8]. We recently reported on spatial variations in radon exposure in the state, and found high mean radon activity in several well-defined regions [9]. As tobacco smoking rates continue to decline and awareness of the health risks associated with radon increases [10], research is needed to identify additional environmental exposures that may contribute measurably to the attributable risk fraction of lung cancer.

Differences in histopathology may provide clues to whether environmental exposures contribute to Maine's high rates of lung cancer. For example, tobacco smoking had long been associated with squamous cell carcinoma and small cell carcinoma of the lung [11]. Connecticut data from 1959 to 1991 identified increasing

* Corresponding author at: Albert Einstein College of Medicine, Division of Epidemiology and Population Health, Division of Epidemiology 1300 Morris Park Ave, Belfer 1309, Bronx, New York 10461, United States. Tel.: +1 718 430 2138; fax: +1 718 430 8780.
E-mail address: dean.hosgood@einstein.yu.edu (H.D. Hosgood III).

0169-5002/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.lungcan.2013.06.018

trends of lung cancer for an additional pathological subtype, adenocarcinoma, that correlated with increasing cigarette smoking by men and women, but not birth cohort or changes in lung cancer diagnosis criteria [12]. Further, a geospatial study of the high lung cancer rates in the Appalachia regions of coal mining yielded new information regarding the increased diversity of environmental exposure risks in disadvantaged populations with increased lung cancer incidence and mortality rates [13], suggesting more effective risk mitigation for public health strategies.

To generate hypotheses of novel environmental exposures associated with lung cancer, we used Maine Cancer Registry data to investigate if there was non-random spatial distribution of lung cancer in Maine. Research into the etiology of lung cancer in Maine will help not only reduce the more than 1200 newly diagnosed cases in Maine each year, but also provide very important clues for lung cancer mitigation and prevention throughout North America.

Table 1
Number of Maine Cancer Registry lung cancer cases among whites from 1995 to 2006 for selected histologies, by gender.

	Males		Females		All subjects	
	N	%[a]	*N*	%[a]	*N*	%[a]
Non-small cell lung carcinoma						
Adenocarcinoma	1875	34.6	1805	42.5	3680	38.1
Squamous cell carcinoma	1899	35.0	902	21.2	2801	29.0
Large cell carcinoma	676	12.5	519	12.2	1195	12.4
Small cell lung carcinoma						
Small cell carcinoma	971	17.9	1020	24.0	1994	20.6
Total[b]	5424	56.1	4246	43.9	9670	100.0

[a] Percent of cases observed in males ($n = 5424$), females ($n = 4246$), and all subjects ($n = 9670$), respectively.
[b] Percent of all cases observed ($n = 9670$).

2. Methods

This protocol was approved by the Eastern Maine Medical Center Institutional Review Board and US Office for Human Protection (OHP) Human Research Protections Office (HRPO). All lung cancer cases in Maine, 1995–2006 ($n = 14{,}038$) were obtained from the Maine Cancer Registry as a de-identified dataset under a Data Use Agreement. Data elements used from these records were: histology, ZIP code of residence at diagnosis, age at diagnosis, sex, and race. The ability to evaluate the contribution of race or ethnicity to disease incidence is limited in Maine since it is predominantly white, and minority populations, comprised of Native Americans, Blacks, Hispanics and Asians, constitute only about 5% of the total population [14]. Therefore, we restricted our spatial analyses of lung cancer to white Maine residents ($n = 13{,}917$). Exclusion of cancers not originating in the lung resulted in 11,153 white lung cancers for analysis. Lung cancer subtypes selected for this investigation were non-small cell lung carcinoma (NSCLC) [adenocarcinoma, large cell carcinoma, squamous cell carcinoma] and small cell lung carcinoma (SCLC) (combined $n = 9670$). Based on patient ZIP code of residence at diagnosis, we aggregated patient counts for each Maine ZIP code by subtype, gender, and age category. We used Esri ArcMap to process geographic data. This included geocoding cases to the ZIP code level of resolution and aggregating cases to each ZIP code. We used the US Census 2000 SF3 file [15], with population counts for each ZIP code by age category and gender to perform spatial analyses. Age was categorized as under 5, 5–9, 10–14, 15–19, 20–24, 25–34, 45–54, 55–59, 60–64, 65–74, 75–84, and 85 & over. Supporting GIS data for this study included: feature class *Roads of Maine* demarcating county boundaries at 1:100,000 scale, feature class *Mountains*, representing the mountainous regions of Maine, and feature class *ZIP code* demarcating 2000 postal ZIP code boundaries at 1:100,000 scale, including the latitude and longitude of each ZIP code's centroid. The latter feature class was used to generate a locator reference file to geocode the patient data. Feature classes were obtained from the Maine Office of GIS online database [16]. Spatial scan statistic [17] was applied via SaTScan software [18] to detect regions with significant differences ($p < 0.05$) in disease incidence, assuming a discrete Poisson distribution [19]. This method imposes Monte Carlo randomization to determine the expected number of cases per unit zonal region and adjusts for the multiple testing inherent in the many zones considered. Our cluster detection method identified clusters of high and low incidence, with a maximum scanning window size that included up to 10% of the total ZIP codes under investigation. This scanning window was selected as optimal after evaluation of clusters using scanning windows up to 50% (software default) of the total population [20]. Secondary clusters were reported if they had no geographic overlap with more likely clusters. *p*-Values were derived from 999 simulated Monte Carlo replications, assuming a null hypothesis of spatial randomness of histopathology subtype. Spatial variations by gender were assessed for each of the four lung cancer subtypes and for overall lung cancer cases. Each analysis assumed a discrete Poisson distribution, using Census 2000 data as the underlying population, adjusted for population density and age. Rejecting the null hypothesis indicated a significant variation in the distribution of lung cancer cases overall, or by subtype, for each gender. Figures show colored spatial regions on cancer histology maps as: Red [Regions where spatial incidence was significantly higher, $p < 0.05$], Blue [Regions where spatial incidence was significantly lower, $p < 0.05$], or Yellow [Regions where differences in incidence were not statistically significant]. We also investigated spatiotemporal trends to determine if there were changes in incidence of each subtype. Because of changes in American College of Pathologists' diagnostic criteria due to new insights into NSCLC carcinogenesis over the past two decades [21,22], we analyzed variations in spatial incidence over time to determine if there were significant changes in incidence of each subtype between 1995 and 2006. We performed various sensitivity analyses to determine if significant regions of specific histopathology were associated with radon or tobacco exposures. Because certain geological formations generate radon concentrations well above the EPA maximum contaminant level (MCL) for localized regions of Maine [9,23], and radon has been implicated in risk of lung cancer [24–26], spatial autocorrelation results from our Maine radon dose study [9] were compared to the results derived from this research. Finally, because our data, like most cancer registry data, did not include comprehensive records on smoking status of patients, we compared the identified significant regions to a map of smoking rates, at the ZIP code level, reported using Behavioral Risk Factor Surveillance System (BRFSS) data, aggregated from 2005 to 2009 [27], to determine if smoking might contribute in part to the higher incidence. Due to privacy requirements associated with the sparse population of Maine, BRFSS only provided survey data for the ZIP codes with sufficient population size.

3. Results

We carried out in-depth geospatial analyses of lung cancer in the state of Maine (Fig. 1). Our review of the Maine Cancer Registry data yielded a total of 9670 white subjects diagnosed with lung cancer of a major histological subtype (Table 1). Adenocarcinoma ($n = 3680$) was the most prevalent histology, followed by squamous cell carcinoma ($n = 2801$), SCLC ($n = 1994$), and large cell carcinoma ($n = 1195$). Overall, 56.1% of all cases were observed in males and 43.9% were observed in females.