Original Article

Regional trends in diagnosis of advanced lung cancer in Michigan over 33 years

Hollis E. Hutchings[1], Sue C. Grady[2], Qiong Zhang[3], Erik Schwarze[1], Andrew Popoff[1], Kamil Khanipov[4], Ikenna C. Okereke[1]

[1]Department of Surgery, Henry Ford Health System, Detroit, MI, USA; [2]Department of Geography, Environment and Spatial Sciences, Michigan State University, East Lansing, MI, USA; [3]Department of Public Health Sciences, Henry Ford Health, Detroit, MI, USA; [4]Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, TX, USA

Contributions: (I) Conception and design: HE Hutchings, SC Grady, Q Zhang, IC Okereke; (II) Administrative support: SC Grady, Q Zhang, K Khanipov, IC Okereke; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: HE Hutchings, SC Grady, Q Zhang, K Khanipov, IC Okereke; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Ikenna C. Okereke, MD. Vice Chairman, Department of Surgery, System Director of Thoracic Surgery, Henry Ford Health System, 2799 W. Grand Blvd, Detroit, MI 48202, USA. Email: iokerek1@hfhs.org.

Background: Lung cancer is the most common cancer killer worldwide. Nearly 80 percent of lung cancers are diagnosed at advanced stages. Lack of access to medical care and undwerutilized lung cancer screening are key reasons for advanced diagnoses. We sought to understand the regional differences in presentation of lung cancer across Michigan. Utilizing a comprehensive cancer registry over 33 years, our goal was to examine associations between sociodemographic patient factors and diagnoses at advanced stages.

Methods: The Michigan Cancer Registry was queried from 1985 to 2018 to include all new diagnoses of non-small cell lung cancer (NSCLC) using International Classification of Diseases for Oncology (ICD-O) version 3 codes. NSCLC was categorized as early, regional and distant disease. Advanced disease was considered to be any disease that was regional or distant. NSCLC rates were calculated and mapped at the zip code level using the 2010 population as the denominator and spatial empirical Bayes methodology. Regional hospital service areas were constructed using travel time to treatment from the patient's zip code centroid. Logistic regression models were estimated to investigate the significance of rural *vs.* urban and travel time on level of disease at presentation. Kaplan-Meier and multivariate survival analysis was performed to evaluate the association between distance from the nearest medical center and length of survival controlling for known risk factors for lung cancer.

Results: From 1985 to 2018, there were 141,977 patients in Michigan diagnosed with NSCLC. In 1985, men were 2.2 times more likely than women to be diagnosed but by 2018 women and men developed disease at equal rates. Mean age was 67.8 years. Among all patients with known stage of disease, 72.5% of patients were diagnosed with advanced disease. Regional and distant NSCLC rates were both higher in the northern parts of the state. Longer drive times in rural regions also significantly increased the likelihood of advanced NSCLC diagnoses, in particular regional lung cancer. Patients with longer drive times also experienced overall worse survival after controlling for other factors.

Conclusions: Regional disparities exist in Michigan for diagnoses of NSCLC at advanced stages. Factors such as lack of screening in urban regions and distances to treating institutions in rural areas likely contribute to the increased likelihood of advanced NSCLC. Future interventions should target the specific needs of residents to detect disease at earlier stages and improve overall outcomes.

Keywords: Lung cancer; epidemiology; diagnosis

Submitted Feb 03, 2024. Accepted for publication Mar 29, 2024. Published online May 23, 2024.

doi: 10.21037/jtd-24-205

View this article at: https://dx.doi.org/10.21037/jtd-24-205

Introduction

Lung cancer is the most common cause of cancer-related death for men in the world, and the second most common cause of cancer-related death for women in the world (1-3). Smoking is the most common risk factor, with 80% to 90% of patients having a history of tobacco use (4). Increasingly, however, environmental factors like radon and poor air quality have been associated with an increased risk of lung cancer (5-8). The advent of low dose computed tomography (CT) scans for lung cancer screening has reduced the number of deaths from lung cancer in recent years (9). In addition, advances in chemotherapy and new immunotherapy agents have helped to improve survival (10-12).

Despite these advances, disparities persist in the incidence and mortality from lung cancer. Mortality rates are higher in males, African Americans, and those of lower socioeconomic statuses (13,14). Additionally, there are regional variations in lung cancer outcomes. In the United States, mortality is higher in the mid-South (Kentucky, Mississippi, Arkansas, and Tennessee) compared to the rest of the country (2). Patients of lower socioeconomic statuses participate less frequently in lung cancer screening, experience delays in time from initial diagnosis to treatment and experience higher mortality (15-17).

Our goal was to examine regional variations in lung cancer incidence and outcome over time. We chose to focus on a single state to perform a detailed analysis of patient specific and regional factors. Michigan is a Midwestern state characterized by several urban areas and large expanses of rural areas. The state is divided into the Lower Peninsula, which contains all of the major cities, and the remote Upper Peninsula. Most of the state's major medical institutions are located directly within the major cities. Residents in the northern part of the Lower Peninsula and the Upper Peninsula typically may drive 2–4 hours to get to the nearest large academic medical center. We present this article in accordance with the STROBE reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-205/rc).

Highlight box

Key findings

- The ratio of lung cancer incidence among men and women has largely equalized in recent years.

What is known and what is new?

- There are specific regional variations in the presentation of lung cancer.
- Longer drive times to the nearest medical center are associated with more advanced diagnoses of non-small cell lung cancer.

What is the implication, and what should change now?

- Future interventions should focus on detecting disease at earlier stages with programs such as lung cancer screening.

Methods

Patient selection

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethical approval for the study was obtained prior to data collection from the Henry Ford Health Institutional Review Board (No. IRB-16448). Individual consent for this retrospective analysis was waived.

The Michigan Cancer Surveillance Program (MCSP) is a government-funded program used to report all new cancer diagnoses in the state of Michigan. Using the Michigan Cancer Registry, which is the associated registry of the MCSP, we queried this database for all patients with a new diagnosis of non-small cell lung cancer (NSCLC) from 1985 to 2018. Demographic information, status of disease at presentation (early, regional, distant), survival data, comorbidity data and patient zip code of residence were collected for all included patients. The Surveillance, Epidemiology and End Results (SEER) program 2000 and 2018 categories were used to define early, regional, and advanced disease states. Regional NSCLC was defined as SEER categories 2, 3, 4, 5 implying "Regional by direct extension and/or lymph node involvement" and corresponded with stage II and/or stage III disease. Distant NSCLC was defined by category 7 and implied stage IV disease. Advanced NSCLC was defined as regional or distant disease. Any records with unknown or missing stage information were removed. Those removed records however, were mapped using the accompanying zip code to visualize their spatial patterns. This map did not show clusters of data on unknown or missing stage information, demonstrating that these data were more random across Michigan. Incidence and disease status at presentation were mapped at the zip code level throughout the state. Notably, smoking history was not recorded for over 80% of patients in the database. More consistent smoking data was recorded over the last 3 years of the study period. Given the inconsistent recording of smoking history, smoking status was excluded from the study.

All hospitals that existed in the state during this period were also recorded. Hospitals were included if they had capacity to diagnose and direct initial work up and treatment of lung cancer. We considered centers with computed tomography imaging capabilities and the physical presence of oncology services, such as medical oncologists, to fit this description. Medical centers with computed tomography capability without oncology services, such as urgent care centers, did not meet the above criteria and were not included. Hospital service areas defined by groupings of drive times in minutes were calculated using Michigan's street network. Patients were assigned a drive time based on their zip code of residence within the medical center's service area. The patient's residence was also assigned an urban or rural designation.

Statistical analysis

Rates of regional and distant NSCLC were calculated at the zip code level using the 2010 census population as the denominator and empirical Bayes smoothing methodology in GeoDa version 1.2 (18). The rates were mapped in ArcGIS Pro v. 3.0.4 (19). Zip codes were coded as urban *vs.* rural using the 2010 urban boundaries, defined as having 50,000 or more people (20). Rural areas encompassed all people not included in urban areas. Hospital service areas were calculated in Network Analyst ArcGIS Pro using the Michigan Street Network provided by ESRI and distance-time breakpoints of 0 to less than 15 minutes, 15 to less than 30 minutes, 30 to less than 60 minutes, 60 to less than 90 minutes and 90 to 120 minutes. Logistic regression models were estimated to investigate the direct association of rural versus urban residency and travel time on each patient's level of disease at diagnosis controlling for sex, age, African American race and year of diagnosis. These models were also stratified by stage of disease. These regression analyses were conducted using proc logistic in SAS v. 9.4 (21).

The same models were also used to measure the associations of increasing travel time to a regional medical facility for treatment, controlling for sex, age, and year of diagnosis. Survival analysis was conducted in R version 4.1.0 using Survival and ggsurvfit packages (22-24). The date of last contact was used to calculate survival. Patients alive at the date of last contact were censored. Additionally, patients who died within 30 days of their initial diagnosis were excluded from the analysis, since only the month and year of diagnosis is recorded. A total of 116,070 participants were used for the analysis. A period of 4 years was set up as a follow-up period. Patients who were first diagnosed in 2017 were last contacted in 2021. Kaplan-Meier Plots were generated for time to the nearest medical center, race, level of disease, and sex. Cox proportional-hazards model was calculated to find multivariate associations between survival time and distance to the nearest medical center, level of disease, race, and sex. The proportional hazards assumption was tested using scaled Schoenfeld residuals against the transformed time. Hazard ratios, 95% confidence levels, and P values were reported.

Propensity matching

The analysis was also repeated on a propensity-matched cohort using the MatchIt package (25). Nearest neighbor matching based on propensity score differences was performed. Participants were matched based on age at diagnosis, level of disease, race, and sex. Matching was performed using the nearest 1:1 matching with the propensity score as the distance metric. Age was taken as a continuous variable. A total of 86,570 participants remained in the analysis after propensity matching. Tables S1-S3 show a summary of the matching data.

Results

Demographics

Table 1 shows overall patient demographics. A total of 141,977 patients were diagnosed with NSCLC in Michigan during the study period. *Figure 1* shows the percentage of cases by sex over time. In 1985, men were 2.2 times more likely than women to be diagnosed but by 2018 women and men developed disease at equal rates. Mean age was 67.8 years. Overall 85.2% of patients were White and 13.6% of patients were Black. After removal of patients with unknown or missing records about stage, there were a total of 124,009 cases.

Disease status at presentation

During the study period, early disease was diagnosed in 23.3% (33,060/141,977) of patients, regional disease in 24.9% (35,380/141,977) of patients and distant disease in 39.1% (55,569/141,977) of patients (*Table 1*). Rates of advanced disease rose steadily over time (*Figure 2*).

Geographic area versus level of disease

Figure 3 details the geographic distribution of patients with advanced disease at presentation. There was a strong tendency for rates of regional disease to be highest in the more rural parts of the state, particularly in the northeast Lower Peninsula and the Upper Peninsula. *Table 2* shows the relationship of distance from the treating medical center to level of disease at presentation. A majority of NSCLC patients living in urban areas traveled less than 30 minutes to the nearest available regional hospital. The majority of similar patients in rural areas, in contrast, had a travel time of at least 60 minutes.

Table 3 shows factors that were associated with a regional or distant diagnosis. Patients living in rural areas were more likely to be diagnosed with regional disease controlling for sex, age, African American race and year of diagnosis. Rural *vs.* urban residence was not significant for advanced disease.

Table 1 Patient demographics (N=141,977)

Variable	Values
Male, n (%)	81,328 (57.3)
Age at diagnosis, years, mean	67.8
Race, n (%)	
White	120,964 (85.2)
African American	19,309 (13.6)
Hispanic	1,030 (0.7)
Other	674 (0.5)
Stage at diagnosis, n (%)	
Early	33,060 (23.3)
Regional	35,380 (24.9)
Distant	55,569 (39.1)
Unknown	17,968 (12.65)

Survival compared to distance from nearest medical center

Kaplan-Meier survival analysis of patients before and after propensity matching revealed an association of decreased survival as drive times increased from the nearest medical center (*Figure 4*). Multivariate Cox proportional hazard modeling, after controlling for sex, race and stage of disease, showed that patients living 15–60 minutes away had significantly worse survival (*Table 4*). However, patients living 60 to less than 90 minutes away from the nearest medical center did not have significantly increased risk.

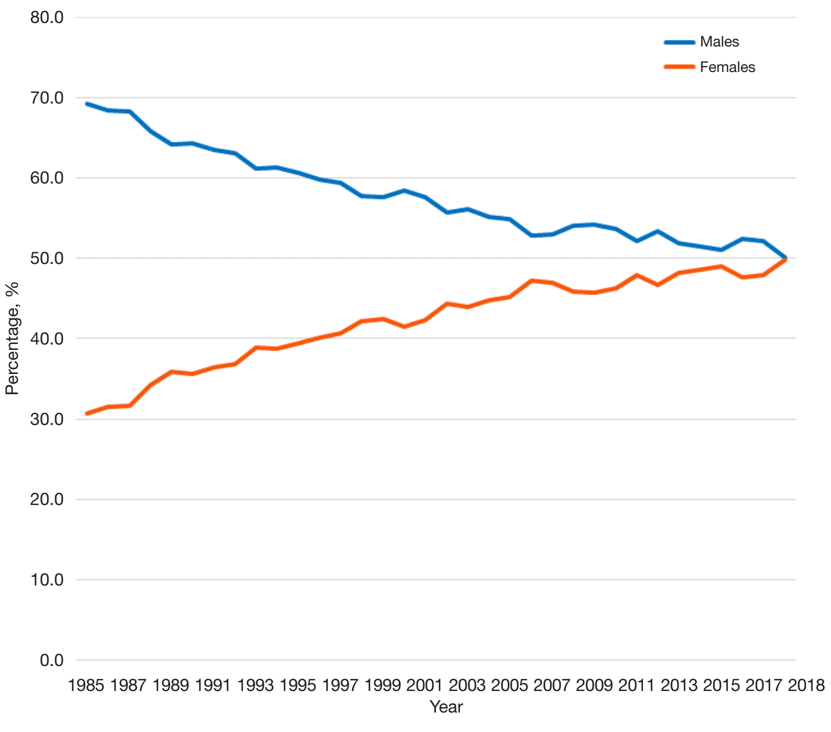


Figure 1 Percentages of NSCLC cases in men and women over time. NSCLC, non-small cell lung cancer.

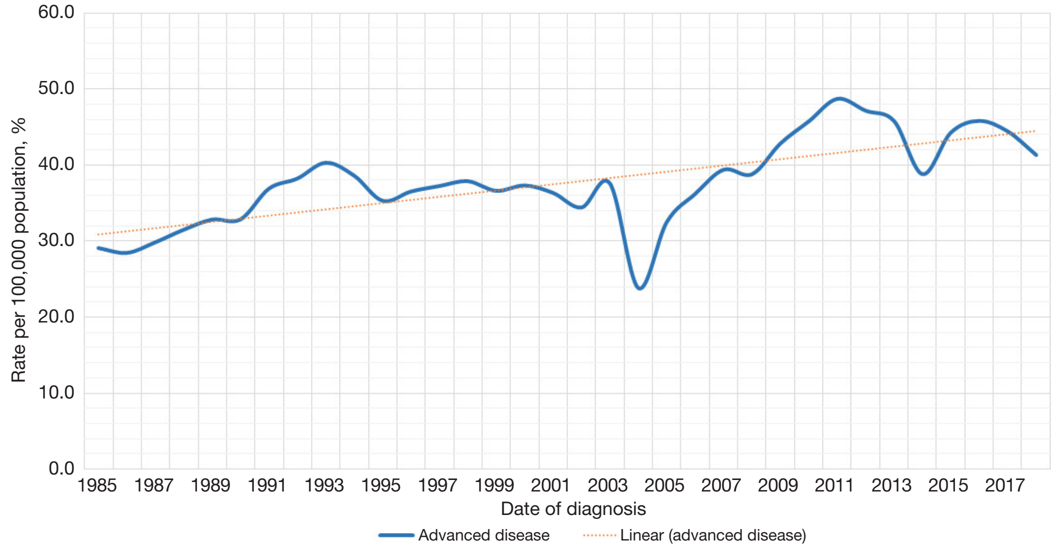


Figure 2 Rates of advanced disease over time.

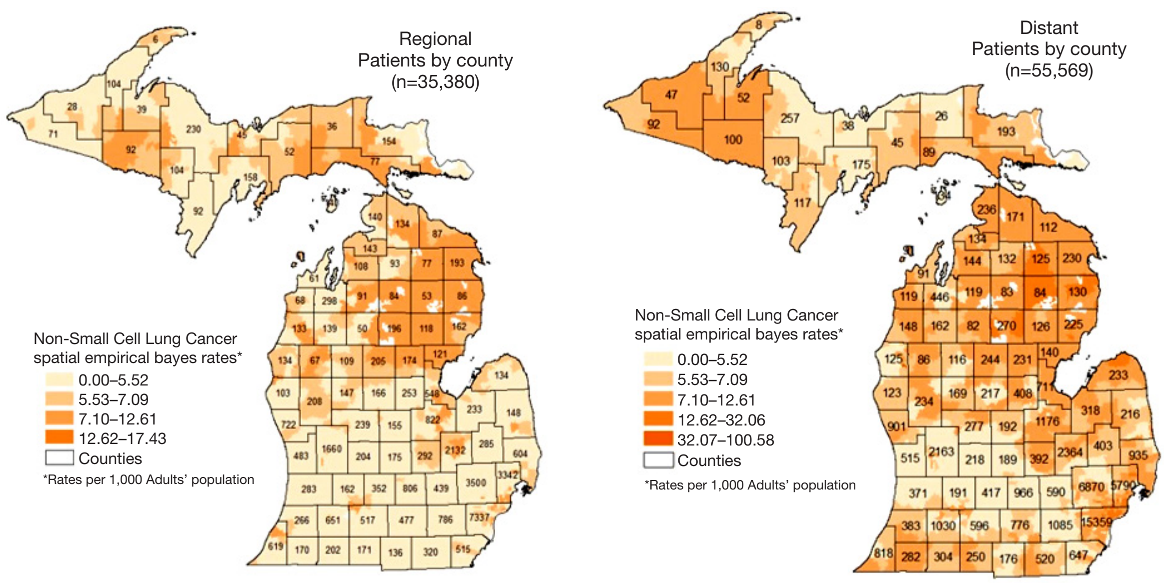


Figure 3 Geographic distribution of patients with regional and advanced disease at presentation.

Survival by sex, race and disease stage

Figure 5 shows Kaplan-Meier analyses of patients by race. Black patients had a lower overall survival probability than Hispanic and White patients. Males also had a significantly lower survival outcome probability than females.

Discussion

Our study demonstrated there are regional differences in lung cancer stage at diagnosis in the state of Michigan. We chose to use a state database for several reasons. Firstly, this database is comprehensive and every lung cancer diagnosed in our state was included. We felt that this database would not have a selection bias that other national databases may have. Secondly, our state is characterized by few urban centers and large rural areas surrounding these urban centers. Large parts of our state are quite remote and access to medical centers is limited in these areas. We felt that

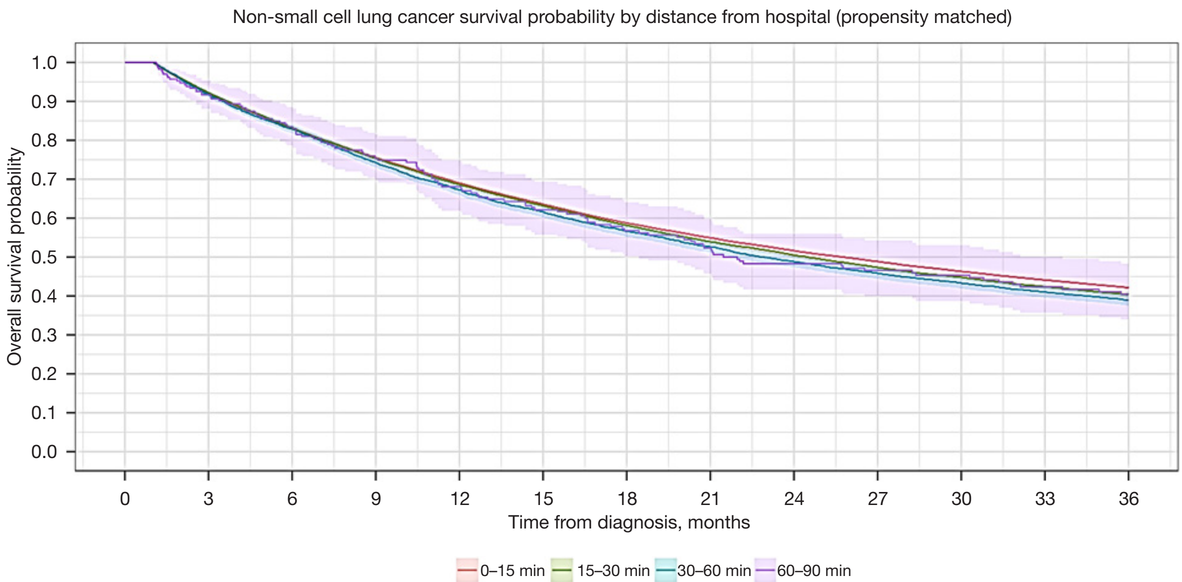
Table 2 Number of regional and distant diagnoses by drive time and rural/urban status in Michigan

Level of disease	Rural, n (%)	Urban, n (%)	Total, n (%)
Regional			
0 to <15 min	2,741 (26.19)	19,351 (79.05)	22,092 (63.22)
15 to <30 min	5,522 (52.77)	4,811 (19.65)	10,333 (29.57)
30 to <60 min	2,052 (19.61)	318 (1.30)	2,370 (6.78)
60 to <90 min	110 (1.05)	0	110 (0.31)
90–120 min	40 (0.38)	0	40 (0.11)
Total	10,465 (100.00)	24,480 (100.00)	34,945 (100.00)
Distant			
0 to <15 min	3,628 (26.13)	33,674 (82.12)	37,302 (67.95)
15 to <30 min	7,353 (52.95)	6,928 (16.89)	14,281 (26.02)
30 to <60 min	2,704 (19.47)	406 (0.99)	3,110 (5.7)
60 to <90 min	147 (1.06)	0	147 (0.27)
90–120 min	55 (0.40)	0	55 (0.10)
Total	13,887 (100.00)	41,008 (100.00)	54,895 (100.00)
Advanced stage			
0 to <15 min	6,369 (26.15)	53,025 (80.97)	59,394 (66.11)
15 to <30 min	12,875 (52.87)	11,739 (17.93)	24,614 (27.40)
30 to <60 min	4,756 (19.53)	724 (1.11)	5,480 (6.10)
60 to <90 min	257 (1.06)	0	257 (0.29)
90–120 min	95 (0.39)	0	95 (0.11)
Total	24,352 (100.00)	65,488 (100.00)	89,840 (100.00)

Table 3 Logistic regression models estimating risk factors for regional and distant diagnoses

Effect	Regional *vs.* early lung cancer			Distant *vs.* early lung cancer			Advanced *vs.* early lung cancer		
	OR	95% Wald CI	P value	OR	95% Wald CI	P value	OR	95% Wald CI	P value
Variable									
Rural *vs.* urban	1.093	1.056–1.131	<0.01	0.891	0.863–0.920	<0.01	0.969	0.941–0.998	0.30
Females	0.822	0.797–0.847	<0.01	0.800	0.778–0.822	<0.01	0.810	0.789–0.831	<0.01
Age	0.983	0.981–0.984	<0.01	0.981	0.980–0.982	<0.01	1.285	1.235–1.338	<0.01
African American	1.074	1.023–1.127	<0.01	1.422	1.363–1.483	<0.01	0.982	0.981–0.983	<0.01
Diagnosis year	1.001	0.989–1.002	0.35	1.017	1.015–1.018	<0.01	1.010	1.009–1.012	<0.01
Distance from healthcare center									
0 to <15 min	–	–	–	1.194	1.128–1264	<0.01	1.089	1.035–1.147	<0.01
15 to <30 min	1.064	1.028–1.102	<0.01	1.061	1.000–1.126	0.051	1.044	0.989–1.102	0.12
30 to <60 min	1.024	0.962–1.089	0.46	–	–	–	–	–	–
60 to <90 min	1.297	0.972–1.729	0.08	–	–	–	–	–	–
Females	0.821	0.796–0.847	<0.01	0.799	0.777–0.822	<0.01	0.908	0.789–0.830	<0.01
Age	0.983	0.981–0.984	<0.01	0.981	0.980–0.982	<0.01	0.982	0.981–0.983	<0.01
African American	1.068	1.018–1.121	<0.01	1.404	1.346–1.465	<0.01	1.274	1.224–1.326	<0.01
Diagnosis year	1.001	0.999–1.002	0.38	1.017	1.016–1.019	<0.01	1.011	1.009–1.012	<0.01

OR, odds ratio; CI, confidence interval.



	0	3	6	9	12	15	18	21	24	27	30	33	36
At risk													
0 to <15 min	56621	49951	42820	37105	32699	29149	26345	24211	22408	20937	19635	18536	17587
15 to <30 min	24178	21379	18227	15706	13797	12325	11092	10112	9320	8611	8063	7566	7126
30 to <60 min	5532	4841	4115	3515	3094	2762	2480	2269	2078	1933	1807	1692	1595
60 to <90 min	231	199	168	146	127	116	104	93	84	79	77	69	65
Events													
0 to <15 min	0	4414	9182	12988	16141	18681	20848	22532	23999	25190	26279	27209	28035
15 to <30 min	0	1870	3907	5558	6949	8016	8993	9787	10433	11009	11465	11896	12257
30 to <60 min	0	449	900	1317	1647	1903	2126	2297	2461	2587	2693	2788	2874
60 to <90 min	0	19	36	52	66	77	87	95	102	105	107	112	115

Figure 4 Kaplan-Meier survival based on drive time to nearest hospital.

utilizing this state database would allow us to map trends and analyze any differences based on urban versus rural settings. Over 75% of patients diagnosed with NSCLC presented in advanced stage disease. Patients residing in rural areas were more likely to be diagnosed with regional disease, and patients residing in urban areas were more likely to be diagnosed with distant and advanced stage disease. However, those patients living in areas with a travel time up to 30 minutes were more likely to be diagnosed with regional and distant stage disease. This demonstrates that both screening in urban areas and distance to a treatment facility in rural areas are important factors to target future interventions.

We hypothesize distance from a medical center is a surrogate marker for access to medical care. We mapped out regional hospitals that have capabilities to perform lung cancer operations as well as manage chemotherapy and radiation regimens. This does not address any free-standing primary care clinics. Further work is required to understand adherence to lung cancer screening guidelines by primary care providers, with a specific focus on screening adherence in rural areas. In northeast Michigan where regional diagnoses were elevated, counties such as Oscoda and Montgomery had very high rates of uninsured persons (17% and 13% respectively). The ratio of patients to primary care physicians is also extremely high in Presque Isle County, located in the northernmost part of the Lower Peninsula. In this county the ratio is 4,250 to 1 compared to the overall state ratio of 1,270 to 1 (26). In future studies, we would like to understand the impact of distance from major medical centers in rural areas has where patients are more likely reliant on community practice physicians for care.

In review of the literature, lung cancer screening remains underutilized. Overall, only a small percent of eligible patients actually complete lung cancer screening (27,28). When comparing urban and rural areas, 5 percent of patients in rural regions lack access. Rates of utilization are not different between urban and rural regions, however (29).

Table 4 Cox proportional-hazards model estimating survival, Michigan 1985–2018

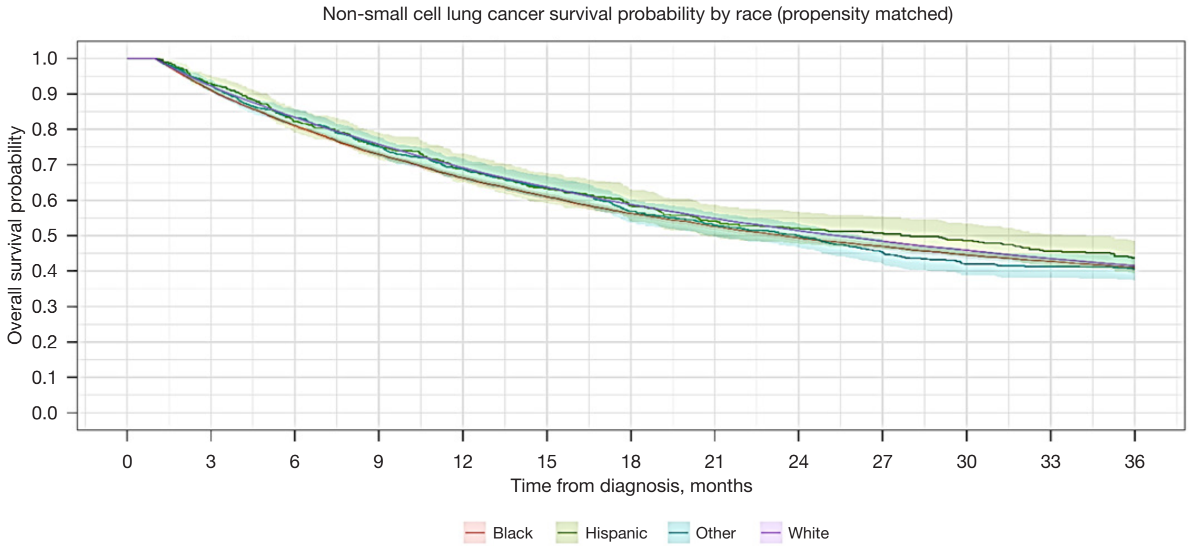
Characteristic	Before matching			After matching		
	HR	95% CI	P value	HR	95% CI	P value
Distance from healthcare center (min)						
0 to <15	–	–	–	–	–	–
15 to <30	1.03	1.01–1.05	<0.01	1.05	1.02–1.07	<0.01
30 to <60	1.07	1.03–1.10	<0.01	1.10	1.06–1.15	<0.01
60 to <90	1.00	0.87–1.15	>0.90	1.00	0.84–1.21	>0.90
Sex						
Male	–	–	–	–	–	–
Female	0.87	0.86–0.88	<0.01	0.87	0.85–0.89	<0.01
Stage						
Early	–	–	–	–	–	–
Regional	1.92	1.88–1.97	<0.01	1.89	1.84–1.93	<0.01
Distant	4.36	4.27–4.45	<0.01	2.35	2.29–2.41	<0.01
Race						
African American	–	–	–	–	–	–
Hispanic	1.09	1.00–1.19	0.053	0.96	0.85–1.08	0.50
Other	0.89	0.83–0.96	<0.01	0.97	0.89–1.06	0.50
White	1.08	1.06–1.10	<0.01	1.01	0.99–1.04	0.30

HR, hazards ratio; CI, confidence interval.

Targeted interventions to ensure access to screening and inform screening eligible patients of the benefits of lung cancer screening are required to improve lung cancer outcomes. It may be helpful to address obstacles such as transportation with travel vouchers, for example. Lung cancer screening is also very related to referrals from primary care physicians. Efforts to increase awareness among primary care physicians will also help to increase rates of lung cancer screening. It should be stressed, however, that lung cancer screening rates are very low across all demographics and should be increased broadly. The low rates of lung cancer screening alone do not explain our findings, however. Lung cancer screening was not approved by the United States Preventive Services Task Force (USPSTF) until December 2013. Most institutions in Michigan began their lung screening programs several years after this approval. Overall screening rates in Michigan are relatively low and consistent with national levels of screening. Given that our study was performed over multiple decades, the effect of screening on incidence or stage at diagnosis is likely minimal.

At the start of our study period men were more likely to be diagnosed with NSCLC. This disparity was eliminated at the end of the study period, however, as men and women were equally likely to develop lung cancer. This equalization likely occurred because the incidence of lung cancer decreased more rapidly in men than in women. In review of the literature, there are epidemiologic differences in presentations between men and women. Women tend to be diagnosed at a younger age and have a lower dose exposure to tobacco than men (29). Additionally, women tend to have a higher rate of diagnosis in never-smokers than men (30,31). There is a paucity in the scientific literature regarding trending rates of diagnosis in men and women over time, and further studies are needed to examine this relationship in other geographic regions.

This study examined the relationship between patients' spatial location and incidence of late-stage lung cancer diagnoses. We observed spatial disparities in late-stage diagnoses of NSCLC, most notably more rural parts of



At risk													
Black	11936	10440	8876	7617	6690	5951	5366	4938	4568	4296	4026	3826	3641
Hispanic	540	485	411	363	324	289	260	239	228	220	209	194	186
Other	1088	956	806	676	590	519	451	413	380	340	311	305	297
White	72998	64489	55237	47816	42113	37593	33944	31095	28714	26704	25036	23538	22249

Events													
0 to <15 min	0	1054	2172	3036	3723	4251	4707	5049	5354	5571	5795	5957	6111
15 to <30 min	0	37	92	126	156	182	204	223	232	238	247	260	268
30 to <60 min	0	80	172	249	304	347	401	432	455	491	515	520	525
60 to <90 min	0	5581	11589	16504	20620	23897	26742	29007	30954	32591	33987	35268	36377

Figure 5 Kaplan-Meier survival based on race.

the state having a higher incidence of late-stage diagnoses. This is consistent with the literature demonstrating spatial variations in cancer outcomes, with a predominance toward rural areas (2). Concerning survival, patients living 15 to less than 60 minutes from a medical center had significantly increased risk of mortality. Patients living 60 to less than 90 minutes from a hospital, however, did not experience increased risk. The paradox is likely due in part to the lack of individual-level data for patients and treating institutions from our database. Additionally, we did not examine patient level socioeconomic factors that likely contribute to increased mortality in more rural areas. Patients in more rural areas may have differences in education level compared to similar patients in urban regions which can impact compliance to lung cancer screening guidelines. Political affiliations, primary payer and insurance status were not captured in this dataset, but may be important in understanding the disparity in mortality experienced for patients in more rural regions of the state compared to urban regions.

We used zip code level data and mapped to all regional treating hospitals. In the future, recording a database with the hospital at which a patient is treated would strengthen this correlation to allow for more robust analysis. Despite this limitation, our data reproduced what has been demonstrated in the literature that men have an overall worse survival than females, and Black patients had an overall worse survival than Whites. Continued work is needed to understand these disparities better and eliminate modifiable factors contributing to worse survival in specific patient populations.

As our study was conducted over a very long time period, it is important to note the changes in treatment patterns that has occurred. During the later part of the study more targeted treatments became available. In addition, novel techniques such as minimally invasive surgery have continually increased in prevalence and have allowed for resection with less morbidity. Despite these advances, there still may be a stigma against lung cancer being "brought upon oneself" by a modifiable behavior like smoking.

This stigma may dampen referral patterns to appropriate treating institutions, particularly for residents in rural or disadvantaged communities.

In a population-based study on smoking among patients with lung cancer in 7 states (Alaska, Colorado, Florida, Idaho, Louisiana, North Carolina, and Rhode Island), investigators observed 87.5%, 36.7% and 50.8% of lung cancer patients reported ever, current, or former smoking history, respectively (32). These states were selected because their patient data on lung cancer had less than 15% of cases with unknown smoking status. The findings varied for men and women who ever smoked (90.4% *vs.* 84.3%), currently smoked (37.5% *vs.* 35.8%) and formerly smoked (52.8% *vs.* 48.5%). These findings were used to inform our study.

This study has limitations. Data was collected retrospectively through a state-managed database. We relied on the completeness of the database for our conclusions. But some key variables, such as smoking, were missing in most patients making those conclusions weaker. This study however, was informed by the previously noted study that demonstrated that a high percentage of patients with lung cancer had a smoking history (32). We need to emphasize the importance of complete data collection for better database completion and completion of future studies. Also, there were approximately 17,000 patients that were excluded from analysis due to incomplete data. A prior study has suggested propensity matching for missing data, but weigh the advantages and disadvantages of this approach (33). In addition, this was a retrospective study looking only at the state of Michigan. Future studies using a national cancer database may offer further insights to more national geographic disparities. As mentioned, the association between drive time to a medical center and survival was determined using patient zip code level data and not specific treating institution. In the future, utilizing treating institution may be more useful in determining accessibility in rural areas. Since patient-level treating hospital data was unavailable, zip-code level allowed for an analysis of access. This lack of data may have led to the paradox we witnessed regarding survival for patients who lived 60 to 90 minutes away from a treating institution. Future studies examining patient distance to their specific treating institution would be helpful to understand if this plays a role in survival.

Conclusions

A statewide analysis of lung cancer trends over 4 decades revealed shifting trends in lung cancer incidence. Female patients now develop disease as frequently as male patients. Patients in rural areas tend to develop advanced disease more commonly than patients in urban areas. Survival is independently associated with drive time to the nearest institution. Future interventions should improve accessibility and lung cancer screening particularly in rural and underserved areas.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-205/rc

Data Sharing Statement: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-205/dss

Peer Review File: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-205/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-205/coif). I.C.O. serves as an unpaid editorial board member of *Journal of Thoracic Disease* from February 2023 to January 2025. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethical approval for the study was obtained prior to data collection from the Henry Ford Health Institutional Review Board (IRB-16448). Individual consent for this retrospective analysis was waived.

References

1. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. Ann Glob Health 2019;85:8.
2. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol 2016;893:1-19.
3. Henley SJ, Gallaway S, Singh SD, et al. Lung Cancer Among Women in the United States. J Womens Health (Larchmt) 2018;27:1307-16.
4. Walser T, Cui X, Yanagawa J, et al. Smoking and lung cancer: the role of inflammation. Proc Am Thorac Soc 2008;5:811-5.
5. Yang D, Liu Y, Bai C, et al. Epidemiology of lung cancer and lung cancer screening programs in China and the United States. Cancer Lett 2020;468:82-7.
6. Lorenzo-González M, Torres-Durán M, Barbosa-Lorenzo R, et al. Radon exposure: a major cause of lung cancer. Expert Rev Respir Med 2019;13:839-50.
7. Hosgood HD, Cai Q, Hua X, et al. Variation in oral microbiome is associated with future risk of lung cancer among never-smokers. Thorax 2021;76:256-63.
8. Warren GW, Cummings KM. Tobacco and lung cancer: risks, trends, and outcomes in patients with cancer. Am Soc Clin Oncol Educ Book 2013;359-64.
9. Becker N, Motsch E, Trotter A, et al. Lung cancer mortality reduction by LDCT screening-Results from the randomized German LUSI trial. Int J Cancer 2020;146:1503-13.
10. Howlader N, Forjaz G, Mooradian MJ, et al. The Effect of Advances in Lung-Cancer Treatment on Population Mortality. N Engl J Med 2020;383:640-9.
11. Reck M, Heigener DF, Mok T, et al. Management of non-small-cell lung cancer: recent developments. Lancet 2013;382:709-19.
12. Elci OC, Akpinar-Elci M. The trend of small cell lung cancer among young men. Lung Cancer 2007;57:34-6.
13. Hovanec J, Siemiatycki J, Conway DI, et al. Lung cancer and socioeconomic status in a pooled analysis of case-control studies. PLoS One 2018;13:e0192999.
14. Sin MK. Lung cancer disparities and African-Americans. Public Health Nurs 2017;34:359-62.
15. Shah M, Parmar A, Chan KKW. Socioeconomic disparity trends in diagnostic imaging, treatments, and survival for non-small cell lung cancer 2007-2016. Cancer Med 2020;9:3407-16.
16. Wong ML, Clarke CA, Yang J, et al. Incidence of non-small-cell lung cancer among California Hispanics according to neighborhood socioeconomic status. J Thorac Oncol 2013;8:287-94.
17. Kawaguchi T, Matsumura A, Fukai S, et al. Japanese ethnicity compared with Caucasian ethnicity and never-smoking status are independent favorable prognostic factors for overall survival in non-small cell lung cancer: a collaborative epidemiologic study of the National Hospital Organization Study Group for Lung Cancer (NHSGLC) in Japan and a Southern California Regional Cancer Registry databases. J Thorac Oncol 2010;5:1001-10.
18. GeoDa Center. 2023. Retrieved on January 10, 2023. Available online: https://geodacenter.github.io/
19. Environmental Research Systems, Inc. (ESRI). ArcGIS Pro v. 3.0.4. Retrieved on December 2, 2022. Available online: https://esri.com
20. U.S. Bureau of the Census. 2010 Census Urban and Rural Classification and Urban Area Criteria. Retrieved on March 1, 2023. Available online: https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html
21. SAS Institute. 2023. Retrieved on December 3, 2022. Available online: https://www.sas.com/en_us/home.html
22. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016. Available online: http://www.R-project.org/
23. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer, 2000.
24. Daniel D. Sjoberg, Mark Baillie, Steven Haesendonckx and Tim Treis (2022). ggsurvfit: Flexible Time-to-Event Figures. R package version 0.2.1. Available online: https://CRAN.R-project.org/package=ggsurvfit
25. Ho D, Imai K, King G, Stuart E (2011). "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." _Journal of Statistical Software_, *42*(8), 1-28. Available online: https://doi.org/10.18637/jss.v042.i08
26. University of Michigan, Michigan Poverty and Wellbeing Map. Accessed November 1, 2023. Available online: https://poverty.umich.edu/research-funding-opportunities/data-tools/michigan-poverty-well-being-map/
27. Zgodic A, Zahnd WE, Advani S, et al. Low-dose CT lung cancer screening uptake: A rural-urban comparison. J Rural Health 2022;38:40-53.
28. Hasson RM, Phillips JD, Fay KA, et al. Lung Cancer Screening in a Surgical Lung Cancer Population: Analysis of a Rural, Quaternary, Academic Experience. J Surg Res

2021;262:14-20.
29. Sahar L, Douangchai Wills VL, Liu KKA, et al. Geographic access to lung cancer screening among eligible adults living in rural and urban environments in the United States. Cancer 2022;128:1584-94.
30. Mederos N, Friedlaender A, Peters S, et al. Gender-specific aspects of epidemiology, molecular genetics and outcome: lung cancer. ESMO Open 2020;5:e000796.
31. Stapelfeld C, Dammann C, Maser E. Sex-specificity in lung cancer risk. Int J Cancer 2020;146:2376-82.
32. Siegel DA, Fedewa SA, Henley SJ, et al. Proportion of Never Smokers Among Men and Women With Lung Cancer in 7 US States. JAMA Oncol 2021;7:302-4.
33. Shiba K, Kawahara T. Using Propensity Scores for Causal Inference: Pitfalls and Tips. J Epidemiol 2021;31:457-63.

Cite this article as: Hutchings HE, Grady SC, Zhang Q, Schwarze E, Popoff A, Khanipov K, Okereke IC. Regional trends in diagnosis of advanced lung cancer in Michigan over 33 years. J Thorac Dis 2024;16(5):2936-2947. doi: 10.21037/jtd-24-205

