Spatial and spatio-temporal clusters of lung cancer incidence by stage of disease in Michigan, United States 1985-2018

Qiong Zhang,[1] Shangrui Zhu,[2] Sue C. Grady,[2] Anqi Wang,[1] Hollis Hutchings,[3] Jessica Cox,[4] Andrew Popoff,[3] Ikenna Okereke[3]

[1]*Department of Public Health Sciences, Henry Ford Health, Detroit, Michigan;* [2]*Department of Geography, Environment and Spatial Sciences, Michigan State University, East Lansing, Michigan;* [3]*Department of Surgery, Henry Ford Health, Detroit, Michigan;* [4]*School of Medicine, University of Texas Medical Branch, Galveston, Texas, USA*

Abstract

Lung cancer is the most common cause of cancer-related death in Michigan. Most patients are diagnosed at advanced stages of the disease. There is a need to detect clusters of lung cancer incidence over time, to generate new hypotheses about causation and identify high-risk areas for screening and treatment. The Michigan Cancer Surveillance database of individual lung cancer cases, 1985 to 2018 was used for this study. Spatial and spatio-temporal clusters of lung cancer and level of disease (localized, regional and distant) were detected using discrete Poisson spatial scan statistics at the zip code level over the study time period. The approach detected cancer clusters in cities such as Battle Creek, Sterling Heights and St. Clair County that occurred prior to year 2000 but not afterwards. In the northern area of the lower peninsula and the upper peninsula clusters of late-stage lung cancer emerged after year 2000. In Otter Lake Township and southwest Detroit, late-stage lung cancer clusters persisted. Public and patient education about lung cancer screening programs must remain a health priority in order to optimize lung cancer surveillance. Interventions should also involve programs such as telemedicine to reduce advanced stage disease in remote areas. In cities such as Detroit, residents often live near industry that emits air pollutants. Future research should therefore, continue to focus on the geography of lung cancer to uncover place-based risks and in response, the need for screening and health care services.

Correspondence: Sue C. Grady, Department of Geography, Environment, and Spatial Sciences, Michigan State University, USA.
Tel.: +1.517432-9998 - Fax:+1.5174321671.
E-mail: gradys@msu.edu

Key words: lung cancer, lung cancer stages, lung cancer clusters, SaTScan, Michigan, USA.

Conflict of interest: the authors declare no potential conflict of interest, and all authors confirm accuracy.

Availability of data and materials: all data generated or analyzed during this study are included in this published article.

Received: 30 June 2023.
Accepted: 15 January 2024.

Geospatial Health 2024; 19:1219
doi:10.4081/gh.2024.1219

Introduction

Lung cancer is the second most common cancer worldwide with more than 2.2 million new cases of lung cancer in 2020 and with an age-standardized rate (ASR) of 22.4 per 100,000 population at risk (World Cancer Research Fund International, 2022). The ten countries with the highest ASR are Hungary (50.1), Serbia (47.3), France, New Caledonia (42.9), French Polynesia (40.4), Turkey (40.0), Montenegro (39.7), Belgium (38.3), Bosnia-Herzegovina (37.8), North Korea (37.0) and Denmark (36.8) (World Cancer Research Fund International, 2022). The World Health Organization's (WHO) Global Burden of Disease estimates show the lung cancer ASR for the United States (35.1) in 2020 was slightly higher than that of the United Kingdom and Northern Ireland (32.5). Lung cancer mortality ranged from 42.4 per 100.000 in Hungary to 25.8 in Germany (Eldridge, 2022).

Smoking is the leading risk factor for development of lung cancer but people are also at an increased risk if they have previous lung disease, experienced harmful occupational exposures, or exposed to indoor and/or outdoor air pollution (Lee *et al.*, 2011; Alberg *et al.*, 2013; Pui *et al.*, 2014; Zhang *et al.*, 2015; de Groot *et al.*, 2018). Lung cancers typically start in the cells lining the bronchi and parenchyma of the lung. Non-small cell lung cancer (NSCLC) comprises 80% to 85% of all lung cancer diagnoses and includes adenocarcinoma, squamous cell carcinoma and large cell carcinoma, while small cell lung cancer (SCLC) comprises only 10% to 15% of all lung cancers diagnoses and tends to be more aggressive but responds well to treatment if diagnosis occurs early (American Cancer Society, 2023b).

In U.S., the year 2023 is expected to add an estimated 238,234 new cases of lung cancer despite a recent 5-year (2015-2019) declining trend in incidence (ASR = 56.3, trend -2.6, 95% CI -3.5, -1.8) (American Cancer Society, 2023a). Previous studies in the U.S. (2016-2020) have shown that lung cancer rates are higher in males (61.1 per 100,000 population) compared to females (48.6) although this gap varies by race and ethnicity, non-Hispanic African American males (71.4), non-Hispanic white males (64.4),

non-Hispanic American Indian and Alaskan Native males (56.4) and non-Hispanic white females (53.9) having the highest rates. Overall, the gender gap has narrowed over time (Centers for Disease Control and Prevention, United States Cancer Statistics 2023) in part because women started smoking in large numbers later than men and have also been slower to quit smoking (Thun *et al.*, 2013). The lifetime risk of developing lung cancer is about 6.2% for men and 5.8% for women, *i.e.* 1 in 16 men and 1 in 17 women are expected to develop lung cancer during their lifetime and this risk is influenced by smoking history (American Cancer Society, 2023a). Lung cancer develops over many years before symptoms appear, therefore most (83%) patients are diagnosed at ages 65 years and older (American Cancer Society, 2023a). There are also social determinants of risk, in particular educational levels, ranging from 166.6 per 100,000 in men without a high school diploma to 57.6 with a college degree (Torre *et al.*, 2016). Higher education levels are associated with higher incomes and lower tobacco exposure (Wang *et al.*, 2018), likely explaining part of the lower incidence of lung cancer. In contrast, people of lower socioeconomic statuses tend to have less access to health care and may be diagnosed at more advanced stages (Barta *et al.*, 2019; Van der Heyden *et al.*, 2009; Walters *et al.*, 2013; Goss *et al.*, 2014; Siegel *et al.*, 2018). Finally, survival rates for lung cancer are highly dependent on the stage of diagnosis with lower survival rates likely reflecting less access to early detection, curative-intent surgeries and new therapies (Landsdorp-Vogelaar *et al.*, 2012; Javid *et al.*, 2014; Lovly, 2022). Public education among the general population and patient education and screening for lung cancer among high-risk groups are thus very important interventions for the early detection and survival of lung cancer.

In 2018 the rate of new lung cancer cases (59.7 per 100,000) in Michigan was higher than the national rate (54.9) using a direct method of standardization and the 2000 U.S. standard population despite a dramatic decline in lung cancer incidence since 1985 (84.1 per 100,000 population at risk) and recent (2015-2019) 5-year trend (-1.6, 95% CI -1.8, -1.4). In Michigan, the percentage of all patients diagnosed with lung cancer at an early (localized) stage (26.6%) was lower than that of the national rate (28.8%). In contrast, the percentage of patients diagnosed with lung cancer at a late stage in Michigan, i.e. regional (23.8%) or distant (49.6%), were higher than that of the national rates (regional 23.7%) or distant (47.6%). Late-stage lung cancer also varies by county (2017-2019) ranging from 47.6% to 79.1%. While most research on lung cancer in Michigan has been population-based, there is a need to further understand the spatial, temporal, and spatio-temporal patterns of lung cancer at finer resolutions. Understanding these patterns will help to generate new hypotheses about risk factors for lung cancer - *e.g.*, environmental pollutants and reasons underlying the high proportion of late diagnoses as well as target areas for early screening and treatment interventions. Detection of clusters is important for disease surveillance and spatial epidemiological research (Torabi & Galloway, 2018). A purely spatial cluster is defined as the unusual aggregation of case counts within certain group(s) of people in a geographic area. A space-time cluster is the same aggregation but in a geographic area over a specific time period (Lawson, 2006; Amin *et al.*, 2010). This study aimed to detect spatial and spatio-temporal clusters of lung cancer overall and for each of the three stages of disease (localized, regional or distant) at time of diagnosis, with findings intended to inform future epidemiological and health system studies and public health and health care interventions.

Materials and Methods

Study area

Michigan, a U.S. state located in the Midwest Region in Health and Human Service Region 5 (Figure 1), served as the study area. This study was conducted at the zip code level across 83 counties and 10 Prosperity Regions. Prosperity Regions are economic regions with an array of services distributed within boundaries used as reference when describing lung cancer clusters as well as when generating new hypotheses about processes underlying the spatial patterns found.

Data

The Michigan Cancer Surveillance Program (MCSP) database is a state government-funded surveillance program that collects demographic and clinical information on all patients diagnosed with cancer at a participating treating institution. This study utilized the MCSP data on newly diagnosed lung cancer patients between 1985 and 2018 (n=253,572). Lung cancer diagnoses were comprised of the following International Classification of Disease (ICD-10) codes: C34.00, C34.01, C34.02 and C34. Demographic information collected for analysis included sex, age and zip code of residence. Tumour information was categorized according to the program surveillance, epidemiology and end results (SEER). According to SEER categorization, localized disease was SEER 1. Regional disease was SEER 2, 3, 4 and 5. Distant disease was SEER 7. Advanced disease (late stage) was defined as the presence of either regional or distant disease. Counts of lung cancer cases overall and by stage were summed by zip code (residential locations of patients at time of diagnosis) as the numerator for subsequent cluster detection of high relative risk.

Population counts at the zip code level were used to calculate expected lung cancer counts as outlined below for the year 2010 (Summary File 1, decennial census) to identify purely spatial clusters. Population counts were used for the years 1980, 1990, 2000, 2010, and 2020 (Summary File 1 decennial census) for the detection of temporal, spatial and spatio-temporal clusters. Since all

Figure 1. Reference map of the State of Michigan by the Health and Human Service Regions in the United States.

patients diagnosed with lung cancer within the Michigan Cancer Surveillance database were adults, the population data utilized in this study were referred to people >18 years of age. Since some zip codes changed over time, the zip code map for 2010 was used as the baseline from which all other zip codes were matched. Zip codes that did not match were re-coded by searching their historical locations in Google Maps and assigning a timely zip code for that area. An online global positioning system (GPS) program was used to search for 336 such zip codes. Of the total, 16 zip codes were recoded. The remaining 320 zip codes were either not found (305) or removed (15). A population threshold of 50 residents was used in this study to ensure a stable denominator. The 320 records removed from the dataset represented less than 0.5% of the patient data (n=2,887 patient records). The counties of these missing records were queried, and it was found that the missing data were not clustered in space or time and would thus not impact subsequent spatial, temporal and spatio-temporal cluster detection. The final dataset comprised a total of n=250,685 lung cancer cases for analysis.

Spatial cluster detection

SaTScan (v.9.7) software by Martin Kulldorf (2023) was used to detect spatial, temporal and spatio-temporal clusters of lung cancer relative risk (RR) at the zip code level in Michigan from 1985 to 2016. SaTScan was used because of its functionality to detect clusters from within points patterns (centroids of zip codes) rather than area-level clusters, e.g., using spatial autocorrelation cluster detection techniques and its ability to model lung cancer (relatively rare numbers of cases in many underlying zip code level populations) using a Poisson modelling process. First, a retrospective discrete Poisson model was implemented to detect areas with high lung cancer RR = observed/expected lung cancer cases, assuming that the number of lung cancer cases in each zip code was Poisson-distributed. Under the null hypothesis, the expected number of lung cancer cases was assumed to be similar to that of the population size. As noted above, the population for the purely spatial scan was the 2010 adult population within each zip code:

$$E[c] = p^* C/P \quad \text{Eq. 1}$$

where c was the observed number of lung cancer cases; p the number of adult population in the zip code of interest; C the total number of lung cancer cases; and P the adult population in Michigan. The RR was derived by dividing the observed cases by the expected number of cases. The alternative hypothesis is that there was an elevated risk within the window compared to the outside. Under the Poisson assumption, the likelihood function for a specific window was proportional to:

$$\left(\frac{c}{E[c]}\right)^c \left(\frac{C-c}{C-E[c]}\right)^{C-c} I0 \quad \text{Eq. 2}$$

where C was the total number of lung cancer cases; c the observed number of cases within the window; and E [c] the adjusted expected number of cases within the window under the null hypothesis. C-E [c] was the expected number of lung cancer cases outside the window, and I() an indicator function. When SaTScan scans for clusters with high rates, I() was equal to 1 when the window had more cases than expected under the null hypothesis, and otherwise zero. This study focused on the high risk of lung cancer (in contrast to low risk). Hypothesis testing was conducted using 999 Monte Carlo simulations. Finally, a test statistic was calculated for each random replication and the real dataset (herein defined by the log-likelihood ratio). If the latter was among the 5% highest, then the test was significant at the 0.05 level. The significance level of clusters is herein defined by their log-likelihood ratio and p-values. In this study, SaTScan scanned for geographic sizes that would capture between zero and 3% of the adult population at risk for lung cancer. A 3% threshold was used to ensure the sensitive capture of residents <50 years of age who are at lower-risk for lung cancer, compared to older residents residing in the same area. Here, there were no geographic overlap of clusters.

Temporal cluster detection

Temporal clusters of overall lung cancer and lung cancer defined as localized, regional and distant were estimated across Michigan for the time period 1985 to 2020 using a retrospective purely temporal Poisson process and 1-year time aggregation (cluster range, 1-5 years). Here the total number of lung cancer cases was summed and the population groups averaged over the study time period. Years with significantly high lung cancer rates were reported.

Spatial-temporal cluster detection

This kind of scan explored the geographical distribution of lung cancer risk clusters from 1985 to 2020. The RR values were calculated based on 1990, 2000, 2010 and 2020 adult populations, corresponding to incidence data from 1985-1994, 1995-2004, 2005-2014 and 2015-2018. Considering that different criteria for reporting high RR clusters may result in different cluster collections, we conducted a sensitivity analysis by applying 10, 30, 40, 50, and 90 percent of the study period and 3, 5, and 10 years as the maximum temporal cluster size. The boundaries for Michigan's counties (n=83) and Prosperity Regions (n=10) were overlaid onto the spatial and spatio-temporal cluster maps for reference and to help generate new hypotheses about lung cancer.

Results

Spatial clusters of lung cancer

Figure 2 shows 31 significant spatial clusters of lung cancer across the time period 1985 to 2018.The most significant cluster was in Region 10, encompassing the city of Detroit, with a RR=1.74 (n=11,521 patients). There were also highly significant clusters in the city of Battle Creek in Calhoun County Region 8, with a RR=2.44 (n=1,296 patients) and Mount Clemens in Macomb County Region 10, with a RR=2.0 (n=859 patients). Other significant clusters with lower RRs but high patient-counts were in south-western Detroit, with a RR=1.6 (n=10,046 patients), west Detroit, RR=1.39 (n=9,000 patients) and Sterling Heights in southern Macomb County, RR=1.52 (n=11,173 patients). There were also significant spatial clusters with high case counts in rural locations, in particular in the northern part of the lower peninsula (Regions 2 and 3) and the central and eastern parts of the upper peninsula (Regions 1b and 1c). Supplemental Table 1S shows the RR and patient count for each of the 31 purely spatial clusters for lung cancer in Michigan.

Temporal clusters of lung cancer

Across the study's time period the rate of localized lung cancer was 18.6 per 100,000 followed by regional lung cancer (22.0) and distant lung cancer (41.1). These findings of temporal clusters are consistent with the space-time clusters of lung cancer across the study time period as outlined below. The first temporal cluster that emerged was regional lung cancer (25.1) occurring between 1993 to 1997 (RR = 1.17, n=8,960). Thereafter, a significant cluster of distant lung cancer (52.9) emerged between 2009 and 2013 (RR = 1.36, n=19,945). Most recently between 2015 to 2017 there was a significant cluster of localized lung cancer (23.4) (RR = 1.29, n=5,536). These findings show that distant lung cancer remains the highest form of lung cancer diagnosed while also suggesting some earlier detection of localized lung cancer in more recent years.

Spatio-temporal clusters of lung cancer

Figure 3 shows 31 significant space-time clusters of lung cancer in Michigan between 1985 and 2018. Of these, 17 (61.3%) clusters were detected prior to year 2000 including a cluster encompassing Detroit (RR=1.64, n=3,762 patients) 1985-1994, Battle Creek (RR=3.37, n=545 patients) 1988-1997, Sterling Heights (RR=1.55, n=3,480 patients) 1989-1998 and St. Clair County (RR=1.75, n=695 patients) 1986-1990. Importantly, after year 2000, the most significant clusters were in rural areas in the northern part of the lower peninsula in Regions 3 and 5 (RR=1.7, n=3,967 patients) 2008-2017; Regions 2 and 4 (RR=1.35, n=2,902 patients) 2008-2017 and the upper peninsula Region 1 (RR=1.36, n=1,347 patients) 2010-2018. There were two clusters detected in the city of Detroit (RR=1.28, n=1,673 patients) 2004-2013 and south-western Detroit (RR= =1.61, n=3,180 patients) with the later cluster beginning in 1999 and ending in 2008. There was also a significant cluster in Lapeer County (Region 6) (RR=1.35, n=3,001 patients) 2009-2018 near Otter Lake. Other clusters of lung cancer detected after 2000 were in specific zip codes of elevated risk. Supplemental Table 2S shows the RR and patient count for all 31 space-time clusters for lung cancer (before and after year 2000) in Michigan.

Spatio-temporal clusters by stage of disease

The prevalence of *localized* lung cancer was 18.6 per 100,000 population. Figure 4a shows 20 significant space-time clusters of localized lung cancer. Eleven (55.0%) of these clusters occurred prior to year 2000including a cluster in Battle Creek, which was the most significant cluster (RR=5.36, n=160 patients) followed by Jackson in Calhoun County Region 8 (RR=2.26, n= 531 patients). There were also rural clusters prior to year 2000 in Region 4 and the upper peninsula (Region 1b). In Region 3 there was an expansive rural cluster beginning in 1996 and ending in 2005 (RR=1.74, n=665 patients). After 2000, there were 9 significant clusters of

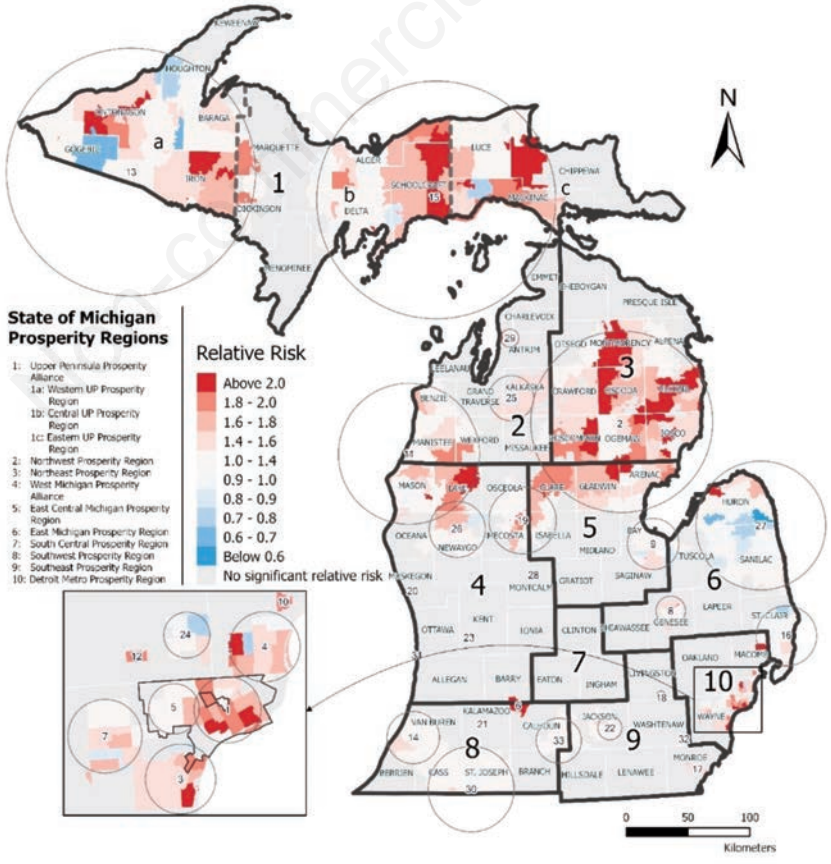

Figure 2. All lung cancer patients: discrete Poisson purely spatial clusters in relation to Michigan zip codes 1985-2018. Data Source: The Michigan Cancer Registry ICD-O version 3 codes 1985-2018.

localized lung cancer (Table 1) including the emergence of clusters in rural Cheboygan in Region 3 (RR=2.21, n=62 patients) 2001-2010 and Mecosta County in Region 4 (RR=1.84, n=131 patients) 2012-2017. There were also three clusters that emerged in Region 6 (n=1,081 total patients) and three clusters within localized areas outside of Detroit in Region 10 (n=1,109 total patients). Supplemental Table 3S shows the RR and patient count for all 20 space-time clusters for localized lung cancer (before and after the year 2000) in Michigan. The prevalence of *regional lung cancer* was 22.0 per 100,000 population. Figure 4b shows 22 significant space-time clusters of regional lung cancer. Of these, 10 clusters were detected prior to year 2000 including those in Battle Creek, Detroit, Sterling Heights, St. Clair and the upper peninsula. Another 9 clusters originated prior to year 2000 and ended after year 2000 including those in the upper peninsula in Regions 1b and 1c, Sterling Heights in Macomb County and south-western Detroit in Region 10 and new sites emerging in the city of Saginaw in Region 6, south Grand Rapids and Muskegon and Newago in Region 4 and Kalamazoo and Calhoun and Branch Counties in Region 8. Only three of these clusters occurred after year 2000. The most expansive space-time cluster of regional lung cancer after 2000 was in the rural areas of Region 3 (RR=1.85, n=930 patients) 2000-2009 followed by the cities of Flint and Otter Lake Township in Region 6 (RR=1.69, n=813 patients) during the same time period. Mecosta County in Region 4 also contained a cluster of regional lung cancer (RR=1.88, n=119 patients) 2004-2013. Supplemental Table 4S shows the RR and patient count for all 22 space-time clusters for localized lung cancer (before and after year 2000) in Michigan.

The prevalence of *distant lung cancer* was 41.1 per 100,000 population. Figure 4c shows 24 significant space-time clusters of distant lung cancer. Of these, 5 clusters were detected before year 2000 in Detroit, Battle Creek, St. Clair, Livingston and Kalamazoo. There were 2 clusters that originated prior to year 2000 and ended after year 2000 in Sterling Heights and southern Oakland County. There were 17 (70.8%) clusters detected after year 2000. The most significant cluster of distant lung cancer after the year 2000 was in south-western Detroit (RR=2.03, n=1,677 patients) 2005-2014. There was another cluster in east Detroit (RR=1.75, n=1,184 patients) 2006-2014. All other clusters of distant lung cancer in Region 10 were in the surrounding areas of Detroit. The second most significant cluster of distant lung cancer was in Region 3 (RR=1.91, n=1,806 patients) 2004-2013. There was also a significant cluster of distant lung cancer in the upper peninsula in Region 1b and 1c that included Cheboygan in Region 3 (RR = 1.23, n=991 patients) 2009-2017. Other clusters of distant lung cancer were in Region 6 including the cities of Saginaw and Otter Lake and Region 4 Mecosta County and its surrounding areas. Supplemental Table 5S shows the RR and patient count for all 24 space-time clusters for localized lung cancer (before and after year 2000) in Michigan.

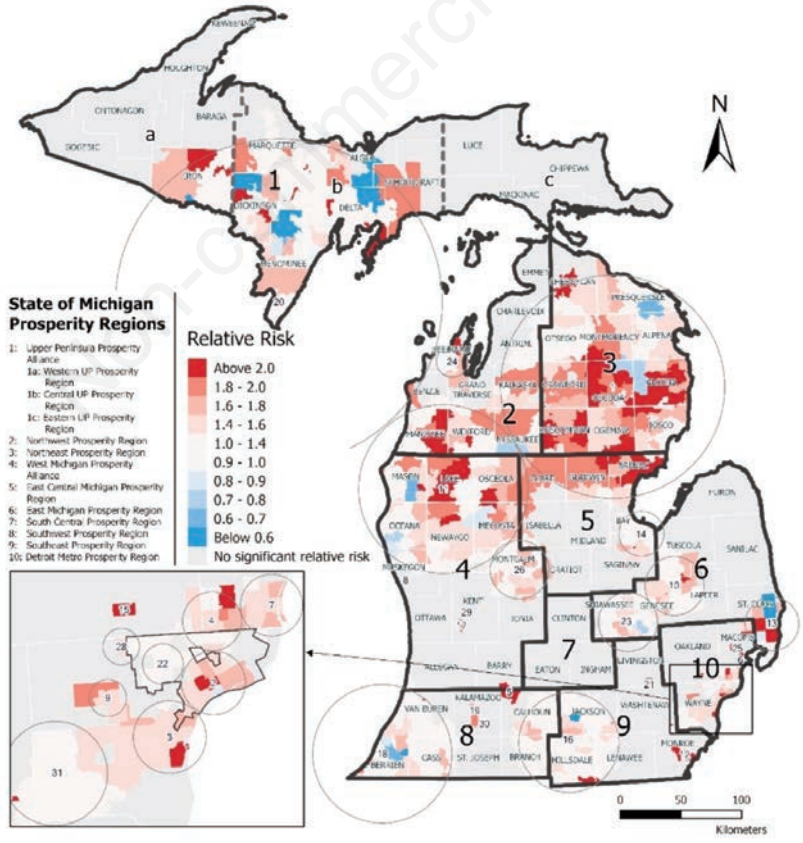

Figure 3. All lung cancer patients: discrete Poisson purely spatio-temporal clusters in relation to Michigan zip codes 1985-2018. Data Source: The Michigan Cancer Registry ICD-O version 3 codes 1985-2018.

Discussion

This study does not attempt to explain the processes underlying the clusters of lung cancer detected but hypotheses are generated for future studies. Important findings from this study are further discussed below. First, lung cancer clusters prior to year 2000 generally had higher RRs compared to clusters after year 2000 demonstrating that lung cancer diagnoses are declining across Michigan. There were many cities in particular that had high RR of lung cancer prior to year 2000 but not afterwards, in particular the central area of Detroit and the cities of Battle Creek, Jackson, Sterling Heights and St. Clair. These highly populated cities should therefore be studied further to verify that risk for lung cancer has actually reduced versus the need for improved screening and diagnosis. Battle Creek prior to year 2000, for example, was contained

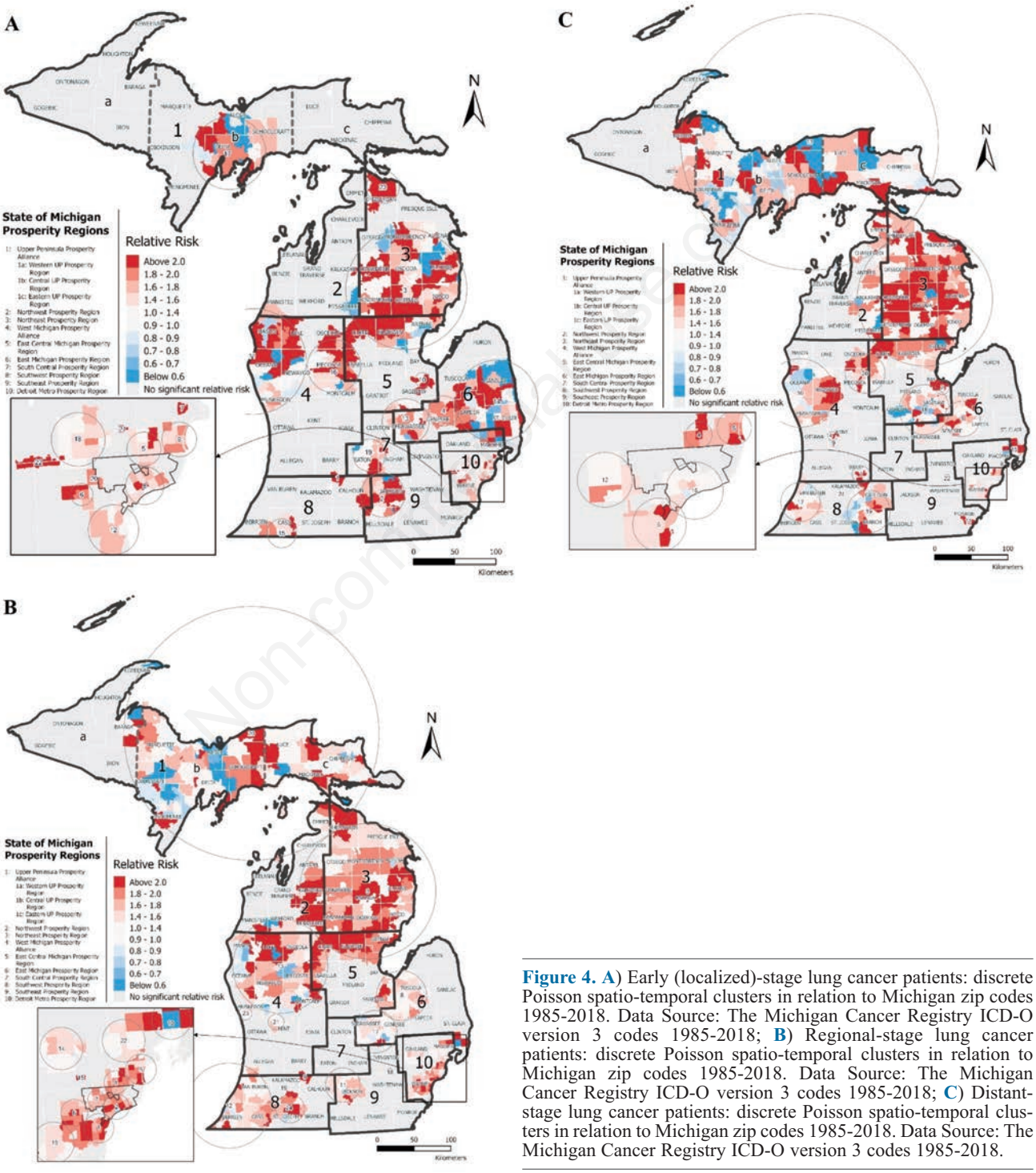

Figure 4. A) Early (localized)-stage lung cancer patients: discrete Poisson spatio-temporal clusters in relation to Michigan zip codes 1985-2018. Data Source: The Michigan Cancer Registry ICD-O version 3 codes 1985-2018; **B**) Regional-stage lung cancer patients: discrete Poisson spatio-temporal clusters in relation to Michigan zip codes 1985-2018. Data Source: The Michigan Cancer Registry ICD-O version 3 codes 1985-2018; **C**) Distant-stage lung cancer patients: discrete Poisson spatio-temporal clusters in relation to Michigan zip codes 1985-2018. Data Source: The Michigan Cancer Registry ICD-O version 3 codes 1985-2018.

Table 1. Discrete Poisson space-time clusters of lung cancer for all patients and by stage of disease for clusters after year 2000 in Michigan 1985-2018.

Cluster no.	Observed patients (no.)	Population (no.)	Relative risk	Log-likelihood ratio	p	Start year	End year	Prosperity region
All Patients								
1	3,967	219,451	1.70	468.89	<0.0001	2008	2017	3
10	2,902	219,236	1.38	132.53	<0.0001	2009	2018	6
11	3,001	200,720	1.35	121.77	<0.0001	2009	2018	2,4
14	1,820	131,862	1.37	82.15	<0.0001	2000	2009	5
16	2,760	208,986	1.27	70.70	<0.0001	2006	2015	9
18	2,468	191,094	1.26	61.14	<0.0001	2009	2018	8
20	1,347	108,357	1.36	58.22	<0.0001	2010	2018	1
22	1,673	147,368	1.28	46.03	<0.0001	2004	2013	10
23	1,762	133,577	1.42	43.01	<0.0001	2008	2017	6
26	616	42,736	1.33	22.79	<0.0001	2009	2018	4
28	344	53,740	1.40	17.71	0.0031	2005	2009	10
31	806	60,305	1.23	16.53	0.010	2004	2013	10
Localized (early) stage lung cancer								
4	665	21,8049	1.74	85.40	< 0.0001	2008	2017	6
12	594	21,5606	1.51	44.11	< 0.0001	2008	2017	10
13	252	88,009	1.64	26.27	< 0.0001	2009	2017	6
14	164	217,844	1.83	24.49	< 0.0001	2016	2017	6
15	127	42,966	1.88	20.54	0.0003	2010	2017	8
16	131	59,023	1.84	19.92	0.0005	2012	2017	4
18	484	209,496	1.35	19.44	0.0008	2009	2017	10
21	31	10,129	3.44	16.31	0.015	2013	2017	10
23	62	14,114	2.21	15.17	0.031	2001	2010	3
Regional stage lung cancer								
1	930	219,451	1.85	142.25	< 0.0001	2000	2009	3
3	813	219,236	1.69	93.76	< 0.0001	2000	2009	6
20	119	26,210	1.88	19.34	0.0006	2004	2013	4
Distant stage lung cancer								
1	1,677	213,171	2.02	327.816	< 0.0001	2005	2014	10
3	1,806	219,451	1.91	303.48	< 0.0001	2004	2013	3,2
5	1,467	200,987	1.85	225.25	< 0.0001	2005	2014	10
6	1,482	209,228	1.79	205.33	< 0.0001	2006	2015	10
7	1,184	196,199	1.75	153.43	< 0.0001	2006	2015	10
8	1,209	216,932	1.55	100.50	< 0.0001	2005	2013	6
11	743	113,373	1.56	63.09	< 0.0001	2004	2013	9
12	1,217	206,950	1.40	61.75	< 0.0001	2008	2017	8
13	784	190,103	1.52	60.41	< 0.0001	2008	2013	4
14	1172	217,929	1.37	52.13	< 0.0001	2009	2017	10
16	435	60,305	1.60	41.11	< 0.0001	2006	2015	10
17	1013	191,434	1.29	29.75	< 0.0001	2008	2017	6
20	991	207,788	1.23	20.23	0.0004	2009	2017	1,3
21	269	105,244	1.48	18.32	0.0023	2009	2012	4
22	637	202,720	1.28	18.26	0.0024	2010	2014	10
23	381	83,963	1.35	15.19	0.0330	2005	2012	4
24	83	32,865	1.96	15.10	0.0340	2009	2011	8

in the most significant cluster for localized lung cancer (RR=5.36, n=160 patients), the most significant cluster for regional lung cancer (RR=3.62, n=123 patients) and the second most significant cluster for distant lung cancer (RR=3.29, n=196 patients). While the numbers of patients diagnosed with lung cancer was relatively low compared to other cities, the RRs were more elevated which indicates that further investigation into the potential of unrecognized lung cancer risk *vs*. a true decline in risk after 2000 is warranted. Since 2000, the rate of lung screening has increased nationwide. The rate of lung cancer screening is still very low compared to screening for other solid organ tumours such as breast cancer. Despite this increased screening rate, however, the risk of lung cancer in locations like Battle Creek has decreased. Other phenomena, such as decreased smoking rates in these areas, may account for the decreased risk.

Second, there were also areas in Michigan in which cancer clusters emerged after year 2000. In particular are the rural areas of the northern lower peninsula and the upper peninsula. Region 3 in north-eastern Michigan emerged as the most significant cluster for all lung cancer in 2008 to 2017, with a cluster of regional lung cancer detected in years 2000 to 2009 and a cluster of distant lung cancer detected in years 2005 to 2014. A new cluster of localized lung cancer was also detected in Cheboygan in the farthest northern area of Region 3. Rural clusters were also detected in Region 2 that extended into the upper areas of Region 4 in particular near Mecosta County, which was highly significant for regional lung cancer in years 2004 to 2013 and distant lung cancer in 2008 to 2013. Finally, noteworthy persistent clusters of lung cancer after 2000 also occurred in Region 6 in the area of Otter Lake Township (localized 2016-2017, regional 2000-2009 and distant 2005-2013). The emergence of these late-stage lung cancer diagnoses in the northern regions of rural Michigan requires further investigation to evaluate individual and environmental risk factors and to ensure continued screening and access to high quality health care.

Third, while there were only 3 clusters that extended across the decades for all lung cancer patients in the space-time cluster analysis, including the cities of Wyoming/Grand Rapids in Kent County Region 4 (RR=1.9, n=19,342 patients) 1995-2004, East pointe and Roseville in Macomb County north of Grosse Pointe (RR=1.56, n=2,121 patients) 1996-2005 and south-western Detroit (RR=1.61, 3,180 patients) 1998-2008 in Region 10, there were more clusters when the stage of disease was studied. In particular, there were trends in significant clusters of lung cancer by stage of disease within the city of Detroit defined by eastern, central (downtown) and south-western Detroit. In central Detroit significant clusters of lung cancer were defined as localized in 1985-1991, regional in 1985-1994 and distant in 1985-1994 all occurring before year 2000; whereas in eastern Detroit significant clusters of lung cancer were defined as distant (2006-2015); and in south-western Detroit significant clusters of lung cancer were defined as localized (2008-2017), regional (1992-2001) and distant (2005-2014) demonstrating the range in stage of diagnosis—with timely emphasis on localized and distant lung cancer in this area.

Finally, the socio-demographic and environmental contexts of these spatio-temporal clusters of lung cancer are also important to understand to generate new hypotheses about individual and environmental risks and the need for improved screening and/or health care. For example, the city of Battle Creek has a slightly younger (15.5% population < 65 years) and poorer (poverty rate, 22.0%) population compared to Sterling Heights (17.8% and 9.8%) and St. Clair County (19.8% and 11.1%) (U.S. Bureau of the Census, 2022). It may be important therefore, to explore access to lung cancer screening and health care in Battle Creek and changes in behavioural risks compounded by aging in Sterling Heights and St. Clair County to further explain why late-stage lung cancer clusters in these areas were detected prior to year 2000 and not afterwards. Furthermore, the emerging clusters of regional and distant lung cancer in the northern area of the lower peninsula as well as the upper peninsula of Michigan after year 2000 may in part be due to increased risk vs. improved lung cancer screening—both explanations challenged by its rural environment and aging population (Demographic Data Northeast Prosperity Alliance, 2023). The increased risk may be a result of an aging population in rural Michigan with older people having less ability to drive long distances and/or commit to frequent trips to doctor's appointments including the doctor's appointments for the treatment of lung cancer following a diagnosis. The city of Detroit is the largest urban area of Michigan that was once an economic centre but with the decline of the auto industry there has led to a substantial out-migration of population and rise in unemployment, particularly among its racially segregated and poor residents. Within Detroit, there were clusters of lung cancer defined by geography as eastern, central and south-western Detroit. Downtown in Detroit is currently undergoing gentrification, which may partly explain why clusters of all lung cancer and lung cancer by all stages were detected before year 2000 but not thereafter. Health systems in eastern Detroit may have more recently improved outreach of lung cancer screening to explain the emergence of distant stage lung cancer. In contrast, south-western Detroit is an industrial hub with numerous industries emitting regulated and unregulated amounts of air pollutants, which may in part explain the persistence of all lung cancer (1999-2008), regional lung cancer (1992-2001), distant (2005-2014) and localized (2008-2017) lung cancer. The more recent cluster of localized lung cancer in south-western Detroit may be due to improved screening in this area. Future research will need to further evaluate socio-demographic, environmental and health system explanations for the spatial and spatio-temporal clusters of lung cancer found in this study.

Previous literature has examined spatio-temporal clusters of cancer in other areas of the world. A study of Pennsylvania identified 5 significant clusters between 2010 and 2017 (Camina *et al.*, 2022). In contrast, a 15 year study of South Korea identified consistently high clusters in the rural and southern parts of the country (Nguyen *et al.*, 2023). However, our study is distinguished from these previous studies in that we examined cancer rates over a very long time period. Given the comprehensive nature of Michigan's cancer registry, we had access to every case of lung cancer recorded over 33 years. In addition, Michigan is unique in having few urban areas surrounded by large rural areas and being a state made up of two distinct peninsulas.

Limitations

This is the first study to detect clusters of lung cancer in Michigan. The clusters detected were therefore, unadjusted in order to explore their spatial patterns and generate new hypotheses about underlying risks and reasons for the high incidence of advanced-stage disease. A lung cancer incidence of 3.0% was used as a sensitive spatial measure by which to define the underlying population at risk, with the intention of ensuring the capture of people <50 years with a lung cancer diagnosis. While age-adjustment can be a beneficial approach in many studies, it can also limit the ability to detect clusters in age-specific populations. Future

research will adjust for age and other individual-level covariates to further learn about the processes underlying these spatial patterns. Also, this study did not account for changes in environmental pollutant levels. Measuring the correlations between pollutant levels and lung cancer risk are important and future studies will need to focus on the aetiology of lung cancer, the range in lag periods from exposure to disease for various pollutants and where to target screening and treatment interventions. Further, it will be important to focus these studies at regional and local scales. Further, future studies on lung cancer should also focus the local scale and the spatio-temporal variation of late-stage disease for residents living in cities near industries vs. gentrification in relation to available health care.

Conclusions

Temporal trends in lung cancer are declining in the United States and Michigan. This statewide study of Michigan at the zip code level demonstrated the value of detecting spatial and spatio-temporal trends. Future research will focus on the etiology of lung cancer, the range in lag periods from exposure to disease for various pollutants and where to target screening and treatment interventions. Further, it will be important to focus these studies at regional and local scales. Prosperity regions in Michigan are defined as economic regions that can help to understand the major industries and labor participation and the health system at a smaller scale. In the rural area of Prosperity Region 3 where there were significant clusters of late-stage disease (regional and distant) and for these, future research should focus on improved access to telemedicine that can enable early diagnoses and referral for treatment. Future studies on lung cancer should also focus the local scale and the spatio-temporal variation of late-stage disease for residents living in cities near industries vs. gentrification in relation to available health care such as in southwest Detroit. This study highlights the importance of requiring a strong health system by which the public and patients are educated about lung cancer prevention and screening and treatments are provided by the health care sector with the goals of preventing lung cancer, and early diagnosis if it occurs.

References

Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD, 2013. Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 143:e1S-e29S.

American Cancer Society, 2023a. Key Statistics for Lung Cancer. Retrieved January 25, 2023. Available from: https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html#:~:text=The%20American%20Cancer%20Society%27s%20estimates,men%20and%2059%2C910%20in%20women.

American Cancer Society, 2023b, January. What is Lung Cancer? Retrieved January 26, 2023a Available from: https://www.cancer.org/cancer/types/lung-cancer/about/what-is.html#:~:text=About%2080%25%20to%2085%25%20of,(outlook)%20are%20often%20similar

Amin R, Bohnert A, Holmes L, Rajasekaran A, Assanasen C, 2010. Epidemiologic mapping of Florida Childhood Cancer Clusters. Pediatr. Blood Cancer 54:511-8.

Barta J.A, Powell CA, Wisnivesky JP, 2019. Global Epidemiology of Lung Cancer. Ann Glob Health 85:2419.

Camina N, McWilliams TL, McKeon TP, Penning TM, Hwang WT, 2022. Identification of spatio-temporal clusters of lung cancer cases in Pennsylvania, USA: 2010-2017. BMC Cancer 22:555.

Centers for Disease Control and Prevention (CDC) United States Cancer Statistics 2023. Retrieved on February 9, 2024. Available from https://gis.cdc.gov/Cancer/USCS/#/AtAGlance/

De Groot PM, Wu CC, Carter BW, Munden RF, 2018. The epidemiology of Lung Cancer. Transl. Lung Cancer Res 7:220-33.

Demographic Data Northeast Prosperity Alliance, 2023. Retrieved on June 21, 2023. Available from http://www.discovernortheastmichigan.org/demographics.asp

Eldridge L, 2022. Which countries have the highest rate of lung cancer? Latest data from the Global Burden of Disease Study. Verywell Health. Available from: https://www.verywellhealth.com/where-in-the-world-is-lung-cancer-most-common-2248876?print.

Goss PE, Strasser-Weippl K, Lee-Bychkovsky BL, Fan L, Li J, Chavarri-Guerra Y, Liedke PE, Pramesh CS, Badovinac-Crnjevic T, Sheikine Y, Chen Z, 2014. Challenges to effective cancer control in China, India, and Russia. Lancet Oncol 15:489-538.

Javid SH, Varghese TK, Morris AM, Porter MP, He H, Buchwald D, Flum DR, 2014. Guideline-concordant cancer care and survival among American Indian/Alaskan Native patients. Cancer 120:2183-90.

Landsdorp-Vogelaar I, Kuntz KM, Knudsen AB, Van Ballegooijen M, Zauber AG, Jemal A, 2012. Contribution of screening and survival differences to racial disparities in colorectal cancer rates. Cancer Epidemiol. Biomarkers Prev 21:728-36.

Lawson AB, 2006. Disease cluster detection: a critique and a Bayesian proposal. Statistics Medicine 25(5):897-916.

Lee YJ, Kim J, Kim SK, Ha S, Mok TS, Mitsudomi T, Cho BC 2011. Lung cancer in never smokers: Change of a mindset in the molecular era. Lung Cancer 72:9-15.

Lovly CM, 2022. Expanding Horizons for Treatment of Early-Stage Lung Cancer. N Engl J Med 386:2050-1.

Nguyen CT, Song I, Jung I, Choi YJ, Kim SY, 2023. Changes in spatial clusters of cancer incidence and mortality over 15 years in South Korea: implication to cancer control. Cancer Med 12(:17418—27.

Pui DY, Chen S, Zuo Z, 2014. PM 2.5 in China: Measurements, sources, visibility and health effects, and mitigation. Particuology 13:1-26.

Siegel RL, Miller KD, Jemal A, 2018. Cancer statistics, 2018. CA Cancer J Clin 68;7-30.

Thun MJ, Carter BD, Feskanich D, Freedman ND, Prentice R, Lopez AD, Hartge P, Gapstur SM, 2013. 50-year trends in smoking-related mortality in the United States. N. Engl. J. Med., 368:351-64.

Torabi M, Galloway K, 2018. Examining the impact of the number of regions used in cluster detection methods: An application to childhood asthma visits to a hospital in Manitoba, Canada. Geospat. Health 13:696.

Torre LA, Siegel RL, Jemal A, 2016. Lung Cancer Statistics. Adv Exp Med Biol 893:1-19.

U.S. Bureau of the Census, 2022. Retrieved on June 21, 2023.

Available from https://www.census.gov/quickfacts/battlecreekcitymichigan.
Van der Heyden JH, Schaap MM, Kunst AE, Esnaola S, Borrell C, Cox B, Leinsalu M, Stirbu I, Kalediene R, Deboosere P, Mackenbach JP, 2009. Socioeconomic inequalities in lung cancer mortality in 16 European populations. Lung Cancer 63:322-330.
Walters S, Maringe C, Coleman MP, Peake MD, Butler J, Young N, Bergström S, Hanna L, Jakobsen E, Kölbeck K, Sundstrøm S, 2013. Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: A population-based study, 2004–2007. Thorax 68:551-564.
Wang Q, Shen JJ, Sotero M, Li CA, Hou Z, 2018. Income, occupation and education: Are they related to smoking behaviors in China? PLoS One 13:e0192571.
World Cancer Research Fund International, 2022, April 14. Lung cancer statistics. Retrieved January 28, 2023. Available from: https://www.wcrf.org/cancer-trends/lung-cancer-statistics/
Zhang X, Wang L, Wang W, Cao D, Wang X, Ye D, 2015. Long-term trend and spatiotemporal variations of haze over China by satellite observations from 1979 to 2013. Atmos Environ 119:362-73.

Online Supplementary Materials
Table 1S. All lung cancer patients: discrete poisson purely spatial clusters, Michigan 1985-2018.
Table 2S. All lung cancer patients: discrete poisson space-time clusters, Michigan 1985-2018.
Table 3S. Localized (early) stage lung cancer: discrete poisson space-time clusters, Michigan 1985-2018.
Table 4S. Regional stage lung cancer: discrete poisson space-time clusters, Michigan 1985-2018.
Table 5S. Distant stage lung cancer: discrete poisson space-time clusters, Michigan 1985-2018.