Journal of Carcinogenesis

Special issue on Cancer Prevention and Detection

Original Article

Toward the identification of communities with increased tobacco-associated cancer burden: Application of spatial modeling techniques

Noella A. Dietz*, Recinda Sherman[1], Jill MacKinnon[1], Lora Fleming[2] , Kristopher L. Arheart, Brad Wohler[1], David J. Lee

Department of Epidemiology and Public Health, [1]Florida Cancer Data System, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, [2]European Centre for Environment and Human Health, Peninsula College of Medicine and Dentistry, Turo, Cornwall, UK

E-mail: ndietz@med.miami.edu
*Corresponding author

Published: 21 September, 2011
Journal of Carcinogenesis 2011, 10:22
This article is available from: http://www.carcinogenesis.com/content/10/1/22

Received: 04 June, 2011
Accepted: 07 August 2011

Abstract

Introduction: Smoking-attributable risks for lung, esophageal, and head and neck (H/N) cancers range from 54% to 90%. Identifying areas with higher than average cancer risk and smoking rates, then targeting those areas for intervention, is one approach to more rapidly lower the overall tobacco disease burden in a given state. Our research team used spatial modeling techniques to identify areas in Florida with higher than expected tobacco-associated cancer incidence clusters. **Materials and Methods:** Geocoded tobacco-associated incident cancer data from 1998 to 2002 from the Florida Cancer Data System were used. Tobacco-associated cancers included lung, esophageal, and H/N cancers. SaTScan was used to identify geographic areas that had statistically significant (P<0.10) excess age-adjusted rates of tobacco-associated cancers. The Poisson-based spatial scan statistic was used. Phi correlation coefficients were computed to examine associations among block groups with/without overlapping cancer clusters. The logistic regression was used to assess associations between county-level smoking prevalence rates and being diagnosed within versus outside a cancer cluster. Community-level smoking rates were obtained from the 2002 Florida Behavioral Risk Factor Surveillance System (BRFSS). Analyses were repeated using 2007 BRFSS to examine the consistency of associations. **Results:** Lung cancer clusters were geographically larger for both squamous cell and adenocarcinoma cases in Florida from 1998 to 2002, than esophageal or H/N clusters. There were very few squamous cell and adenocarcinoma esophageal cancer clusters. H/N cancer mapping showed some squamous cell and a very small amount of adenocarcinoma cancer clusters. Phi correlations were generally weak to moderate in strength. The odds of having an invasive lung cancer cluster increased by 12% per increase in the county-level smoking rate. Results were inconsistent for esophageal and H/N cancers, with some inverse associations. 2007 BRFSS data also showed a similar results pattern. **Conclusions:** Spatial analysis identified many nonoverlapping areas of high risk across both cancer and histological subtypes. Attempts to correlate county-level smoking rates with cancer cluster membership

Access this article online	
Quick Response Code:	Website: www.carcinogenesis.com
	DOI: 10.4103/1477-3163.85184

yielded consistent results only for lung cancer. However, spatial analyses may be most useful when examining incident clusters where several tobacco-associated cancer clusters overlap. Focusing on overlapping cancer clusters may help investigators identify priority areas for further screening, detailed assessments of tobacco use, and/or prevention and cessation interventions to decrease risk.

Keywords: Cancer cluster, cancer, spatial analysis, tobacco use

INTRODUCTION

Tobacco use is the single most preventable cause of morbidity and mortality in the United States. Tobacco products contain over 60 known carcinogens in mainstream smoke and nearly as much in sidestream smoke.[1]Cancers with the strongest association with tobacco use include lung, esophageal, and head and neck. The proportion of lung, esophageal, and head and neck cancer deaths attributable to smoking range from 71% to 87% in men and 45% to 70% in women.[2]

While smoking rates have decreased over time, this decline has begun to level off, with the median rate for all US states at 19.8%.[3] The adult smoking prevalence rate in Florida is similar to the median rate (19.3%).[4] Initiatives in Florida to increase smoking cessation often are implemented at the state level, broadly targeting the population to encourage quit readiness. However, some of the state's population subgroups have smoking rates that exceed the state average. These population subgroups bear a higher share of the burden from tobacco-associated cancers than others. Hence, there is often substantial geographic variation in cancer risks. Identifying areas with higher than average cancer risk and smoking rates, then targeting those areas for enhanced intervention, is one approach to more rapidly lower the overall tobacco disease burden in a given state. A frequently underutilized tool to identify these geographic areas and populations at risk is the use of spatial models for the identification of communities with high rates of tobacco-associated cancers.

Spatial analysis uses a statistical approach to answer questions about the complex pathway of cancer development by integrating the analysis of physical, social, and cultural environments.[5] Spatial analysis, such as desktop geographic information systems (GIS) software, allows researchers to see patterns and relationships in the data based on geography, with results helping researchers postulate about a community's health, focus public health action, and choose the best interventions.[6] Using surveillance data from central cancer registries and GIS software, the identification of individuals' at risk for disease based on geographic community of residence is a relatively inexpensive undertaking. GIS technology also permits the linkage, based on geographic location, of otherwise incongruent data sources for analysis, such as patient level cancer registry data with sociodemographic data from the US Census. Linkages with census data have demonstrated, for example, that late-stage breast cancer clusters are more likely to be located in communities with high rates of poverty.[7]

At present, there are no studies which have attempted to correlate tobacco-associated cancer clusters with variations in community-level smoking rates. To better understand this relationship, our research team used spatial modeling techniques to identify areas in Florida with higher than expected tobacco-associated cancer incidence clusters. Specifically, we examine, at the census block group level, the association between tobacco-associated cancer clusters and community-level smoking rates. We also sought to examine the overlap in the geographic location of the various tobacco-associated cancer clusters. Specifically, we hypothesized that there should be a fair amount of geographic overlap given the cancers we selected for modeling, with each having a high smoking-attributable risk. Finally, we examined associations between location of identified tobacco-associated cancer clusters and county-level smoking rates estimated from population-based state surveillance data.

MATERIALS AND METHODS

We used geocoded tobacco-associated incident cancer data from 1998 to 2002 from the Florida Cancer Data System (FCDS), Florida's incidence cancer registry. Household residence at the time of diagnosis was used to geocode the cases. The tobacco-associated cancers included invasive lung, invasive esophageal, and invasive head and neck cancers. These cancers were further classified by histological subtypes; however, the lung cancer histological type was restricted to squamous cell and adenocarcinoma due to its association with tobacco. Therefore, all three cancer sites were classified as either squamous cell or adenocarcinoma.

Because the only population estimates available at the subcounty level were from the 2000 Census, we used tobacco-associated cancers diagnosed from 1998 to 2002 to most closely align with the available demographic data. Using the block group as the smallest area of geography, SaTScan version 5.0 was used to identify geographic areas within Florida that had statistically significant ($P<0.10$) excess age-adjusted rates of tobacco-associated cancers. We employed SaTScan, a publically available cluster software program developed for the National Cancer Institute (NCI) which utilizes a Poisson-based spatial scan statistic to identify cancer clusters. The spatial scan creates an infinite number of discreet, circular windows (which vary in size and location) across geographic areas. Each circle was evaluated as a possible cancer cluster; the ratios of observed versus expected rates are calculated and tested for significance. Phi correlation coefficients were computed to examine associations among the block groups with and without overlapping cancer clusters. Logistic regression was used to assess associations between county-level smoking prevalence rates and tobacco-related cancer clusters. Estimates of community-level smoking rates were obtained from the 2002 Florida Behavioral Risk Factor Surveillance System (BRFSS). Logistic regression analyses were repeated using 2007 BRFSS data to examine the consistency of associations.

RESULTS

Overall, we modeled lung, esophageal, and head and neck cancers and found clusters for each cancer site and histological type. Lung cancer clusters were geographically larger for both squamous cell and adenocarcinoma cases in Florida from 1998 to 2002, than for esophageal or head and neck clusters [Figures 1 and 2]. There were very few squamous cell esophageal cancer clusters in Florida and a limited number of adenocarcinoma esophageal cancer clusters, both being fairly small in geographical size. Head and neck cancer mapping showed some squamous cell head and neck cancer clusters and a very small amount of adenocarcinoma cancer clusters. In Table 1, we present phi correlations documenting the level of overlapping tobacco-associated cancer clusters. High correlations would reflect a large geographic overlap of these cancers; in this study, the correlations were generally weak to moderate in strength, with the strongest correlations seen within the same cancer sites.

Figure 3 displays all invasive cancer clusters overlapping with other tobacco-associated cancers. The lung cancer and head and neck cancer overlay [Figure 3a] produced the largest number of overlapping clusters. The number of overlapping clusters for lung and esophageal [Figure 3b] produced fewer overlapping clusters, although both figures showed similar geographic patterns. Finally, Figure 3 displays the overlay of all three identified tobacco-associated cancer clusters, which also showed similar geographic patterning.

In examining the 2002 BRFSS data, the odds of having an invasive lung cancer cluster significantly increased by 12% per increase in the county-level smoking rate [Table 2]. Results were inconsistent for esophageal and head and neck cancers. There was an increased odds ratio of 1.052 for adenocarcinoma esophageal cancer, but for squamous cell esophageal cancer a significant reduced odds of being diagnosed within versus without a cancer cluster has the county smoking level increased (OR = 0.863). Further, there were significant and inverse associations (i.e., lower risk) noted for all head and neck cancers (OR = 0.916, 95% CI = 0.903 −0.930; OR=0.665, 95% CI=0.644 −0.687; and OR=0.876, 95% CI=0.860 −0.892, respectively). Logistic regressions using the 2007 data also showed similar results, thereby confirming the stability of the associations based on 2002 BRFSS estimates [Table 2]. For instance, the odds of having an invasive lung cancer increased by 13% per increase in the county-level smoking rate. Invasive esophageal cancer increased by 3% and adenocarcinoma esophageal cancer increased by 11% (OR=1.109, 95% CI=1.085−1.133). The decline in squamous cell esophageal cancer (OR=0.944, 95% CI=0.923 −0.965) per increase in the county-level smoking rate also was present in the 2007 data. Additionally, there remained significant and inverse associations for all head and neck cancers.

DISCUSSION

Despite the relatively high smoking attributable risk for these

Table 1: Correlations among lung, esophageal, and head and neck cancer clusters in Florida

	Inv lung	Inv Esoph	Inv H/N	Adeno lung	Adeno Esoph	Adeno H/N	Squam lung	Squam Esoph	Squam H/N
Invasive lung	1.00								
Invasive esophageal	0.219*	1.00							
Invasive head and neck	0.177*	0.359*	1.00						
Adenocarcinoma lung	0.317*	0.111*	0.249*	1.00					
Adenocarcinoma esophageal	0.050*	–0.017	–0.122*	–0.041*	1.00				
Adenocarcinoma head and neck	–0.103*	0.021+	0.284*	–0.059*	–0.072*	1.00			
Squamous cell lung	0.409*	0.122*	0.136*	0.374*	–0.043*	–0.104*	1.00		
Squamous cell esophageal	0.135*	0.568*	0.258*	0.091*	–0.054*	0.060*	0.050*	1.00	
Squamous cell head and neck	0.312*	0.412*	0.660*	0.197*	–0.098*	0.288*	0.154*	0.306*	1.00

**P*<0.01; +*P*<0.05, Inv, invasive; Esoph, esophageal; H/N, head and neck; Adeno, adenocarcinoma; Squam, squamous.

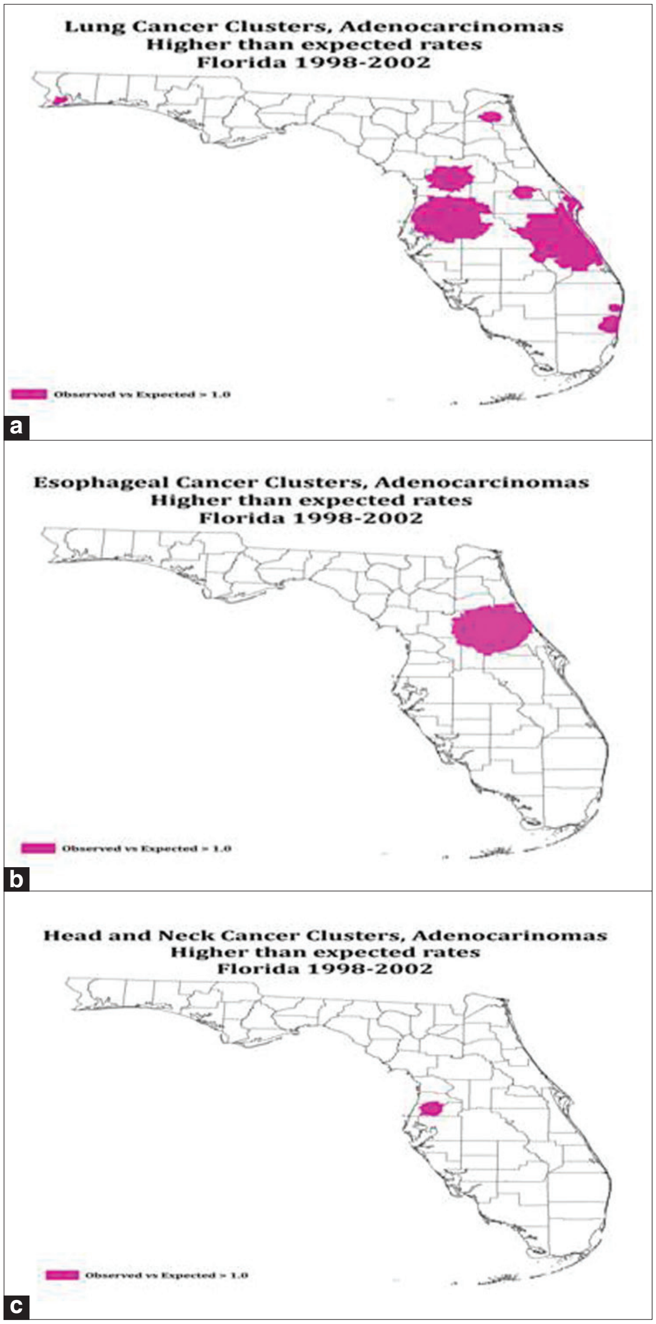

Figure 1: (a) Lung cancer clusters (b) Esophageal cancer clusters, (c) Head and Neck cancer clusters

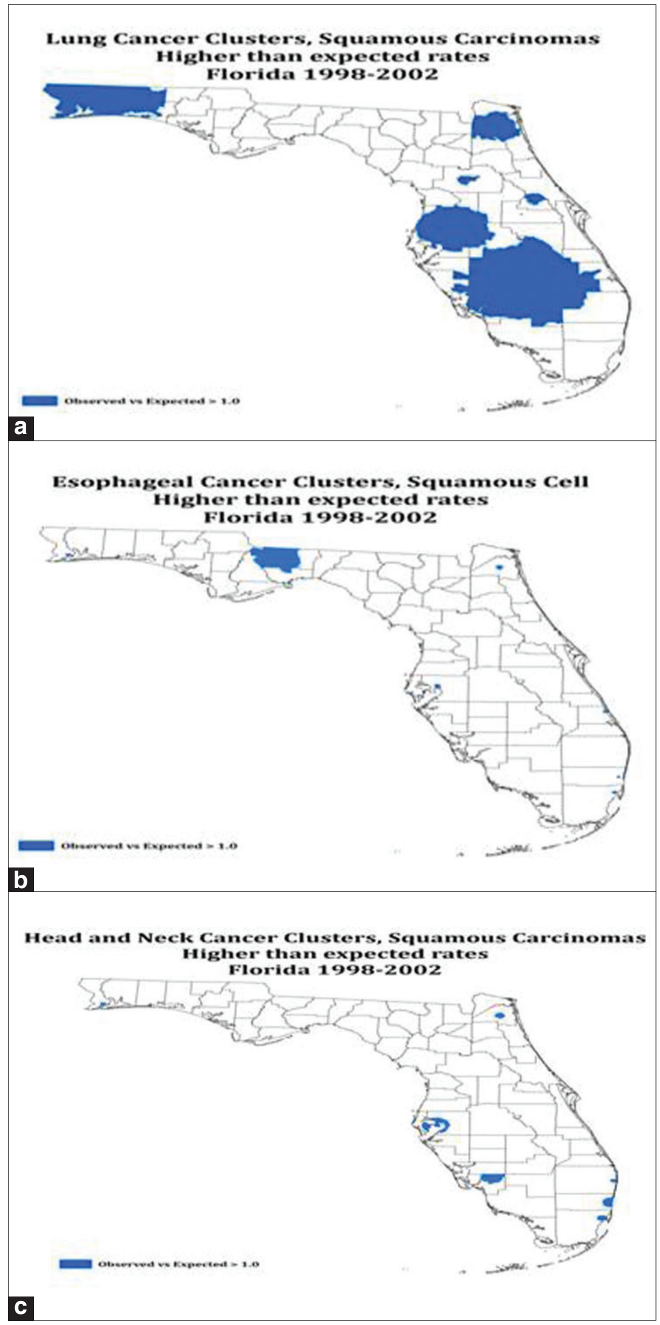

Figure 2: (a) Lung cancer clusters, (b) Esophageal cancer clusters, (c) Head and Neck cancer cluster

cancers, our spatial analysis identified many nonoverlapping areas of high risk across both cancer and histological subtypes. Furthermore, attempts to correlate county-level smoking rates with cancer cluster membership yielded consistent results only for lung cancer, thereby raising questions as to the validity of this approach for examining associations among low incident cancers, which tend to generate clusters of smaller size and often include only portions of a county or counties. Smoking rates have been shown to vary considerably in communities located within counties with substantial sociodemographic heterogeneity. Correlations with larger clusters which span multiple counties may be less susceptible to this form of error and may explain why we found associations between county-level smoking rates and lung cancer clusters.

However, spatial analyses may be most useful when examining incident clusters where several tobacco-associated cancer clusters overlap [Figure 3]. In this instance, we identified multiple overlapping clusters of lung, esophageal, and head and neck cancer incidence throughout Florida. These overlapping

Table 2: Odds ratios showing the association of cancer cluster membership and county-level smoking rates, 2002 and 2007

	Odds ratio	95% confidence interval	*P*- value
2002 Lung			
Invasive lung	1.124	1.110/1.139	0.000
Adeno lung	1.045	1.031/1.058	0.000
Squam lung	1.179	1.163/1.196	0.000
Esophageal			
Invasive esophageal	1.017	0.988/1.046	0.261
Adeno esophageal	1.052	1.023/1.082	0.000
Squam esophageal	0.863	0.839/.887	0.000
Head and neck			
Invasive head/neck	0.916	0.903/.930	0.000
Adeno head/neck	0.665	0.644/.687	0.000
Squam head/neck	0.876	0.860/.892	0.000
2007 Lung			
Invasive lung	1.133	1.121/1.146	0.000
Adeno lung	1.163	1.150/1.177	0.000
Squam lung	1.268	1.251/1.285	0.000
Esophageal			
Invasive esophageal	1.028	1.004/1.053	0.020
Adeno esophageal	1.109	1.085/1.133	0.000
Squam esophageal	0.944	0.923/.965	0.000
Head and neck			
Invasive head/neck	0.938	0.926/.950	0.000
Adeno head/neck	0.730	0.710/0.751	0.000
Squam head/neck	0.949	0.935/0.962	0.000

clusters were more often identified when comparing lung cancer, the more common cancer, and head and neck cancer, which occurs less often. The overlapping clusters raise the possibility of a shared underlying risk factor profile for these cancers in the various identified communities. Comparison of lower incidence cancers, which yield smaller size clusters, may not provide adequate strength to be considered independently; however, focusing on those clusters that incorporate several types of tobacco-related cancers may help investigators identify priority areas for further screening, detailed assessments of tobacco use, and/or prevention and cessation interventions to decrease risk [Figure 3].

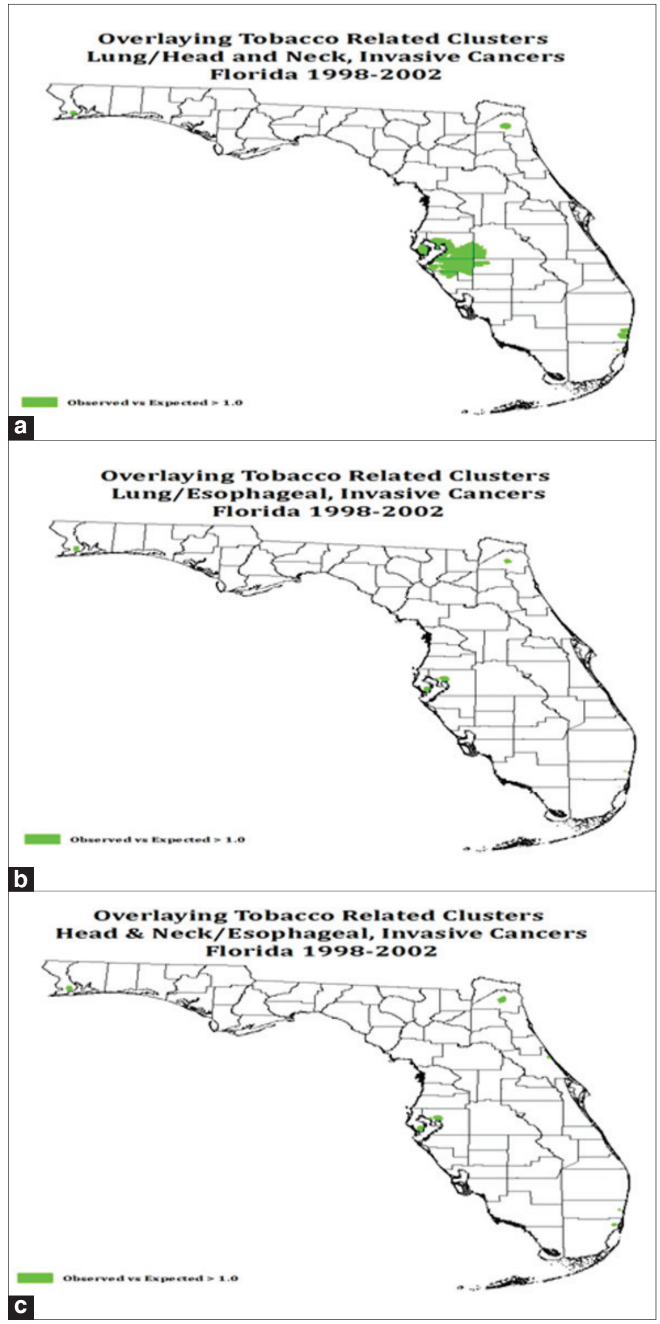

Figure 3: (a) Overlaying lung and head and neck cancer clusters, (b) Overlaying lung and esophageal cancer clusters, (c) Overlaying head & neck and esophageal cancer clusters

Several limitations of the study should be addressed. First, this is an ecological study with no individual-level smoking status information, no individual-level secondhand smoke (SHS) exposure information, and a large temporal distance between tobacco use behavior and cancer diagnosis. In other words, community-level smoking rates may or may not reflect individual-level behaviors and exposures to explain current cancer cases, although it is possible that individuals who grew up in areas with higher than expected tobacco prevalence rates may well be smokers or be exposed to SHS.

Secondly, not all cancer records in the FCDS database were geocoded to the block group level. There are approximately 3% of the cases which are not geocoded and 4% that have incomplete records (2% are geocoded to the zip code centroid of a PO Box address and 2% are geocoded to the zip code centroid of a street address). This likely introduces a level of geographic selection bias because in Florida, like many registries, there are a higher proportion of cases geocoded to the zip code centroid from rural and lower socioeconomic areas.[14,15]

Third, we obtained the community-level smoking rates from the 2002 and 2007 BRFSS. The estimates we used were crude estimates since we were unable to obtain stable rates for each of Florida's 67 counties. Future research examining community-level smoking variables should consider devoting adequate resources for generating stable county level estimates. Additional consideration should be given to generating stable estimates at the subcounty level for communities with substantial sociodemographic variation, including race/ethnicity, and socioeconomic status. Hence, future studies should also consider sociodemographic variables as cancer cluster predictors, in addition to community-level tobacco prevalence. The block group level data on the sociodemographic status of communities are available from the US Census Bureau and can be readily incorporated into spatial analysis to predict cancer cluster membership.[7,16] Finally, future studies should also consider the association between alcohol use and cancer cluster membership. Past research has shown tobacco use and alcohol consumption to have a synergistic effect on cancer risk for esophageal and head and neck cancers; risk factors for these cancers are amplified by alcohol consumption.[17-19] While FCDS data do not contain information on alcohol use, community-level alcohol use rates should be estimated from the BRFSS, or other state level surveys in future studies. Overall, the identification of communities with excess risk for tobacco-associated cancers represents an opportunity to prioritize screening or prevention activities in communities which are suffering from a disproportionate cancer burden.

To summarize, we used spatial analysis to identify many nonoverlapping areas of high risk across both cancer and histological subtypes throughout Florida, although we did identify communities that had excess risk of all three examined tobacco-associated cancers. Attempts to correlate county-level smoking rates with cancer cluster membership yielded consistent results only for lung cancer. Nevertheless, spatial analyses may be most useful for the rapid identification of communities with a simultaneous excess burden of several tobacco-associated cancers. Low-cost identification of these high-risk communities represents a unique opportunity to prioritize screening or prevention activities in communities that are suffering from disproportionate cancer burdens.

Competing interests

None of the authors have a competing interest with this study.

Human participant protection

The protocol was approved by the institutional review board of the University of Miami, Miller School of Medicine.

REFERENCES

1. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3:733-44.
2. Centers for Disease Control and Prevention. Smoking-attributable mortality, years of potential life lost, and productivity losses – United States, 2000 – 2004. MMWR Morb Mortal Wkly Rep 2008;57:1226-8.
3. Centers for Disease Control and Prevention. State-Specific Prevalence and Trends in Adult Cigarette Smoking-United States, 1998-2007. MMWR Morb Mortal Wkly Rep 2009;58:221-6.
4. Available from: http://www.doh.state.fl.us/disease_ctrl/epi/BRFSS_Reports/BRFSS_Index_2007.html. [Last accessed on 2011 May 21].
5. Graves BA. Integrative literature review: a review of literature related to geographical information systems, healthcare access, and health outcomes. Perspect Health Inf Manag 2008;5:11.
6. Mitchell, A. The ESRI Guide to GIS Analysis. Vol. 1. Redlands, California: ESRI Press; 1999.
7. MacKinnon JA, Duncan RC, Huang Y, Lee DJ, Fleming LE, Voti L, *et al*. Detecting an association between socioeconomic status and late stage breast cancer using spatial analysis and area-based measures. Cancer Epidemiol Biomarkers Prev 2007;16:756-62.
8. Burns DM, Anderson CM, Gray N. Do changes in cigarette design influence the rise in adenocarcinoma of the lung? Cancer Causes Control 2011; 22:13-22.
9. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3:733-44.
10. Hammond D, Fong GT, Zanna MP, Thrasher JF, Borland R. Tobacco denormalization and industry beliefs among smokers from four countries. Am J Prev Med 2006;31:225-32.
11. Arheart KL, Sly DF, Trapido EJ, Rodriguez RD, Ellestad AJ. Assessing the reliability and validity of anti-tobacco attitudes/beliefs in the context of a campaign strategy. Prev Med 2004;39:909-18.
12. MacDonald G, Starr G, Schooley M, Yee SL, Klimowski K, Turner K. Introduction to Program Evaluation for Comprehensive Tobacco Control Programs. Atlanta, GA: Centers for Disease Control and Prevention; 2001.
13. Department of Health and Human Services. Best Practices for Comprehensive Tobacco Control Programs. Atlanta, GA: National Center for Chronic Disease Prevention and Health Prevention, Centers for Disease Control and Prevention; 1999.
14. Zandbergen PA. Geocoding quality and implications for spatial analysis. Geogr Compass 2009;3:647-80.
15. Oliver MN, Matthews KA, Siadaty M, Hauck FR, Pickle LW. Geographic bias related to geocoding in epidemiologic studies. Int J Health Geogr 2005;4:29.
16. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: the Public Health Disparities Geocoding Project. Am J Public Health 2005;95:312-23.
17. Toh Y, Oki E, Ohgaki K, Sakamoto Y, Ito S, Egashira A, *et al*. Alcohol drinking, cigarette smoking, and the development of squamous cell carcinoma of the esophagus: molecular mechanisms of carcinogenesis. Int J Clin Oncol 2010;15:135-44.
18. Ansary-Moghaddam A, Huxley RR, Lam TH, Woodward M. The risk of upper aero digestive tract cancer associated with smoking, with and without concurrent alcohol consumption. Mt Sinai J Med. 2009;76:392-403.
19. Pelucchi C, Gallus S, Garavello W, Bosetti C, La Vecchia C Alcohol and tobacco use, and cancer risk for upper aerodigestive tract and liver. Eur J Cancer Prev 2008;17:340-4.

AUTHOR'S PROFILE

Dr. Jill MacKinnon, Florida Cancer Data System Sylvester Comprehensive Cancer Center University of Miami Miller School of Medicine, 3rd Floor Fox Building, Miami, Florida

Dr. Kristopher L Arheart, Department of Epidemiology and Public Health Sylvester Comprehensive Cancer Center University of Miami Miller School of Medicine, 1120 NW 14th Street, 10th Floor, Miami, Florida

Mr. Brad Wohler, Florida Cancer Data System Sylvester Comprehensive Cancer Center University of Miami Miller School of Medicine, 3rd Floor Fox Building, Miami, Florida

Dr. David J Lee, Department of Epidemiology and Public Health Sylvester Comprehensive Cancer Center University of Miami Miller School of Medicine, 1120 NW 14 Street, C202, Miam, Florida

Dr. Noella A Dietz, Department of Epidemiology and Public Health, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine 1120 NW 14th Street, 15th Floor C202, Miami, Florida

Ms. Recinda Sherman, Florida Cancer Data System Sylvester Comprehensive Cancer Center University of Miami Miller School of Medicine 3rd Floor Fox Building, Miami, Florida

Dr. Lora Fleming, European Centre for Environment and Human Health Peninsula College of Medicine and Dentistry Truro, Cornwall, UK