Original Article

Clustering of esophageal cancer among white men in the United States

J. H. Rubenstein,^{1,2} H. Morgenstern,³ K. Longstreth⁴

¹*Veterans Affairs Center for Clinical Management Research, Ann Arbor Veterans Affairs Medical Center and,* ²*Barrett's Esophagus Program, Division of Gastroenterology, University of Michigan Medical School and,* ³*Departments of Epidemiology and Environmental Health Sciences, School of Public Health; Department of Urology, Medical School; and Comprehensive Cancer Center, Michigan Medicine, and* ⁴*Spatial and Numeric Data Services, Stephen S. Clark Library, University of Michigan, Ann Arbor, Michigan, USA*

SUMMARY. Studies of the geographic distribution of esophageal cancer in the United States have been limited. We aimed to examine geographic clustering of esophageal cancer in the United States and assess whether that clustering is explained by the distribution of known risk factors for esophageal cancer. We conducted cluster analyses derived from county mortality rates of esophageal cancer, using publicly available datasets. State incidence rates of esophageal adenocarcinoma were from the National Program of Cancer Registries, and county esophageal-cancer mortality rates were from the Vital Statistics Cooperative Program. County prevalences of cigarette use, alcohol use, obesity, education, and income were published estimates derived from the Behavioral Risk Factor Surveillance System and the American Community Survey. The primary outcomes were clusters of high and low esophageal-cancer mortality rates among non-Hispanic white men, both unadjusted and adjusted for risk factors. Age-standardized county rates of esophageal-cancer mortality among non-Hispanic white men ranged from 4.8 to 21.2 per 100,000/year. There was a cluster of high mortality in the Great Lakes states and New England and a cluster of low mortality in the Southeastern United States. State incidence rates of esophageal adenocarcinoma were consistent with this pattern. Adjusting for risk factors did little to change the pattern of observed rates or the clusters derived from them. Among non-Hispanic white men, there are clusters of high and low mortality rates with esophageal cancer within the United States, likely representing esophageal adenocarcinoma; but those clusters were not explained by several known risk factors. Focusing future efforts in the high-cluster areas might improve the efficiency of cancer screening and control.

KEY WORDS: Behavioral Risk Factor Surveillance System, esophageal adenocarcinoma, geography, SEER program.

ABBREVIATIONS

EAC: esophageal adenocarcinoma;
EC: esophageal cancer;
ESCC: esophageal squamous cell carcinoma;

ICDO: International Classification of Diseases for Oncology;
NPCR: National Program of Cancer Registries;
SEER: Surveillance Epidemiology and End Results;
US: United States.

Address correspondence to: Dr. Joel H. Rubenstein, MD, MSc, VA Medical Center 111-D, 2215 Fuller Rd., Ann Arbor, MI 48105, USA. E-mail: jhr@umich.edu

Specific author contributions: Design: Joel H. Rubenstein, Hal Morgenstern; Acquisition of data: Joel H. Rubenstein, Karl Longstreth; Analysis of data: Joel H. Rubenstein, Karl Longstreth; Interpretation of data: Joel H. Rubenstein, Hal Morgenstern, Karl Longstreth; Drafting the manuscript: Joel H. Rubenstein; Critical revision of the manuscript: Hal Morgenstern, Karl Longstreth. Financial support: Joel H. Rubenstein was supported by VHA grant I01-CX000899, NIH grant U01CA199336, and NIH grant U54CA163059. The funding agencies had no role in the collection, analysis, or interpretation of data.

Conflicts of interest: The authors have no relevant financial conflicts of interest.

INTRODUCTION

The incidence rate of esophageal adenocarcinoma (EAC) in the United States (US) rose seven-fold over three decades.¹ Recognizing geographic clustering of the cancer could improve the efficiency of efforts aimed at reducing mortality of the cancer by directing efforts in those areas. It is well known that EAC is rare in Africa and Asia.² But there is also evidence that EAC may cluster in smaller geographic areas. For

instance, in Europe, the incidence of EAC is greatest in the northern countries, particularly Great Britain, Ireland, Iceland, Belgium, the Netherlands, Denmark, and Norway.² Geographic data within the US have been fairly limited. A study of 11 metropolitan or state cancer registries found the greatest incidence of EAC in Seattle, Connecticut, Iowa, and Detroit, with the lowest incidence in Utah, Los Angeles, and Hawaii.³ And a study of the entire US divided into four large census regions found the incidence of EAC among white men aged 45 to 84 years to be greatest in the midwest and northeast (age-standardized rate [ASR] = 18.1 and 17.7 per 100,000/year, respectively) compared to the south and west (ASR = 14.4 and 14.3 per 100,000/year, respectively).⁴ However, the patterns of EAC burden on smaller geographic scales in the US are not known. We therefore sought to estimate the EAC mortality rates within individual counties throughout the US, and to estimate the mortality adjusted for county prevalences of known risk factors.

METHODS

We performed a series of retrospective ecologic analyses using publicly available datasets. County-level data were limited to mortality without incidence data. In addition, histology was not available for counties, so analyses of counties were restricted to non-Hispanic white men from 1999 through 2015 with the expectation that the great majority of esophageal cancers (EC) in that group would be EAC rather than esophageal squamous cell carcinoma (ESCC), since EAC became the predominant histologic type of EC among non-Hispanic white men in 1989.^{5–7} Analyses of state-level incidence data using histologically confirmed cases of EAC were performed to validate the findings at the county level.

Esophageal cancer incidence and mortality rates

The incidence rates of EAC in individual states were obtained from the National Program of Cancer Registries (NPCR) from 2003 through 2013 (the only dates publicly available) using SEER*Stat v.8.3.2 software (The Surveillance Research Program of the Division of Cancer Control and Population Sciences, National Cancer Institute & Information Management Services, Inc., Calverton, MD).⁸ The NPCR entails five Surveillance Epidemiology and End Results (SEER) registries that cover five entire states plus cancer registries for the other 45 states using the same methodology as SEER. The NPCR cover 96% of the US population. We identified incident cases of microscopically confirmed cancers of the esophagus (site C15), and excluded *in situ* cases. Cases of adenocarcinoma (International Classification of Diseases for Oncology 3 [ICDO3] = 8140–8575) and

squamous cell (ICDO3 = 8052–8078) were abstracted separately for each state, and stratified by sex, race, and Hispanic ethnicity. States have the option to suppress race-specific and Hispanic ethnicity-specific data, so data were not available for Delaware, Illinois, Kansas, Kentucky, Massachusetts, New Jersey, New York, North Dakota, Pennsylvania, and Virginia. Release of data is also restricted to states with at least 16 cases in the stratum over the period of interest. The SEER*Stat software provided incidence rates standardized to the age distribution of the 2000 US population.

The mortality rates with EAC in individual counties were estimated using mortality rates with EC abstracted from the Vital Statistics Cooperative Program Multiple Causes of Death Files from 1999 through 2015; those data were obtained from county death certificates in all US counties.⁹ Deaths from EC were identified using ICD codes C15.0–C15.9. Approximately 15% of underlying causes listed in death certificates in the US is a ‘garbage code,’ such as listing a diagnosis that either a) should not be considered a cause of death (e.g. hypertension), b) reflects an intermediate cause of death (e.g. peritonitis) which had an underlying cause (e.g. diverticulitis which had perforated), or c) reflects an immediate cause that is the final pathway to death (e.g. respiratory arrest).¹⁰ Various complex algorithms have been developed to deal with this problem, but most rely in the end on expert opinion.¹⁰ Since the long-term survival with EC is less than 15%, and since we were interested in mortality rates as a surrogate for incidence rates, we classified as a case of EC mortality any death with EC listed as any of the multiple causes of death (not restricted to only those cases listing EC as the underlying cause). The dataset does not provide any information regarding histology, so we restricted the primary analyses to white non-Hispanic men to approximate the estimation of EAC mortality rates. We confirmed that expectation in the NPCR, where we found that EAC accounted for 71.4% to 87.9% of all incident cases of EC among states for non-Hispanic white men (median = 81.7%). In contrast, the median state proportion of EAC among all EC incident cases was 50.0% for non-Hispanic white women and 15.3% for non-Hispanic African American men. Release of data was restricted to counties with at least 20 deaths in the stratum over the period of interest. Mortality rates were age-standardized to the 2000 US population.

County prevalences of risk factors

Prevalences of risk factors for EC were obtained from two primary sources: the Behavioral Risk Factor Surveillance System, which surveyed 4.7 million adults, sampling more than 90% of US counties; and the American Community Survey, which

surveyed 3.5 million US addresses. The validity and accuracy of these surveys have been described previously.^{11,12} Because observed prevalences in individual counties may not be precise, owing to small sample size in some counties, we used published datasets that applied Bayesian techniques to estimate county-specific prevalences, taking into account the prevalences in neighboring counties. Prevalences considered were only available as dichotomized variables without additional levels of categorization. Obesity (defined as body mass index >30 kg/m²) was available for years 2009 through 2013.¹³ Current cigarette smoking was available for years 1996 through 2012, categorized as daily (vs. less than daily or no use), non-daily (vs. daily or no use), and any use (vs. no use).¹⁴ Alcohol use was available for years 2002 through 2012; available variables were any (≥ 1 drink in past 30 days), heavy (>2 drinks per day in past 30 days), and binge use (>5 drinks on single occasion in past 30 days).¹⁵ Obesity, cigarette smoking, and alcohol use were age-standardized to year 2000 US population and restricted to men, but were not race specific. We also considered median household income and prevalence of poverty (not age-standardized or specific to men or race),¹⁶ as well as educational attainment among various age groups restricted to men.¹⁷ Income was not included in the final model because it had little association with the outcome and did not change the estimates for the other predictors in the model.

Statistical analysis

State and county rates were mapped using ArcGIS v.10.4.1 (ESRI, Redlands, CA). A type of cluster analysis was performed using the ArcGIS Optimized Hot Spot Analysis, which relies on the Getis-Ord Gi* statistical test to provide border contours of high- and low-rate clusters with 90%, 95%, and 99% confidence.¹⁸ The Gi* statistic compares the sum of the rates for a particular county and its neighbors, weighted by its distance to the neighbors, to the sum of the rates for all the counties.

Generalized linear regression modeling was performed using SAS v.9.3 (SAS Institute, Cary, NC) to identify a parsimonious model, taking into account prior knowledge of EC risk factors, model fit and specification, to predict county-level age-standardized EC mortality rates for non-Hispanic white men. The final model used the prevalences of obesity, daily cigarette use, binge alcohol use, and a bachelor degree among men older than 25 years of age. The data for obesity, cigarette, and alcohol use were averaged from 2009 through 2012, and the data for education was averaged from 2008 through 2012.

Next we estimated county-specific age-standardized EC mortality rates adjusted for the risk factors in the linear model. Each adjusted rate was the ratio

of the observed-to-predicted age-standardized rate in the county, multiplied by the EC mortality rate in the US as a whole. We then mapped the adjusted rates and performed cluster analysis as done with the original (unadjusted) age-standardized mortality rates. If the county-specific distributions of risk factors fully explain variability in the county EC mortality rates, we would expect no clustering of adjusted rates; each adjusted rate would be approximately the same as the rate for the total US. In contrast, if the risk-factor distributions do not explain any variability in mortality rates, we would expect the same clustering pattern observed with the original age-standardized mortality rates.

RESULTS

From 1999 through 2015, there were 166,795 deaths with EC among non-Hispanic white men in the US with an age-standardized mortality rate of 8.7 per 100,000/year. Of the 3147 total US counties, 162,183 of the deaths occurred in 1509 counties (the study population). The remaining 1638 counties each had fewer than 20 EC deaths during the period, and so mortality rates for those counties were not available. In the 1509 counties with available data, the age-standardized mortality rate ranged from 4.8 to 21.2 per 100,000/year. Mapping the mortality rates revealed that the counties with missing data were primarily among sparsely populated counties of the middle areas of the US (Fig. 1A). Alaska had average rates and Hawaii low rates (counties included in analysis, but not shown in figures).

Cluster analysis better defined the gross geographic pattern of observed rates, demonstrating a cluster of high-mortality rates in the Northeast and Midwest, spanning from Maine to Illinois, north of the Ohio River and west of the Appalachian Mountains (Fig. 1B). There was also a cluster of low mortality rates in the Southeastern US, from North Carolina to eastern Texas. The mortality rates in the cluster of high mortality were approximately 50% greater than the mortality rates in the low mortality cluster. Smaller areas of high mortality rates also appear in western Wyoming and southwestern Oregon, and low mortality rates also appear in northwestern Utah and central California.

The pattern of age-standardized incidence rates of EAC, by state, among white non-Hispanic men from 2003 through 2013 was similar to the pattern of county mortality rates (Fig. 2). In contrast, the state incidence rates of ESCC among non-Hispanic African American men revealed a different pattern, with high incidence in the Southeastern US, though many states had missing data due to sparse populations of non-Hispanic African American men, low mortality rates, or both (Supplemental Fig. S1). This finding indicates

Fig. 1 Esophageal-cancer mortality among non-Hispanic white men, 1999–2015. (A) Observed county rates (per 100,000/year), age-standardized to the 2000 US population and categorized into color-coded quintiles. Counties in white have missing data due to fewer than 20 cases reported. (B) Cluster analysis of the data shown in Panel A. An area of high mortality rates is found in the Great Lakes States and New England. An area of low mortality rates is found in the Southeastern US.

that the geographic clustering of EC mortality among white non-Hispanic men is unlikely due to local differences in the practices of diagnosing or reporting EC mortality.

The individual β coefficients estimated from the ecologic multivariable regression model of the association of county prevalences of risk factors with esophageal cancer mortality are displayed in Supplemental Table S1. There were positive associations of EC mortality with daily cigarette use, binge alcohol

use, and a weaker association with obesity. College education (bachelor degree) was inversely associated with esophageal cancer mortality. No single prevalence pattern of any of these risk factors across counties matched the observed pattern of esophageal cancer mortality (Supplemental Figs S2–S5 vs. Fig. 1A).

Based on the fitted linear model, the predicted esophageal cancer mortality rate was estimated for each county. This resulted in average rates in the

Fig. 2 Observed state incidence rates of esophageal adenocarcinoma among non-Hispanic white men, for 2003–2013. States in white have missing data due to restriction of release of data stratified by race and/or ethnicity. Rates are per 100,000 person years, age-standardized to the 2000 US population, and classified by quintiles. Note difference in scale compared to Figure 1A.

South and much of the West, low rates in Utah, and high rates in Nevada, the Great Plains states and Midwest (Supplemental Fig. S6). Overall, the pattern of observed county rates for esophageal mortality adjusted for the four risk factors (Fig. 3A) was similar to the pattern of observed mortality rates unadjusted for those risk factors (Fig. 1A). Among the 1509 counties, the adjusted and unadjusted county mortality rates were strongly correlated (Pearson $r = 0.91$).

Results of the cluster analysis adjusting for the county prevalences of the four risk factors are displayed in Figure 3B. The pattern of these risk-factor adjusted clusters was similar to the pattern of unadjusted clusters in the Northeastern and Southern states (Fig. 1B), but they differed in the northern Midwest and Northwest. In particular, much of Michigan and Illinois were no longer included in the cluster, and parts of Wisconsin and Minnesota had lower than expected mortality rates. In addition, an area extending from western Washington to northern California had greater than expected esophageal cancer mortality.

DISCUSSION

We found a cluster of high mortality with esophageal cancer among non-Hispanic white men in the Great Lakes States and New England, and a cluster of low mortality in the Southeastern US. Furthermore, the prevalence of risk factors including obesity, cigarette use, alcohol use, and education do

not appear to explain much of the clustering. The findings are important because recognition of these disease clusters could help efforts aimed at controlling the burden of esophageal cancer mortality by tailoring screening paradigms to regional differences in the burden of disease. For instance, programmatic screening for esophageal adenocarcinoma among white non-Hispanic men may be more justified in the areas of the high mortality rates than in the areas of low mortality rates.

The study was limited by the lack of histologic subtypes of esophageal cancer reported in county death certificates; therefore, we cannot be certain whether the clusters are related to adenocarcinoma or squamous cell carcinoma. However, we restricted our analyses to esophageal cancer in non-Hispanic white men, among whom adenocarcinoma during the study period accounted for approximately 82% of all incident cases of esophageal cancer. Furthermore, using data from state cancer registries, the state-level incidence rate of esophageal adenocarcinoma among non-Hispanic white men, verified by histology, broadly matched the observed county patterns of esophageal cancer mortality. Therefore, the clusters of high and low esophageal cancer mortality very likely reflect esophageal adenocarcinoma. These findings are consistent with the two previous studies on the topic, but those studies were limited to 11 small catchment areas of the US or grossly described differences in four broad census regions of the US.^{3,4} This study provides finer level detail, and the cluster analysis provides smoothed contours of the high and low mortality areas. This study highlights the need for

Fig. 3 Adjusted county mortality rates for non-Hispanic white men with esophageal cancer 1999–2015. (A) Rates are per 100,000 person years, age-standardized to the year 2000 US population, and adjusted for county prevalences of daily cigarette use, binge alcohol use, obesity, and bachelor degree. Rates are classified by quintiles of observed rates from Figure 1A. Counties in white have missing data due to fewer than 20 cases reported. (B) Cluster analysis of the data shown in Panel A. The area of high mortality rates remains in the Great Lakes States and New England, but no longer includes Wisconsin and much of Michigan and Illinois. The area of low mortality rates remain in the Southeastern US. In addition, adjusted for the risk factors, there is a cluster of higher than expected mortality in the Northwestern US, and a cluster of lower than expected mortality in part of Wisconsin and Minnesota.

developing a more robust national cancer registry in the US allowing for finer detail in geographic clustering using high quality data abstracted by trained cancer registrars.

In addition, our results suggest that there are risk or prognostic factors that promote esophageal cancer mortality aside from the known risk factors included in these analyses, and that they are unevenly distributed across the country. If the measured risk factors completely explained variability in

county-specific esophageal-cancer mortality rates, we would expect the rates adjusted for these factors to be uniform throughout the country. Instead, we found that the adjusted rates also demonstrated clusters of high mortality in the Great Lakes States and New England and low mortality in the Southeastern US. This indicates that the regional patterns in esophageal cancer mortality are unlikely to be explained by local differences in prevalences of the risk factors examined. Other potential risk factors that we were unable to

examine include symptoms of gastroesophageal reflux disease, *Helicobacter pylori* infection, diet, and nutritional factors. Given the North-South axis of clustering, we additionally estimated county mortality rates adjusted for ultraviolet radiation, which may be a risk factor for esophageal adenocarcinoma.¹⁹ We found, however, that ultraviolet radiation was only weakly associated with the esophageal cancer mortality rate, and its addition did not change the overall pattern of clusters in our analysis (results not shown). It is possible that the geographic clustering could reflect access to medical care, such as endoscopy. However, some of the greatest use of medical resources in the US, including endoscopy, is in areas of clustering of high mortality, so this seems unlikely to explain the observed patterns.²⁰

As described above, this study has certain limitations regarding classification of histology and absence of data on gastroesophageal reflux disease or *H. pylori*. Nonetheless, if geographic clustering of gastroesophageal reflux disease or *H. pylori* is responsible for the observed clustering of esophageal cancer mortality, then it would beg the question why there would be such clustering of gastroesophageal reflux disease and/or *H. pylori*. In addition, race and ethnicity data were not available by category of the measured risk factors, and we did not have data available for prior use of tobacco or alcohol. Data were missing for the sparsely populated counties in the middle of the country, but the state level data indicates that those areas have average incidence rates of esophageal adenocarcinoma. The mortality data reflect county of death and incidence data reflect state of residence at time of diagnosis; individuals may have moved between counties and states from EC diagnosis to death from EC. This might explain why Florida has greater incidence and mortality rates than other states in the Southeastern US. Lastly, we analyzed county mortality rates rather than incidence rates, but since esophageal cancer is highly fatal, mortality rates are a reasonable surrogate for incidence rates.

Finally, we would like to comment on the potential for ecologic bias in our analyses (sometimes referred to as the ‘ecological fallacy’). Ecologic bias results from the use of aggregated data to make inferences about exposure effects at the individual level.²¹ Our aim, however, was not to make such causal inferences in this study, because we included only factors that are known from previous research to be risk factors for esophageal cancer. Rather, our aim was to assess whether geographical clustering of those risk factors explain the geographic clustering of EC mortality.

Our study also has a number of strengths, including the novel use of a large dataset of county death certificates with corroboration of the data with state level cancer incidence documented by cancer registries using robust protocols. In addition, we conducted an ecologic analysis estimating the county esophageal

cancer mortality rates adjusted for important risk factors, and we enhanced our analysis of county mortality rates using a special type of cluster analysis.

In conclusion, we identified geographic clusters of high and low mortality with esophageal cancer—likely adenocarcinoma—among white non-Hispanic men in the US. Known risk factors do not appear to explain these geographic clusters. The efficiency of efforts directed at screening and controlling esophageal adenocarcinoma could be maximized by focusing efforts in the geographic areas of the high-mortality clusters.

SUPPLEMENTARY DATA

Supplementary data are available at [DOTESO](#) online.

References

- 1 Pohl H, Sirovich B, Welch H G. Esophageal adenocarcinoma incidence: are we reaching the peak? *Cancer Epidemiol Biomarkers Prev* 2010; 19: 1468–70.
- 2 Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* 2015; 64: 381–7.
- 3 Kubo A, Corley D A. Marked regional variation in adenocarcinomas of the esophagus and the gastric cardia in the United States. *Cancer* 2002; 95: 2096–102.
- 4 Drahos J, Wu M, Anderson W F *et al*. Regional variations in esophageal cancer rates by census region in the United States, 1999–2008. *PLoS One* 2013; 8: e67913.
- 5 Wu X, Chen V W, Andrews P A, Ruiz B, Correa P. Incidence of esophageal and gastric cancers among Hispanics, non-Hispanic whites and non-Hispanic blacks in the United States: subsite and histology differences. *Cancer Causes Control* 2007; 18: 585–93.
- 6 Kubo A, Corley D A. Marked multi-ethnic variation of esophageal and gastric cardia carcinomas within the United States. *Am J Gastroenterol* 2004; 99: 582–8.
- 7 Younes M, Henson D E, Ertan A, Miller C C. Incidence and survival trends of esophageal carcinoma in the United States: racial and gender differences by histological type. *Scand J Gastroenterol* 2002; 37: 1359–65.
- 8 National Program of Cancer Registries SEER*Stat Database: NPCR Incidence - Public Use Data, 2001–2013 - jbk 120616. United States Department of Health and Human Services, Centers for Disease Control and Prevention; Released January 2017, based on the November 2015 submission. Available at: www.cdc.gov/cancer/npcr/public-use. (Accessed March 3, 2017).
- 9 Multiple Cause of Death Files 1999–2015 on CDC WONDER Online Database. Vital Statistics Cooperative Program, Centers for Disease Control and Prevention, National Center for Health Statistics; Released December, 2016; Available at: <http://wonder.cdc.gov/mcd-icd10.html>. (Accessed May 26, 2016).
- 10 Naghavi M, Makela S, Foreman K, O’Brien J, Pourmalek F, Lozano R. Algorithms for enhancing public health utility of national causes-of-death data. *Popul Health Metrics* 2010; 8: 9.
- 11 American Community Survey Accuracy of the Data. 2015. <https://www.census.gov/acs/www/data-tables-and-tools/index.php>. (Accessed June 21, 2018).
- 12 Pierannunzi C, Hu S S, Balluz L. A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004–2011. *BMC Med Res Methodol* 2013; 13: 49.
- 13 United States Diabetes Surveillance System. National Center for Chronic Disease Prevention and Health Promotion, Division of Diabetes Translation; Released May 16, 2016. Available at: <http://www.cdc.gov/diabetes/data/countydata/countydataindicators.html>. (Accessed November 28, 2016).

- 14 Dwyer-Lindgren L, Mokdad A H, Srebotnjak T, Flaxman A D, Hansen G M, Murray C J. Cigarette smoking prevalence in US counties: 1996-2012. *Popul Health Metrics* 2014; 12: 5.
- 15 Dwyer-Lindgren L, Flaxman A D, Ng M, Hansen G M, Murray C J L, Mokdad A H. Drinking patterns in US counties from 2002 to 2012. *Am J Public Health* 2015; 105: 1120–7.
- 16 Small Area Income and Poverty Estimates. United States Census Bureau; Available at: <https://www.census.gov/did/www/saipe/data/statecounty/index.html>. (Accessed April 24, 2017).
- 17 Educational Attainment 2008–2012 American Community Survey 5-Year Estimates. United States Census Bureau; Released 2013. Available at: <https://www.census.gov/acs/www/data/data-tables-and-tools/index.php>. (Accessed April 24, 2017).
- 18 Ord J K, Getis A. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Anal* 1995; 27: 286–306.
- 19 Tatalovich Z, Wilson J P, Cockburn M. A comparison of thiesen polygon, kriging, and spline models of potential UV exposure. *Cartogr Geographic Inf Sci* 2006; 33: 217–31.
- 20 Sheffield K M, Han Y, Kuo Y-F, Riall T S, Goodwin J S. Potentially inappropriate screening colonoscopy in medicare patients. *JAMA Intern Med* 2013; 173: 542–50 10.1001/jamainternmed.2013.2912.
- 21 Morgenstern H. Ecological studies. In: K J Rothman, S Greenland Lash T L (eds). *Modern Epidemiology*, 3rd edn. Philadelphia, PA: Lippincott Williams & Wilkins, 2008.