Cancer Epidemiology 59 (2019) 8–14

Contents lists available at ScienceDirect

Cancer Epidemiology

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

Examining determinants of geographic variation in colorectal cancer mortality in North Carolina: A spatial analysis approach

Tzy-Mey Kuo[a,*], Anne Marie Meyer[a,b], Christopher D. Baggett[a,b], Andrew F. Olshan[a,b]

[a] Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[b] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

ARTICLE INFO

Keywords:
Keywords: Geographic variation
Spatial regression
Colorectal cancer mortality
Spatial diffusion

ABSTRACT

Purpose: A recent study using national data from 2000 to 2009 identified colorectal cancer (CRC) mortality "hotspots" in 11 counties of North Carolina (NC). In this study, we used more recent, state-specific data to investigate the county-level determinants of geographic variation in NC through a geospatial analytic approach.
Method: Using NC CRC mortality data from 2003 to 2013, we first conducted clustering analysis to confirm spatial dependence. Spatial economic models were then used to incorporate spatial structure to estimate the association between determinants and CRC mortality. We included county-level data on socio-demographic characteristics, access and quality of healthcare, behavioral risk factors (CRC screening, obesity, and cigarette smoking), and urbanicity. Due to correlation among screening, obesity and quality of healthcare, we combined these factors to form a cumulative risk group variable in the analysis.
Results: We confirmed the existence of spatial dependence and identified clusters of elevated CRC mortality rates in NC counties. Using a spatial lag model, we found significant interaction effect between CRC risk groups and socioeconomic deprivation. Higher CRC mortality rates were also associated with rural counties with large towns compared to urban counties.
Conclusion: Our findings depicted a spatial diffusion process of CRC mortality rates across NC counties, demonstrated intertwined effects between SES deprivation and behavioral risks in shaping CRC mortality at area-level, and identified counties with high CRC mortality that were also deprived in multiple factors. These results suggest interventions to reduce geographic variation in CRC mortality should develop multifaceted strategies and work through shared resources in neighboring areas.

1. Introduction

Colorectal cancer (CRC) is the third leading cause of cancer death in the United States and is expected to result in more than 50,000 deaths in 2018 [1]. The CRC mortality rate has declined dramatically, by 48% from 1970 to 2011, with northeastern states showing greater decreases than southern states [2]. The highest rates of CRC mortality have also shifted from northeastern states to the southern states [2,3]. A recent study further showed clusters (or hotspots) of elevated CRC mortality rates concentrated in areas of the Lower Mississippi Delta, west central Appalachia, and eastern Virginia and North Carolina (NC) [2]. The hotspot of the 11 northeastern counties of NC are in geographic areas that have continually suffered from the highest CRC mortality in recent decades [4]. To understand the driving forces of the excessive mortality rates and to inform tailored interventions for reducing mortality rates in these areas of NC, we sought to identify county-level determinants that contribute to the geographic variation of CRC mortality rates in NC.

Previous studies have identified several important factors related to CRC mortality. Socioeconomic disparities and area-level socioeconomic disadvantage have been found to associate with CRC mortality [5–8]. Better access to primary care and living in urban areas have also been shown to be associated with lower CRC mortality [7–10]. Some behavioral risk factors such as CRC screening, cigarette smoking, and overweight have also been demonstrated to be important predictors of CRC mortality [11–18].

The findings from these existing studies shed some light on determinants related to CRC mortality, but they may not appropriately explain the geographic variations because the studies did not take into consideration the underlying spatial structure. Examining determinants of geographic variation for CRC mortality is complicated by the existence of hotspots of CRC mortality, as it indicates the rates do not distribute randomly across space. Analysis ignoring such spatial

* Corresponding author at: 101 East Weaver Street, CB 7293, University of North Carolina at Chapel Hill, Carrboro, NC 27599-7293, USA.
E-mail address: tkuo@email.unc.edu (T.-M. Kuo).

https://doi.org/10.1016/j.canep.2019.01.002
Received 29 August 2018; Received in revised form 16 December 2018; Accepted 2 January 2019
Available online 11 January 2019

dependence often produce inaccurate results and invalid inference [8,19–21]. In this study, we aimed to address spatial autocorrelation using spatial econometric models to obtain unbiased estimates for the association between CRC mortality and likely county-level determinants within NC.

2. Methods

2.1. Data sources and measures

County-level CRC mortality rates in NC were created using the Compressed Mortality File (CMF) from Centers for Disease Control and Prevention (http://wonder.cdc.gov/wonder/help/cmf.html). Because CRC is most prevalent in older populations and the age to initiate screening is 50 [22], we focused on CRC mortality for such older populations. Due to the available age categories in the CMF (1–4…, 35–44, 45–54, … etc.), we identified people age 45 years and older with deaths due to CRC in 2003–2013. International Classification of Diseases, Tenth Revision codes C18.x, C19, C20, and C26.0 were used to identify CRC deaths. The mortality rates per 100,000 population were age adjusted to the U.S. year 2000 population for each county.

We used the World Health Organization's social determinants of health model [23] to guide our selection of determinants in the analysis. According to the framework, socioeconomic position is the structural determinant that shapes health outcome through healthcare infrastructure and intermediary determinants of health (such as material circumstance, behavioral and biological factors, and psychological factors). In this study, we focused on the following behavioral risk variables as the intermediary determinants: CRC screening, adult obesity rate, and adult smoking rate. Table 1 lists the source and year for county-level variables included in the analysis. All independent variables, except a CRC screening measure, were either from the Area Health Resource File or Robert Wood Johnson Foundation's County Health Rankings & Roadmaps (CHRR, http://www.countyhealthrankings.org/rankings/data/NC). Data from the CHRR were "adjusted" Z-scores (http://www.countyhealthrankings.org/ranking-methods/calculating-scores-and-ranks), in which lower scores indicate healthier or more advantage whereas higher scores indicate less healthy or more disadvantage.

Two variables were used to capture healthcare access and quality. Health Professional Shortage Area (HPSA) data for primary care was used to reflect geographic areas where there are inadequate supplies of primary care health professionals [24]. Hospitalization rates due to ambulatory care sensitive conditions (ACSCs) was used as a proxy for access to adequate and effective care and to reflect quality of care [25–28].

In addition, we used estimates that combined data from two surveys (see Table 1) to measure the percentage of people age 50+ who ever had CRC screening at the county level [29]. We subtracted the value from 100 to reflect the percent who "never had a CRC screening," so the value is consistent with the coding of all variables from CHRR. We also included risk factors of adults who had body mass index greater than $30\,kg/m^2$ (i.e., obesity) and adults who smoked cigarettes.

The variables of ACSC admission, never had CRC screening, and obesity are correlated with each other (Pearson correlation coefficients ranged from 0.38 to 0.46), making it difficult to estimate their independent effects. Following the approach for creating community health indicators from multiple factors [30,31], we created a cumulative effect from the three variables to represent the combined risk of CRC mortality. Specifically, we summed the quintile scores created from each of the three variables and categorized the summed score into three risk groups (termed "risk group", hereafter): low (score 0–4), moderate (score 5–8) and high (score 9–12) risk group.

For socioeconomic characteristics, we chose to use a composite measure, "social and economic factors" from the CHRR, instead of including several individual variables, because these variables tend to be conceptually similar and highly correlated. It should be noted that the socioeconomic variable also included measures of family and social support and community safety. These may represent the psychological component of the WHO's social determinants of health model [23]. For convenience, we named this composite measure as socioeconomic status (SES) deprivation. We also included the following control variables for area-level characteristics: racial and ethnic composition, and population younger than 65 years of age without health insurance.

Table 1
Source and Year of Data for County-level Factors.

County-level variables	Source and Year of Data
Hospital admission for ambulatory care sensitive condition (Z score)	Hospitalization rate for ambulatory care sensitive conditions per 1000 Medicare beneficiaries in 2006-2007: developed by Dartmouth Institute and available on the Robert Wood Johnson Foundation's County Health Rankings & Roadmaps web site (http://www.countyhealthrankings.org/rankings/data/NC).
% Never had CRC[a] endoscopy or did not have FOBT[b] in the past 2 years	Behavioral Risk Factor Surveillance System (BRFSS) and National Health Interview Survey (NHIS) data in 2008-2010, available on State Cancer Profiles web site (https://statecancerprofiles.cancer.gov/)
Obesity (Z score)	Adults had BMI $\geq 30\,kg/m^2$, in 2008: developed by National Center for Chronic Disease Prevention and Health Promotion and available on County Health Rankings & Roadmaps web site (http://www.countyhealthrankings.org/rankings/data/NC).
Cigarette smoking (Z score)	Adults who smoked at least 100 cigarettes in lifetime and were currently smokers, in 2003-2009: developed by BRFSS and available on County Health Rankings & Roadmaps web site (http://www.countyhealthrankings.org/rankings/data/NC).
Socio-economic status deprivation (Z-score)	Composite measure of socioeconomic factors from county health ranking data (http://www.countyhealthrankings.org/), including: education (North Carolina state data in 2008-2009, and American Community Survey in 2005-2009), employment (Bureau of Labor Statistics in 2009), income (Census Small Are Income and Poverty Estimates in 2008), family and social support (BRFSS and American Community Survey in 2005-2009), and community safety (National Center of Health Statistics in 2001-2007).
% non-Hispanic black	Area Health Resource File, in 2008.
% Hispanic	Area Health Resource File, in 2008.
% Native American	Area Health Resource File, in 2008.
% Asian	Area Health Resource File, in 2008.
% Population age < 65 without health insurance	Small Area Health Insurance of the Census in 2008.
Health professional shortage area (partial or whole) indicator	Health professional shortage area for primary care in 2010: from Area Health Resource File. Variable is dichotomized so counties are either "not designated as HPSA" or "the whole or part of the county was designated as a HPSA."
Urbanicity	Rural-urban continuum in 2003, Area Health Resource File.

[a] CRC: colorectal cancer.
[b] FOBT: fecal occult blood test.

These variables had skewed distributions, so we created dichotomous measures using the median as the cutoff point for each of these variables. Finally, we included urbanicity and classified it into three groups: urban counties (counties in metro areas), large town counties (non-metropolitan counties with urban population 20,000 or more), and small town/rural counties (the rest of non-metropolitan and rural areas). This grouping allowed us to separate the counties not in metropolitan areas (N = 60) into two groups; single rural group [7] would have resulted in over 90% of the rural counties being located in an HPSA.

2.2. Analytic approach

Using ArcMap (Esri, Redlands, CA, USA), we first examined global spatial autocorrelation in CRC mortality rates across NC counties. We used local indicators of spatial association (LISA) to identify hotspots [32]. The LISA map produces 4 clusters: high mortality rate area surrounded by other high mortality rate area (high-high clusters), high rate area surrounded by low rate area (high-low clusters), low rate area surrounded by low rate area (low-low clusters), and low rate area surrounded by high rate area (low-high clusters).

To examine determinants of CRC mortality rates, we first ran an ordinary least squares (OLS) model to obtained non-spatial estimates. Using queen continuity weights to define the neighboring structure of NC counties, we evaluated the spatial autocorrelation from Moran's I statistic for the OLS model residuals. A significant result from this test suggests that independent variables in the OLS regression model cannot completely explain the variation and spatial regression approach that incorporates spatial dependence is needed in order to obtain unbiased estimates and correct results.

To address spatial dependence, we examined two commonly used spatial forms: spatial error and spatial lag models [21]. The spatial error model assumes that spatial dependence is from the unmeasured independent variables and it only affects the error terms. Mathematically, this model can be specified as follows. In Eq. 1, Y is the dependent variable, X is the matrix of independent variables, β is the vector of regression parameter. The spatial error model specifies the spatial component on the error structure as in the Eq. 2. The ρ is the spatial autoregressive parameter, W is the spatial weight matrix that describes neighborhood connectivity, and v is an uncorrelated error term.

$$Y = X\beta + e \tag{1}$$

$$e = \rho We + v \tag{2}$$

The spatial lag model, on the contrary, incorporates a diffusion process across space in which the dependent variable is affected by independent variables in the same area and the neighboring areas, and at the same time is influenced by the dependent variable in the neighboring areas. That is, the spatial lag model specifying the spatial component on the dependent variable as shown in the Eq. 3, in which ρ and W are as in the spatial error model, and e is the error term. WY is the spatially lagged dependent variable, which is the averaged dependent variable over the surrounded neighborhood.

$$Y = \rho WY + X\beta + e \tag{3}$$

To select the spatial form that described the underlying structure of CRC mortality in NC, we compared the Lagrange Multiplier (LM) diagnostic tests for both forms. Following the suggestion from Anselin and Rey [21] we selected the model that was more significant in these test statistics as this suggests which form is more appropriate. In addition, we further checked if the selected spatial model sufficiently addressed the spatial autocorrelation using the Anselin-Kelejian test [33].

We explored the main effect model and the two-way interaction effect models between variables of risk group, SES deprivation, racial/ethnic composition, and urbanicity. Only the model with risk group and SES deprivation interaction had significant interaction terms. The model fit of this interaction effect model was also better than the main effect model (AIC was reduced from 667 to 659). The Anselin-Kelejian test was not significant in the interaction effect model (1.484, p = 0.223) but marginally significant in the main effect model (2.963, p = 0.085), suggesting that the interaction effect model sufficiently addressed the spatial autocorrelation. Therefore, we reported the estimates from this interaction model in the results section.

Table 2
Descriptive Statistics.

County-level variables	Mean (sd)	Range
Age adjusted CRC[a] mortality rate per 100,000 population	45.1 (7.8)	29.4, 75.8
% Never had CRC[a] endoscopy or did not have FOBT[b] in the past 2 years	38.2 (6.6)	18.1, 55.7
Hospital admission for ambulatory care sensitive condition (Z score)	0 (0.99)	−2.1, 3.0
Obesity (Z score)	0 (1.0)	−2.2, 2.6
Cigarettes smoking (Z score)	0 (0.9)	−2.7, 2.9
Socio-economic status deprivation (Z-score)	0 (0.28)	−0.7, 0.7
% non-Hispanic black	20.9 (16.4); 18.6 (median)	0.6, 60.9
% Hispanic	5.6 (3.9); 4.5 (median)	1.1, 21.4
% Native American	1.6 (4.8); 0.4 (median)	0.1, 37.9
% Asian	0.9 (0.9); 0.6 (median)	0.1, 5.8
% Population age < 65 without health insurance	19.0 (2.5); 18.5 (median)	13.6, 29.5
Health professional shortage area		
Partial or whole county	N = 70	n/a
No shortage	N = 30	n/a
Urbanicity		
Urban counties	N = 40	n/a
Large town counties	N = 19	n/a
Small town/rural counties	N = 41	n/a
Risk group from ACSC[c] admission. No CRC[a] screening and obesity		
Group 1 (low risk)	N = 35	n/a
Group 2 (moderate risk)	N = 37	n/a
Group 3 (high risk)	N = 28	n/a

[a] CRC: colorectal cancer.
[b] FOBT: fecal occult blood test.
[c] Ambulatory care sensitive condition.

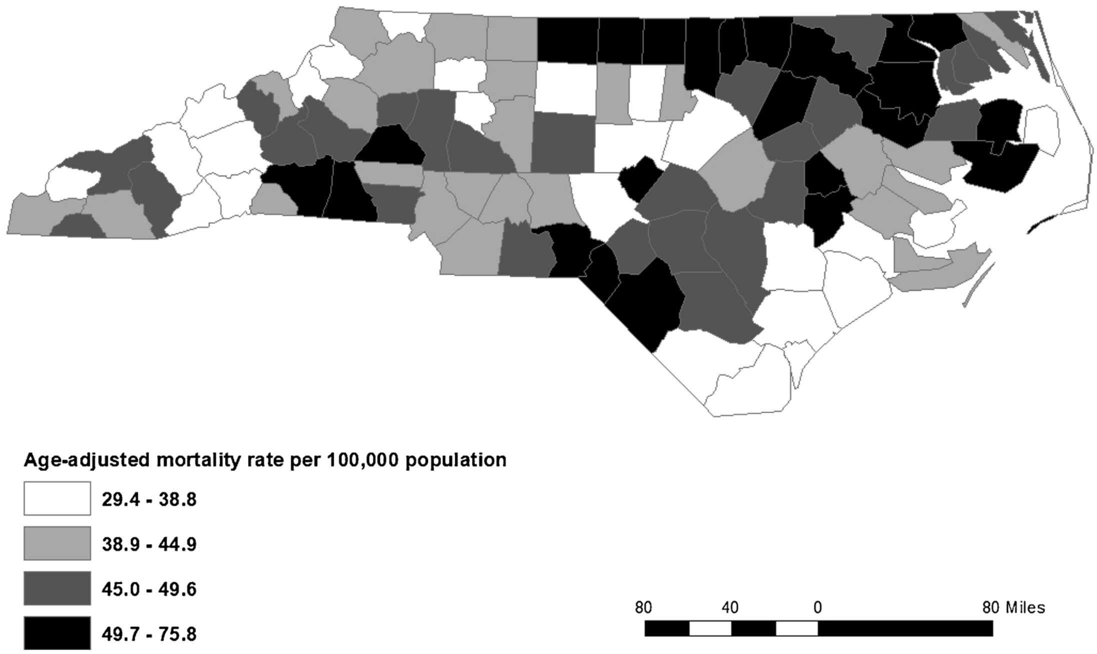

Fig. 1. Colorectal Cancer Mortality Rates: age-adjusted colorectal cancer mortality rates in 2003–2013 in North Carolina counties.

All descriptive statistics were conducted using SAS (9.4, Cary, North Carolina). The OLS and spatial regression models were performed using GeoDaSpace software (https://geodacenter.github.io/GeoDaSpace/).

3. Results

The mean county-level age-adjusted CRC mortality rate was 45.1 per 100,000 population, ranging from 29.4 to 75.8 (Table 2). Spatial distribution of CRC mortality rates showed higher rates in the northeastern areas of the state compared to the other counties (Fig. 1). Global Moran's I was highly significant for spatial autocorrelation (Moran's I = 0.2897, $p < 0.001$), indicating that neighboring counties have similar values. In addition, the LISA map identified the location of high-high clusters in the northeastern counties (Fig. 2).

Table 2 presents summary statistics for all variables included in the analysis. The proportion of people within a county without CRC screening ranged from 18.1% to 55.7%. The Z scores for ACSC admissions, obesity, and cigarette smoking ranged from -2 to +3. The Z scores for SES deprivation were between -0.7 and + 0.7 across counties. The median proportion for race and ethnicity was highest for non-Hispanic black population (18.6%) and lowest for Native Americans (0.4%). Overall, 19% of the population under 65 did not have health insurance. Seventy counties were in an HPSA area for primary care while 40 counties were urban and 41 were small town/rural areas. There were 35 counties in the low CRC risk group and 28 counties in the high risk group.

In the OLS regression model, the Moran's I for residual revealed significant spatial dependence (2.431, $p < 0.05$) confirmed the need

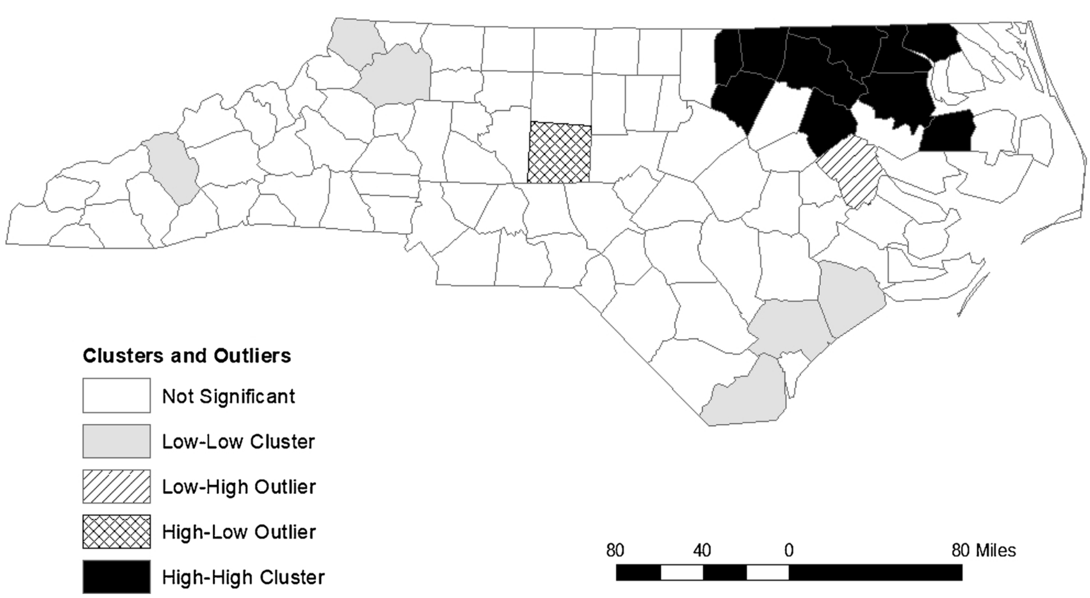

Fig. 2. Clusters of Colorectal Cancer Mortality Rates: local indicators of spatial association (LISA) map of age-adjusted colorectal cancer mortality in 2003–2013 in North Carolina counties.

Table 3
Coefficients and Standard Errors (in Parentheses) from Ordinary Least Squares and Spatial Models.

Variable	Ordinary least squares model	Spatial lag model
Intercept	42. 812 (2.433) ***	13.751 (10.449)
Behavioral risk group		
Moderate vs low	1. 941 (1.848)	1.225 (1.649)
High vs low	1.155 (2.105)	−0.578 (1.954)
Socioeconomic status deprivation	21.783 (5.663)***	21.213 (4.997)***
Interaction terms		
Moderate risk * socioeconomic status deprivation	−18.944 (7.125)**	−21.479 (6.345)***
High risk * socioeconomic status deprivation	−21.739 (7.368)**	−20.753 (6.506)**
Cigarette Smoking	−0.753 (0.708)	−0.175 (0.657)
Health professional shortage area (yes vs no)	1.342 (1.474)	1.824 (1.311)
non-Hispanic black (above vs below median)	3.981 (1.561)*	1.850 (1.567)
Hispanic (above vs below median)	−0.565 (1.446)	−0.136 (1.284)
Native American (above vs below median)	0.379 (1.339)	1.523 (1.248)
Asian (above vs below median)	0.487 (1.551)	0.671 (1.369)
Population age < 65 without health insurance (above vs below median)	−3.832 (1.446)**	−2.513 (1.357)
Urbanicity		
Large town vs urban	4.442 (1.863)*	4.116 (1.647)*
Small town/Rural vs urban	1.570 (1.669)	1.002 (1.485)
Spatial Parameter	NA	0.651 (0.229) **

ACSC: ambulatory care sensitive conditions; CRC: colorectal cancer.
*** $p < 0.001$.
** $p < 0.01$.
* $p < 0.05$.

for a spatial regression model. The LM test was significant for the spatial lag model (5.651, $p < 0.05$) but only marginal for the spatial error model (3.163, $p = 0.075$), suggesting the lag model is more appropriate than the error model. The spatial parameter of the lag model was highly significant (0.651, $p = 0.004$), further confirming the diagnostic tests in the OLS model.

Table 3 presents the results from both the OLS model (non-spatial estimates) and spatial lag model (spatial estimates). Overall, in general, the estimates in the spatial lag model had smaller effects than those in the OLS models. Because the spatial estimates are more accurate than the non-spatial estimates due to the existence of spatial dependence, we focused on the interpretation of the spatial estimates. Significant ($p < 0.05$) results were found for SES deprivation, interaction terms between SES deprivation and risk groups, and large town counties in the spatial lag model.

To interpret the effects of spatial lag estimates, one needs to keep in mind that the CRC mortality rate for a specific county is affected by the independent variables in the same county (direct effect, estimated by β), and by the independent variables in the neighboring counties and their neighbors (indirect effects). The average total effect can be estimated using formula $\beta/(1\text{-}\rho)$ [34,35]. For example, on average, the direct effect of large town counties was 4.116 CRC deaths more per 100,000 population compared to urban counties (Table 3). Given the spatial dependence, the average total effect for a county being a large town was associated with 11.79 more CRC deaths per 100,000 population than that of urban counties.

The significant interaction effects between SES deprivation and risk group make the interpretations for the main effects less meaningful, as the effect of SES deprivation differed by risk group. To directly assess and compare the SES deprivation effect by risk group, we computed the slopes of the SES deprivation effect to estimate the average total effect of SES deprivation for each risk group: 60.78, 0.76 and 1.32, for low, moderate and high risk groups. This indicates that a one standard deviation increase in SES deprivation was associated with, on average, an increase of 61 CRC deaths per 100,000 population for the low risk group, and with a negligible effect in areas where behavioral risk was moderate or high. To better demonstrate these differential effects, we computed the model predicted CRC mortality rates and plotted the prediction by risk group for selected SES deprivation Z scores. Fig. 3 shows that as SES deprivation increased the predicted CRC mortality rates increased, with larger increases in low risk group than moderate and high risk groups (with an increase of 11 people per 100,000 versus 3 and 6, respectively, from SES deprivation Z score -0.25 to 0.25). In addition, the CRC mortality rates were smaller for low risk group as compared to the other risk groups in less SES deprived areas (39.4 for low risk group versus 45.5 and 42.7 for moderate and high risk groups where SES is -0.25). Furthermore, the difference in rates between the risk groups decreased as SES deprivation increased (compared to the data where SES is -0.25, the rate is 50.5 for low risk group versus 48.7 and 49.0 for moderate and high risk groups where SES is 0.25).

4. Discussion

Using NC county-level mortality data from 2003 to 2013, we identified clusters of counties with high CRC mortality in the northeastern area of the state. This finding is consistent with a recent analysis using national data from 2000 to 2009 [2]. We also demonstrated that estimates differed between an OLS model and a spatial lag model, which underscores the importance of accounting for spatial process when the data reveal spatial autocorrelation. This finding confirmed the evidence found in previous studies [8,19–21]. Furthermore, our finding suggests a spatial diffusion process for the CRC mortality rates across NC counties; CRC mortality in one county influenced CRC mortality in the neighboring counties. This process likely worked through shared healthcare resources and informal interactions among individuals in the neighboring areas.

Due to high correlation among ACSC admission, CRC screening, and obesity, we combined these risk factors to represent the cumulative effect of risks on CRC mortality and found the risk effect interacted with SES deprivation in shaping the CRC mortality rates. As SES deprivation increased, the overall CRC mortality rates increased. Furthermore, the differential effect of risk groups on the CRC mortality decreased as SES deprivation increased. When SES deprivation was severe, the mortality rate was high, regardless of the level of risk. We further explored the

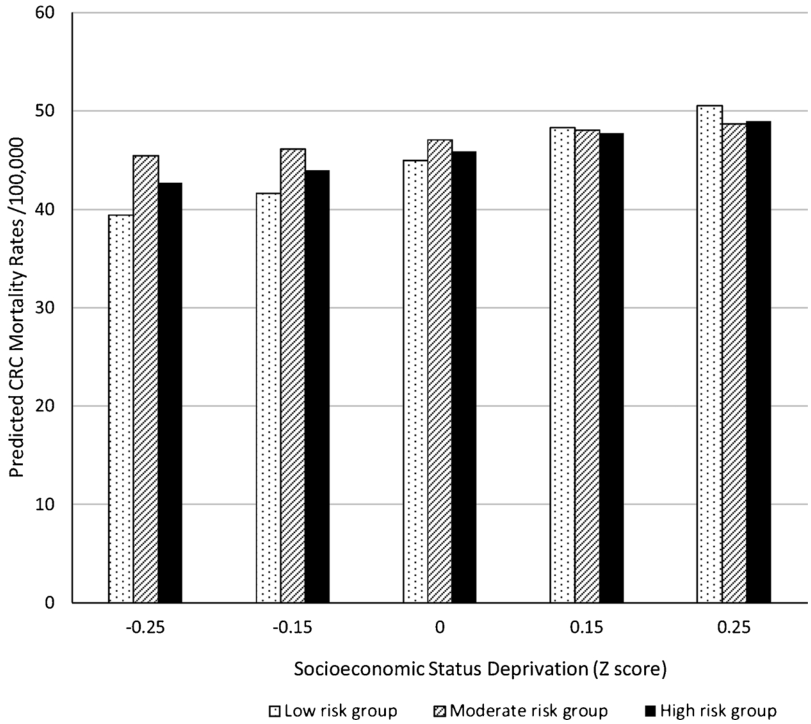

Fig. 3. Spatial Lag Model Predicted Colorectal Cancer Mortality Rates, by Risk Group And Selected Socioeconomic Status Deprivation Scores: model prediction of CRC mortality rates by the three risk groups and selected socioeconomic deprivation scores.

relationship between risk groups and SES deprivation in counties where socioeconomic deprivation was severe ($Z > 0.2$, $N = 23$). Among these counties, 95% were in high (52%) and moderate (43%) risk groups. Five of these counties were also located in the high-high clusters of CRC mortality: Vance, Warren, Halifax, Northampton, and Edgecombe. These counties are the areas in greatest need of intervention. This finding suggests that intervention strategies to reduce CRC mortality for these needed areas should aim to improve multifaceted components—socioeconomic deprivation, low healthcare access and quality, low CRC screening, and high obesity rates. These components are often intertwined with each other in the most needed areas.

In this study we found that the CRC mortality rates were lower in urban counties than that in counties with large non-metropolitan towns, but the rates did not differ between urban and small town/rural counties. This finding is partially consistent with prior studies, in which lower CRC mortality rates were seen in urban counties than both less urban and rural counties [7,10]. It is not clear why we did not find a difference in CRC mortality between urban and small town/rural counties in the regression models. Additional model (data not shown) grouping the non-metro counties into non-metro urban and rural counties and still found no difference between urban and rural counties. Future studies will need to further examine this issue.

Because area-level factors are geographically/regionally specific, our findings may not generalize to other geographic areas. Another limitation of our study is the lack of person-level and tumor level data. This precludes comprehensive analyses to assess the effects from an individual level and the dynamic processes from interactions between individuals and the environment. Unfortunately, there is no state or national data source with complete data on key determinants at both an individual and area-level to allow us to assess geographic variation of CRC mortality. We recognize that without complete data at tumor, person, and area levels, our study is unable to fully distinguish whether the geographic variation is from the characteristics of areas (i.e., contextual) or from the types of individuals living in these areas (i.e., compositional) [36,37]. Until more data become available, a spatial approach that correctly accounts for spatial dependence remains appropriate and valuable to identify determinants at aggregate level to inform population level interventions.

4.1. Conclusion

Results from this study expand our understanding of geographic variation in CRC mortality in NC. Using the spatial lag model our findings depict a spatial diffusion process for the CRC mortality rates across NC counties—CRC mortality rate in a specific county is likely to be influenced by the rates in the neighboring counties. Successful interventions to reduce geographic variation in CRC mortality may need to develop strategies that work through shared resources and informal interactions among individuals living in the vicinity of targeted areas. In addition, the finding of significant interaction effects between SES deprivation and behavioral risk group adds to our understanding of the complicated relationship between CRC mortality and the intertwined effects between area-level SES deprivation and behavioral risks. Finally, we identified clusters of counties with high-high CRC mortality and several of these counties were deprived in multiple factors: socioeconomic deprivation, lack of healthcare access and quality, high obesity rates, and low CRC screening. These areas will be best viewed as priority targets for enhanced surveillance and interventions. These results underscore the importance of developing multifaceted interventions to reduce geographic variation of CRC mortality rates.

Contribution

TM Kuo: design, data processing, analysis, interpretation, writing.
AM Meyer: interpretation of results, discussion, writing.
CD Baggett: interpretation of results, discussion, writing.
AF Olshan: interpretation of results, discussion, writing.
All authors approved the final version of the paper.

Conflicts of interest

None.

Acknowledgements

Work on this study was supported by the Cancer Information and Population Health Resource, UNC Lineberger Comprehensive Cancer Center, with funding provided by the University Cancer Research Fund via the state of North Carolina, USA (A16-0735).

References

[1] American Cancer Society, Key Statistics for Colorectal Cancer (Accessed 29 June 2017), (2018) https://www.cancer.org/cancer/colon-rectal-cancer/about/key-statistics.html.

[2] R.L. Siegel, L. Sahar, A. Robbins, A. Jemal, Where can colorectal cancer screening interventions have the most impact? Cancer Epidemiol. Biomarkers Prev. 24 (8) (2015) 1151–1156, https://doi.org/10.1158/1055-9965.EPI-15-0082.

[3] D. Naishadham, I. Lansdorp-Vogelaar, R. Siegel, V. Cokkinides, A. Jemal, State disparities in colorectal cancer mortality patterns in the United States, Cancer Epidemiol. Biomark. Prev. 20 (7) (2011) 1296–1302.

[4] Health Atlas, North Carolina Cancer Mortality Maps (Accessed 22 June 2017, http://www.schs.state.nc.us/data/hsa/cancer.htm.

[5] J.S. Haas, P. Brawarsky, A. Iyer, G.M. Fitzmaurice, B.A. Nevill, C. Earle, Association of area sociodemographic characteristics and capacity for treatment with disparities in colorectal cancer care and mortality, Cancer 117 (18) (2011) 4267–4276, https://doi.org/10.1002/cncr.26034.

[6] N. Breen, D.R. Lewis, J.T. Gibson, M. Yu, S. Harper, Assessing disparities in colorectal cancer mortality by socioeconomic status using new tools: health disparities calculator and socioeconomic quintiles, Cancer Causes Control 28 (2017) 117–125, https://doi.org/10.1007/s10552-016-0842-2.

[7] G.K. Singh, S.D. Williams, M. Siahpush, A. Mulhollen, Socioeconomic, rural-urban, and racial inequalities in US cancer mortality: part I-all cancers and lung cancer and part II-colorectal, prostate, breast, and cervical cancers, J. Cancer Epidemiol. (2011) 1–27, https://doi.org/10.1155/2011/107497.

[8] P.D. Baade, P. Dasgupta, J.F. Atken, G. Turrell, Geographic remoteness, area-level socioeconomic disadvantage and inequalities in colorectal cancer survival in Queensland: a multilevel analysis, BMC Cancer 13 (2013) 493.

[9] J.M. Ferrante, J.H. Lee, E.P. McCarthy, K.J. Fisher, et al., Primary care utilization and colorectal cancer incidence and mortality among Medicare beneficiaries: a population-based, case-control study, Ann. Intern. Med. 159 (2013) 437–446.

[10] K.D. Blake, J.L. Moss, A. Gaysynsky, et al., Making the case for investment in rural cancer control: An analysis of rural cancer incidence, mortality, and funding trends, Cancer Epidemiol. Biomarkers Prev. (2017) 1–6, https://doi.org/10.1158/1055-9965.

[11] B. Levin, D.A. Lieberman, B. McFarland, K.S. Andrews, D. Brooks, et al., Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology, Gastroenterology 134 (5) (2008) 1570–1595.

[12] M. Pignone, M. Rich, S.M. Teutsch, A.O. Berg, K.N. Lohr, Screening for colorectal cancer in adults at average risk: a summary of the evidence for the U.S. Preventive Services Task Force, Ann. Intern. Med. 137 (2) (2002) 132–141.

[13] W.S. Atkin, R. Edwards, I. Kralj-Hans, K. Wooldrage, et al., Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial, Lancet 375 (2010) 1624–1633.

[14] J.S. Mandel, J.H. Bond, T.R. Church, D.C. Snover, et al., Reducing mortality from colorectal cancer by screening for fecal occult blood: Minnesota colon cancer control study, N. Engl. J. Med. 328 (1993) 1365–1371.

[15] R.E. Schoen, P.F. Pinsky, J.L. Weissfeld, L.A. Yokochi, et al., Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy, N. Engl. J. Med. 366 (2012) 2345–2357.

[16] A.G. Zauber, The impact of screening on colorectal cancer mortality and incidence: has it really made a difference? Dig. Dis. Sci. 60 (2015) 681–691.

[17] A. Chao, M.J. Thun, E.J. Jacobs, et al., Cigarette smoking and colorectal Cancer mortality in the Cancer prevention study II, J. Natl. Cancer Inst. 92 (23) (2000) 888–1896.

[18] A. Shaukat, A. Dostal, J. Menk, T.R. Church, BMI is a risk factor for colorectal cancer mortality, Dig. Dis. Sci. 62 (2017) 2511–2517, https://doi.org/10.1007/s10620-017-4682-z.

[19] D.K. McLaughlin, C.S. Stokes, P.J. Smith, A. Nonoyama, Differential mortality across the United States: the influence of place-based inequality, in: L.M. Lobao, G. Hooks, A.R. Tickamyer (Eds.), The Sociology of Spatial Inequality, State University of New York Press, Albany, NY, 2007, pp. 141–162.

[20] P.J. Sparks, C.S. Sparks, An application of spatially autoregressive models to the study of US county mortality rates, Popul. Space Place 16 (2010) 465–481.

[21] L. Anselin, S. Rey, Properties of tests for spatial dependence in linear regression models, Geogr. Anal. 23 (2) (1991) 112–131, https://doi.org/10.1111/j.1538-4632.1991.tb00228.x.

[22] U.S. Preventive Services Task Force, (2016) (accessed 7 May 2017, http://annals.org/aim/fullarticle/715444/screening-colorectal-cancer-adults-average-risk-summary-evidence-u-s.

[23] O. Solar, A. Irwin, A Conceptual Framework for Action on the Social Determinants of Health. Social Determinants of Health Discussion. Paper 2 (Policy and Practice), World Health Organization, Geneva, 2010.

[24] Health Resources and, Services Administration. Shortage Designation: Health Professional Shortage Areas, Available at (2010) (Accessed 17 May 2017), http://www.hrsa.gov/shortage/.

[25] J. Basu, B. Friedman, H. Burstin, Primary care, HMO enrollment, and hospitalization for ambulatory care sensitive conditions: a new approach, Med. Care 40 (2002) 1260–1269.

[26] A.B. Bindman, K. Grumbach, D. Osmond, et al., Preventable hospitalizations and access to health care, JAMA 274 (1995) 305–311.

[27] J.N. Laditka, S.B. Laditka, J.C. Probst, More may be better: evidence of a negative relationship between physician supply and hospitalization for ambulatory care sensitive conditions, Health Serv. Res. 40 (2005) 1148–1166.

[28] A.J. Trachtenberg, N. Dik, D. Chateau, A. Katz, Inequities in ambulatory care and the relationship between socioeconomic status and respiratory hospitalizations: a population-based study of a Canadian city, Ann. Fam. Med. 12 (5) (2014) 402–407.

[29] T.E. Raghunathan, D. Xie, N. Schenker, et al., Combining information from two surveys to estimate county-level prevalence rates of cancer risk factors and screening, J. Am. Stat. Assoc. 102 (478) (2007) 474–486, https://doi.org/10.1198/016214506000001293.

[30] L.S. Pointer, Z. Al-Qurayshi, D.T. Pointer, et al., Community health indicators associated with outcomes of pancreatectomy, Am. J. Surg. 215 (2018) 120–124.

[31] J.D. Schold, L.D. Buccini, M.W. Kattan, et al., The association of community health indicators with outcomes for kidney transplant recipients in the United States, Arch. Surg. 147 (2012) 520–526.

[32] L. Anselin, Local indicators of spatial association—LISA, Geogr. Anal. 27 (2) (1995) 93–115, https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

[33] L. Anselin, H.H. Kelejian, Testing for spatial error autocorrelation in the presence of endogenous regressors, Int. Reg. Sci. Rev. 20 (1997) 153–182.

[34] L. Anselin, S.J. Rey, Modern Spatial Econometrics in Practice: a Guide to GeoDa, GeoDaSpace and PySAL, GeoDa Press LLC, Chicago, IL, 2014.

[35] C.W. Kim, T.T. Phipps, L. Anselin, Measuring the benefits of air quality improvement: a spatial hedonic approach, J. Environ. Econ. Manage. 45 (2003) 24–39.

[36] A.V. Diez Roux, Investigating neighborhood and area effects on health, Am. J. Public Health 91 (11) (2001) 1783–1789.

[37] S.V. Subramanian, Multilevel modeling, in: M.M. Fisher, A. Getis (Eds.), The Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer, Berlin, 2010.