

RESEARCH ARTICLE

Open Access



Spatial analysis of colorectal cancer outcomes and socioeconomic factors in Virginia

Esther J. Thatcher^{1*} , Fabian Camacho², Roger T. Anderson², Li Li³, Wendy F. Cohn², Pamela B. DeGuzman⁴, Kathleen J. Porter² and Jamie M. Zoellner²

Abstract

Background: Colorectal cancer (CRC) disparities vary by country and population group, but often have spatial features. This study of the United States state of Virginia assessed CRC outcomes, and identified demographic, socioeconomic and healthcare access contributors to CRC disparities.

Methods: County- and city-level cross-sectional data for 2011–2015 CRC incidence, mortality, and mortality-incidence ratio (MIR) were analyzed for geographically determined clusters (hotspots and cold spots) and their correlates. Spatial regression examined predictors including proportion of African American (AA) residents, rural-urban status, socioeconomic (SES) index, CRC screening rate, and densities of primary care providers (PCP) and gastroenterologists. Stationarity, which assesses spatial equality, was examined with geographically weighted regression.

Results: For incidence, one CRC hotspot and two cold spots were identified, including one large hotspot for MIR in southwest Virginia. In the spatial distribution of mortality, no clusters were found. Rurality and AA population were most associated with incidence. SES index, rurality, and PCP density were associated with spatial distribution of mortality. SES index and rurality were associated with MIR. Local coefficients indicated stronger associations of predictor variables in the southwestern region.

Conclusions: Rurality, low SES, and racial distribution were important predictors of CRC incidence, mortality, and MIR. Regions with concentrations of one or more factors of disparities face additional hurdles to improving CRC outcomes. A large cluster of high MIR in southwest Virginia region requires further investigation to improve early cancer detection and support survivorship. Spatial analysis can identify high-disparity populations and be used to inform targeted cancer control programming.

Keywords: Colorectal neoplasms, Incidence, Risk factors, Spatial regression, Virginia

Background

Globally and in the United States (U.S.), colorectal cancer (CRC) is the third most common cause of cancer incidence and the second highest cause of cancer death [1–3]. Epidemiological patterns and risk factors vary

within and among countries and regions of the world. Lifestyle-based risk factors such as diet and smoking vary greatly, as well as preventive measures such as screening test access and mass education campaigns [3]. Disparities in CRC refer to population subgroups that experience worse outcomes due to social, economic, other factors that impede their access to preventive services or to practice healthy lifestyles [4]. Though the

* Correspondence: esther.thatcher@uhhospitals.org

¹University Hospitals, Cleveland, OH, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

characteristics and risk factors of groups experiencing CRC disparities can vary by place, studying and measuring disparities is important for informing efforts to improve health equity.

In the U.S., disparities in CRC incidence and mortality are associated with population characteristics, such as racial and ethnic minority status, low socioeconomic status (SES), rural residence, and lack of access to healthcare [5, 6]. African Americans (AA) have higher rates of cancer-related mortality than whites; this gap has grown even with mortality rates decreasing overall [7]. Populations residing in low SES counties (geographic state subdivisions) have significantly higher CRC mortality than middle or high SES counties [5]. Rural-dwelling populations also have higher CRC incidence than urban populations across most of the U.S., with mortality rates falling at a slower rate [8]. Improved CRC screening rates for CRC is the leading factor in CRC mortality reductions [9]. However, lack of access to primary care providers (PCPs) and gastroenterologists (GIs) is associated with lower screening rates [10, 11].

In the U.S. state of Virginia, the CRC incidence and mortality rates (36.0 and 14.0 per 100,000, respectively) are slightly more favorable than national estimates (39.2 and 14.5 per 100,000, respectively) [12]. However, CRC county-level rates reveal wide variability, with incidence rates ranging from 22.7 to 70.0 per 100,000 and mortality rates from 8.2 to 32.5 per 100,000 [13]. Nationally, county-level incidence ranges 12.4 to 138.6 per 100,000 and mortality ranges 5.0 to 45.1 per 100,000 [12]. State-wide variations in the rates of CRC risk factors highlight the complexities of understanding their contributions to disease disparities. For example, the Appalachian southwest region of Virginia is very rural and has high poverty rates but has a low proportion of AA residents [14]. In contrast, the southeast region has higher densities of AA residents with a wide mix of rural-urban and SES status [14]. Previous studies have applied spatial analysis to understand how these underlying factors impact disparities and outcomes for prostate and breast cancer in Virginia [15, 16]. However, no known spatial analysis study has focused on CRC disparities in Virginia.

In addition to age-adjusted incidence and mortality rates, the mortality-incidence ratio (MIR) is a useful indicator possible cancer disparities. As a simple calculation of mortality rate divided by incidence rate from a simultaneous time period, a higher MIR indicates that more disease-related deaths occurred even if incidence remained steady [17]. MIR is not a direct measure of the likelihood that an individual diagnosed with cancer will survive [18]. However, differences in MIR across a population or region can indicate lack of healthcare access for cancer screening to find cancer in early stages, lack of facilities to treat cancer, or less effective survivorship care

[19]. For example, in one study the amount of access to primary care through federally qualified health centers was found to significantly predict the MIR observed for breast, cervical, and prostate cancers, though non-significant results were found for CRC [20]. However, a multi-nation study found MIR for CRC to be predicted by healthcare access and quality [17].

Hotspot analysis is a method to identify regional disparities by detecting spatial clusters in which adjacent locations share high values, such as cancer incidence. Going beyond simply ranking high values by locality, hotspot analysis identifies regional patterns with multiple adjacent localities where disparities may be harder to overcome without traveling long distances. Hotspot analysis has been used internationally, such as a Brazilian study of CRC and an Iranian study of breast and prostate cancer [21, 22]. Examples from the U.S. include a U.S. national study to identify county clusters of high mortality rates [6] and a state-level study in Florida to identify areas with high rates of late-stage diagnoses [23].

Spatial regression and geographically weighted regression (GWR) provide a unique means of understanding associations between variables. In Tobler's first law of geography, "everything is related to everything else, but near things are more related than distant things" [24]. In addition to identifying clusters, these spatial methods can also test for "non-stationarity" or the extent that relationships between variables, such as poverty and cancer incidence, vary geographically, though not necessarily by proximity [25]. Investigating spatial non-stationarity through analyses, such as GWR, can better identify the relative importance of each factor in a given region [26]. Spatial regression and GWR have been used to analyze disparities in CRC in international settings [27–29] and other cancer types in U.S. states and metropolises [15, 26, 30]; however, few such analyses have focused on CRC in the U.S.

The aim of this study is to assess CRC disparities in Virginia by (1) identifying locality-level clusters of high CRC incidence, mortality, and MIR and (2) describing spatial dynamics of demographic, socioeconomic, and healthcare access contributors to CRC disparities. This study can inform public health practitioners and healthcare administrators about populations and regions that could benefit from targeted CRC prevention, detection, and treatment programs.

Methods

Data sources and variables

Data at the county ($n = 95$) or city ($n = 38$) level were used in this study. Virginia's independent cities have administrative boundaries separate from counties and are considered equivalent to counties in the U.S. Census [31]. The range in population for counties and cities is

2212 to 1,148,433 and 3936 to 450,435, respectively [32]. For the purposes of this study, the term “locality” is used to comprise both counties and independent cities.

Included in the analysis are three variables reflecting CRC outcomes: (1) incidence [12], (2) mortality [12], and (3) MIR [19]. Six independent variables included: (1) proportion of AA residents [32], (2) rural-urban status [33], (3) SES index [34], (4) CRC screening rate [35], (5) PCP density [36], and (6) GI density [36]. Table 1 provides definitions, data sources, and details for the variables. The most recent years of CRC data, 2011–2015, were used. Data for other variables were selected to overlap in timeframe with the CRC data when possible. However, for SES index the 2017 dataset was the closest available year [34], and CRC screening data were available only for 2008–2010 [35]. Data for independent variables were available for all 133 localities.

In all analyses, models were fit using complete case analysis, where localities with missing data ($n = 6$ for incidence, $n = 30$ for mortality, and $n = 30$ for MIR) for incidence, mortality, or MIR were excluded from their respective analyses. The cause of missing data was

suppression of low counts for localities in the dataset, due to concerns for confidentiality and reliability [12].

Getis-Ord G_i^* hotspot detection

A hotspot analysis was performed to identify spatial clustering of localities with high or low values for incidence, mortality, and MIR. These spatial clusters were calculated using the Getis-Ord G_i^* statistic (ArcGIS 10.4 software), which produces a G_i^* statistic (z-score) and p -value assessing high (i.e., hotspot) or low (i.e., cold spot) spatial clustering, as described by Getis and Ord [37]. A pseudo-Plate Carrée linear map projection was used. Two parallel analyses compared results from the sample of Virginia localities only and a sample that included Virginia localities plus adjoining counties in bordering states. Since the results and interpretations were remarkably similar, only the results for the Virginia localities are presented here.

A False Discovery Rate (FDR) was applied to correct for multiple testing [38] to ensure that the rate of false positives in all tests chosen as significant would be at or

Table 1 Variable definitions, sources and measures and characteristics of the Virginia localities

Variable Name	Variable Type	Data Source; details	Measure	Mean [Minimum, Maximum]
Incidence	Outcome	State Cancer Profiles 2011–2015 [12]; annual estimate of invasive cancer diagnoses with 95% confidence interval. Age adjusted incidence based on 2000 US standard population.	Annual CRC diagnoses per 100,000 population	39.4 (22.7,70.0)
Mortality	Outcome	State Cancer Profiles 2011–2015 [12]; annual estimate of CRC attributable deaths with 95% confidence interval. Age adjusted incidence based on 2000 US standard population.	Annual CRC deaths, per 100,000 population	16.5 (8.2,32.5)
Mortality-incidence ratio (MIR)	Outcome	CRC mortality rate divided by incidence rate for each locality for which both data points were available	Ratio expressed as decimal	0.43 (0.26,1.08)
AA population	Independent	Racial category in US Census data 2010 [32]	Percent of population	18.9 (0.10,79.10)
Rural-urban status	Independent	US Department of Agriculture Economic Research Service 2013 [33]; Rural Urban Continuum Codes based on urban population within locality and adjacency to other metropolitan areas	Ordinal scores 1–9; 1–3 are metropolitan, 4–9 are non-metro	3.72 (1,9)
SES index	Independent	Appalachian Regional Commission 2017 [34]; composite index of unemployment, per capita income, poverty.	Index value based on 100 as US average; lower index indicates higher SES	101.6 (47.2,166.7)
CRC screening rate	Independent	Small Area Estimates 2008–2010 [35]; estimates based on Behavioral Risk Factor Surveillance Survey question on self-reported previous screening.	Percent of population	63.1 (33.2,88.8)
PCP density	Independent	HRSA Area Health Resource Files 2011–2015 [36]; count for each year by locality	PCPs per 100,000 population; sum of PCPs for each year, divided by locality population for each year, then multiplied by 100,000	61.9 (0,323.2)
GI density	Independent	HRSA Area Health Resource Files 2011–2015 [36]; count for each year by locality	GIs per 100,000 population; sum of GIs for each year, divided by locality population for each year, then multiplied by 100,000	2.7 (0,43.3)

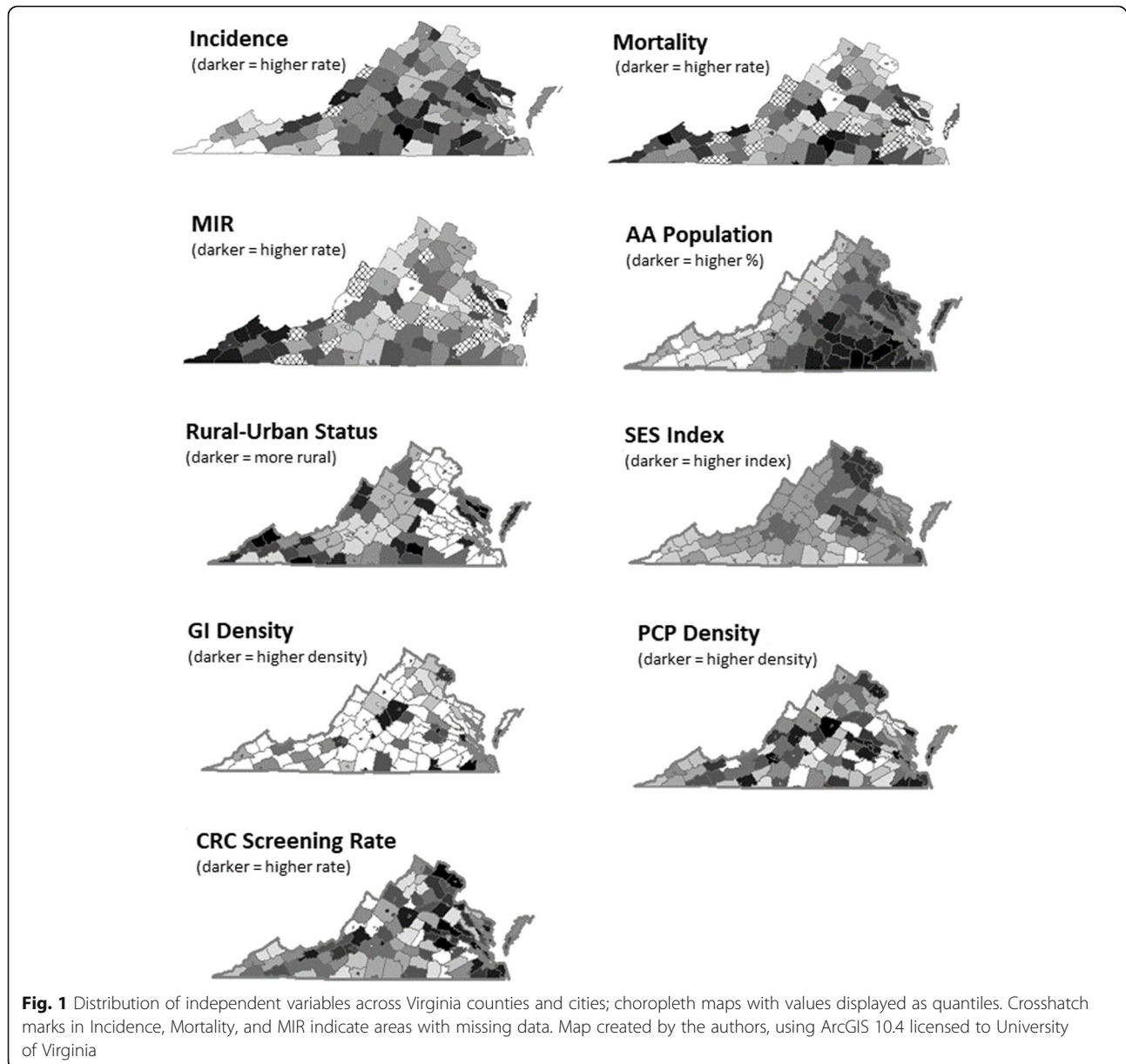
lower than a set level. An FDR of 10% was used in the analysis. Although this criterion may not be as stringent as an FDR criterion of 5%, it is a more conservative approach than choosing the hotspots based on unadjusted *p*-values [39].

Multivariate and spatial regression

Regression analyses were used to assess the contributions of the independent variables to the CRC outcomes. A linear regression was run as a preliminary step, to provide residuals for a test of spatial autocorrelation prior to spatial regression. The two spatial analyses, spatial regression and GWR, incorporated spatial autocorrelation into their models.

In the multivariate linear analyses, the three outcome variables (incidence, mortality, and MIR) were treated separately as a function of five predictors: proportion of AA residents, rural-urban status, SES index, CRC screening rate, and PCP density. The GI density variable was dropped from the analysis to avoid collinearity [40] because its high correlation (~ 0.80) with PCP density.

The residuals in the linear regression were then analyzed for spatial autocorrelation. Moran’s I and Geary’s C tests were performed to determine if spatial autocorrelation was present [41]. Clustering statistics suggested spatial autocorrelation was present for incidence (Geary’s C *p* = 0.0137, Moran’s I < 0.0001) and highly detectable for MIR (Geary’s c and Moran’s I *p* < 0.0001).



As such, the regression model was extended to incorporate spatial autocorrelation. PROC MIXED in SAS was used to specify the covariance matrix, using the syntax described in Xu et al. [41], and using as geographical inputs the county centroid latitude/longitudes [42].

Next, GWR was used to assess for spatial non-stationarity between variables by testing whether the regression coefficients of the linear regression vary as a function of geography (i.e. $\beta(x,y)$ as compared to β , where x,y are geo-coordinates) [43]. A semi-parametric GWR approach, in which the variables with locally varying coefficients were selected while holding other variable coefficients to be constant, was selected as offering the most parsimonious models and simplified interpretation. Locally varying coefficients were identified based on Akaike Information Criterion model fit. Maps display significant findings of local coefficients for each outcome variable (Fig. 3). Colors in the maps represent smoothed values of the non-stationary parameters presented over the geographical region. Red indicates a more positive relationship between the independent and outcome variable, while blue indicates a more negative relationship. In all cases, these parameters were geographically varying regression weights illustrating the local level of association between predictor and outcome.

Results

The statewide distributions of each variable are presented as choropleth displays in Fig. 1. In visual estimation, incidence had higher rates in the central and eastern regions, and pockets of low rates in the northern and southwest regions. Patterns of high mortality were observed in the southwest and south-central regions. For MIR, a high rate was concentrated in the southwest region. The proportion of AA population was highest in the south-central and southeast portions of the state, and lowest in the southwestern areas. With major cities in the northern and eastern parts of the state,

widespread rurality was more pronounced in the southwest and south-central regions, as well as near the eastern seaboard. SES index tended to be lowest in southwest and south-central regions, and highest in northern Virginia. PCP density and CRC screening rates had varying values but without a visible pattern of distribution.

Figure 2 displays hot and cold spots for each of the outcome variables. For incidence, a hotspot was detected in the rural south-central part of the state ($p < 0.10$). Two cold spots were detected, in the far southwest corner ($p < 0.10$) and in the northern area near Washington DC ($p < 0.10$). For mortality, no hotspots or cold spots were detected. The MIR detected a single, large hotspot encompassing the southwest corner ($p < 0.01$).

Table 2 shows the regression parameter estimates between the independent variables and the three CRC outcome variables. Adjusting for all predictors, all three overall models were significantly predictive ($p < 0.001$). In relation to incidence, the only significant single predictor was proportion of AA residents ($p < 0.0001$), such that an increased proportion of AA coincided with increased CRC incidence. For mortality, rural-urban status was a significant predictor ($p = 0.018$), indicating that increased rurality was associated with increased mortality. SES index was also significantly associated ($p = 0.004$), with lower SES index associated with higher mortality. For MIR, the only significant variable in the model was SES index ($p = 0.004$), indicating that lower SES index was associated with higher MIR.

For the spatially correlated regression (Table 2), the overall models for incidence and mortality were significant ($p < 0.001$), while the MIR model was not ($p = 0.142$). For incidence, an increased proportion of AA population ($p = 0.002$) and increased rurality status ($p = 0.018$) had significant positive associations. In terms of mortality, increased rurality status ($p = 0.024$) and lower SES index ($p = 0.007$) had significant positive associations.

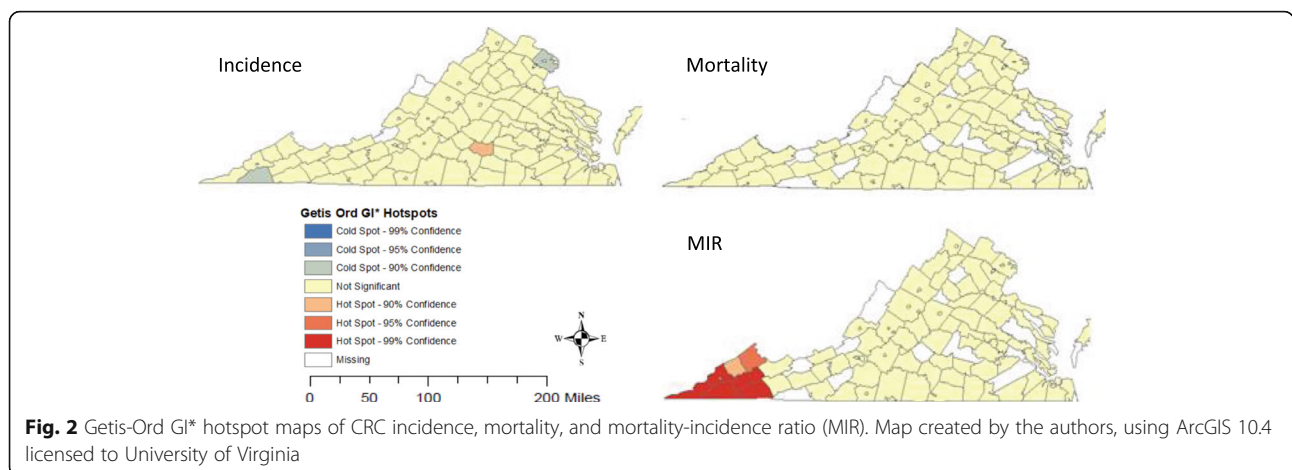


Fig. 2 Getis-Ord GI* hotspot maps of CRC incidence, mortality, and mortality-incidence ratio (MIR). Map created by the authors, using ArcGIS 10.4 licensed to University of Virginia

Table 2 Multivariate and spatially correlated regression model estimates for CRC incidence, mortality, and MIR

	Linear Model with independent errors				Spatially Correlated Model			
	Beta	Std Err	Std Beta	p-value	Beta	Std Err	Std Beta	p-value
Incidence								
Intercept	35.87	8.24	39.3	<.000	28.90	10.32	35.38	0.008
Rural-urban status	0.560	0.334	1.57	0.094	0.700	0.289	1.95	0.018
CRC screening rate	-0.046	0.109	-0.35	0.675	0.020	0.113	0.16	0.859
AA population	0.206	0.049	3.45	<.001	0.230	0.073	3.86	0.002
PCP density	-0.008	0.013	-0.42	0.558	-0.015	0.015	-0.83	0.310
SES index	0.008	0.031	0.24	0.790	-0.008	0.048	-0.23	0.869
Overall Test p-val	< 0.001				< 0.001			
Mortality								
Intercept	10.91	5.184	16.63	0.035	10.91	5.34	16.63	0.044
Rural-urban status	0.398	0.168	1.12	0.018	0.398	0.173	1.12	0.024
CRC screening rate	-0.013	0.064	-0.10	0.834	-0.013	0.068	-0.10	0.842
AA population	0.047	0.025	0.79	0.063	0.047	0.024	0.79	0.050
PCP density	-0.013	0.008	-0.71	0.089	-0.013	0.008	-0.71	0.119
SES index	0.049	0.017	1.44	0.004	0.049	0.018	1.44	0.007
Overall Test p-val	< 0.001				< 0.001			
MIR								
Intercept	0.068	0.141	0.43	0.633	0.329	0.181	0.51	0.085
Rural-urban status	0.006	0.007	0.02	0.360	0.008	0.004	0.02	0.085
CRC screening rate	0.003	0.002	0.02	0.126	0.002	0.002	0.01	0.332
AA population	-0.001	0.001	-0.02	0.129	0.001	0.001	0.02	0.208
PCP density	-0.0002	0.0003	-0.01	0.357	-0.0001	0.0002	-0.01	0.566
SES index	0.002	0.001	0.05	0.004	0.0003	0.001	0.01	0.085
Overall Test p-val	< 0.001				0.142			

Std Err Standard error, Std Beta Standardized beta coefficient

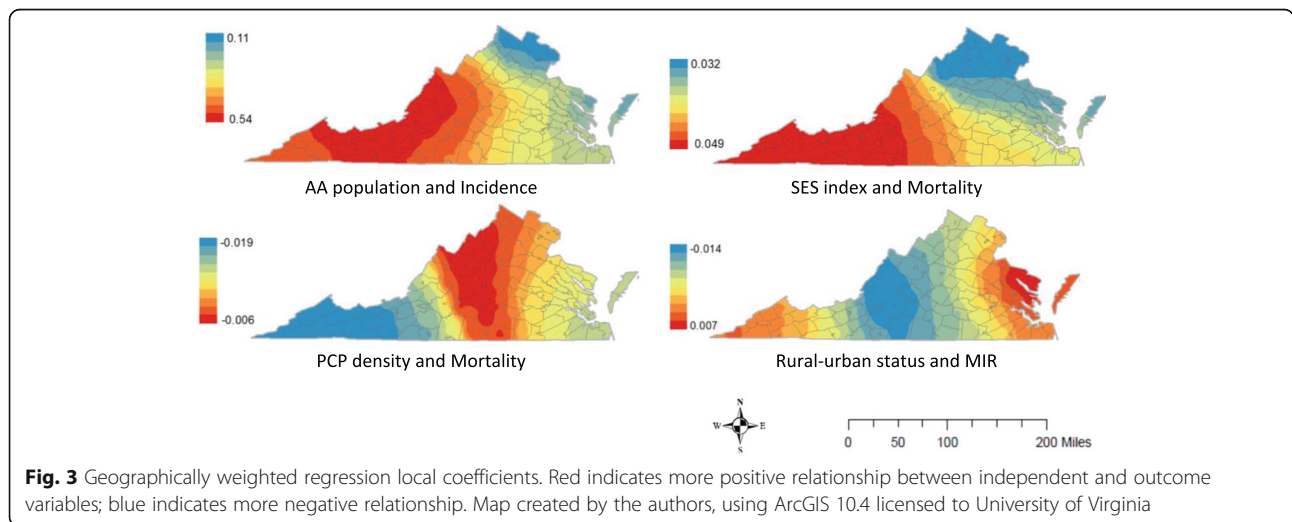


Fig. 3 Geographically weighted regression local coefficients. Red indicates more positive relationship between independent and outcome variables; blue indicates more negative relationship. Map created by the authors, using ArcGIS 10.4 licensed to University of Virginia

Under the GWR model, non-stationarity was observed in the varying strength of association between variables. The association between percentage of AA population and incidence was found to vary geographically but remain positive (GWR local betas from 0.14–0.52, median = 0.24), with a stronger association in southwestern Virginia compared to the rest of the state (Fig. 3). The association of SES index with mortality was found to vary geographically and in the same direction (GWR local betas from 0.032 to 0.049, median = 0.034) such that SES was a stronger predictor of mortality in southwestern Virginia, than other regions. Similarly, the association of PCP density with mortality was found to vary geographically in the same direction of increasing PCP supply with decreasing mortality (GWR local betas from -0.019 to -0.006, median = -0.009). The effect of PCP supply on mortality was strongest in southwestern Virginia and weakest in central Virginia. For rural-urban status, the GWR model detected geographical variation in its association with MIR in varying directions (local betas from -0.014 to 0.007, median -0.002) with positive associations in the eastern seaboard and southwest, and negative associations in central Virginia, suggesting that increasing rurality was associated with increasing MIR in the eastern and southwestern regions of the state but decreasing MIR in the central part.

Summarizing results for incidence, the hotspot in south-central Virginia represents a rural and lower SES region with a high proportion of AA population relative to other regions. The cold spots for incidence were found in two widely contrasting regions including southwest Virginia as rural with high poverty and mostly white residents, and the northern region as urban, more racially diverse and one of the wealthiest parts of the state [14]. Of the independent variables, increasing rurality and proportion of AA residents were positively associated with incidence. Rural-urban status association was significant in the spatial regression, while the proportion of AA residents was significant in the linear and spatial regression model. The GWR model for CRC incidence showed a stronger association for the proportion of AA residents in the western part of the state.

For mortality, no significant hotspots were identified. However, relative to the state average, locality-level mortality rates were generally higher in the southwest region and lower in the northern region. Lower SES index was associated with increased mortality in the linear and spatial regressions. The GWR indicated non-stationarity of this association, showing a stronger association between SES index and mortality in southwestern VA. Rurality was also associated with increased mortality in the linear and spatial regression models. PCP density in association with mortality was found to be non-stationary in the GWR local coefficient, indicating that

an inverse association was stronger in southwestern Virginia and weaker in the central regions.

For MIR, a large and strongly significant hotspot was found in southwestern Virginia. The GWR found the association of rurality and MIR to be non-stationary, with stronger positive associations in the southwestern and eastern edges of the state, and slightly negative associations in the central regions. Low SES was associated with higher MIR in the linear but not the spatial regression.

Discussion

This study found disparities in CRC outcomes in the southwest and south-central regions of Virginia, and that key factors of disparities are rurality, low SES index, and high proportion of AA population. An important finding from our study is that the strength of the association of SES index, race, and PCP density with both incidence and mortality varies across regions of the state. Likewise, rurality was linked to both increased and decreased mortality depending on region. Collectively, findings indicated that many risk factors for CRC incidence or mortality may gain or lose strength on a regional basis, perhaps by combining with other factors or even as a result of community-level interventions. Future studies could assess CRC prevention and control services in relation to risk factors and outcomes.

The MIR hotspot in southwestern Virginia was concerning in terms of potential widespread disparities in access to the continuum of cancer care from detection through treatment. The finding of a cluster of low CRC incidence (a cold spot) in an overlapping location was unexpected and does not reflect other studies finding incidence disparities in rural and especially in Appalachian regions. Previous studies of cancer disparities comparing Appalachian and non-Appalachian populations have found higher overall cancer incidence in Appalachian populations [44, 45] as well as higher CRC incidence [44]. However, a Virginia state cancer report found southwestern health districts had relatively low incidence but high mortality as well as poor rates of detecting cancer at an early stage [13]. Low incidence with high mortality, as denoted by the MIR, could indicate lack of early diagnosis and under-diagnosis, possibly due to low CRC screening rates. Additional assessment is warranted to better understand the dynamics of low incidence rates and high MIR in southwest Virginia. The use of spatial hotspot analysis bolstered evidence that regions with concentrations of social or economic disadvantage may face additional hurdles to improving CRC outcomes. A national CRC mortality study found hotspots occurred in regions characterized by poverty, rurality, and/or high proportion of AA residents [6]. The current study added depth to the understanding of these geographic

disparities through state-level analysis of incidence, mortality, and MIR.

The associations of population characteristics with CRC outcomes reinforced previous studies. Rurality and low SES index were predominant factors in CRC mortality. Both of these population types are characterized by social and structural barriers to cancer prevention and accessing healthcare services along the cancer control continuum [46–48]. The finding of higher incidence among populations with greater proportions of AA residents bears further assessment of local risk factors. However, this finding is also reflective of larger trends of obstacles to healthcare access among AA populations such as lower average SES status than white populations and higher levels of distrust in healthcare organizations [49–51].

The more proximal factors of health care access, PCP density and CRC screening rates, offered additional insights. The finding that PCP density had a stronger negative association with mortality in southwest Virginia compared to the central region is consistent with other studies noting the important role of PCPs in cancer screening and care access, particularly in rural areas where specialists are less accessible [20, 52, 53]. It is notable that CRC screening rate itself was not associated with any of the CRC outcome variables. Previous studies have found strong evidence that screening rates are associated with lower rates of both incidence and mortality and that implementing screening programs results in reductions in these outcomes [44, 54, 55]. Rurality has also been associated with lower screening rates in a national study [56]. The Behavioral Risk Factor Surveillance Survey (BRFSS), the source for screening rates in the current study, might also be a factor. A past study found BRFSS responses overestimated CRC screening rates compared to other surveys [57], and another Virginia-based study using BRFSS prostate cancer screening data also found no association with prostate cancer incidence [16].

The non-stationarity results in this study indicate that a “one-size-fits-all” approach may not be ideal due to variations in the strength of association between population characteristics and CRC outcomes. Rather, these results provide new insight on local or regional solutions that could help improve cancer outcomes in high-disparity parts of the state (Fig. 3). For example, increased access to PCPs in the southwestern region could have a larger effect on CRC mortality. Similarly, the results suggest that programs to improve low SES individuals’ access to early detection and treatment could have greater impact on mortality in the southwestern region. Non-stationarity analysis can also shed light on geographic areas and population characteristics for whom targeted cancer interventions might have greater impact. A previous study found that strength of the association of area-based economic deprivation and later stage

breast tumors was place-specific and was stronger in the Appalachian regions of Pennsylvania than in either Ohio or Kentucky [30]. Other non-stationarity studies include an analysis of U.S. prostate cancer mortality that found urologist availability had stronger association in certain regions [58], and a English study that found regions with stronger associations between low SES and cervical cancer incidence [59]. Identifying associations that have spatial variability is important for selecting data analysis approaches that yield more precise results.

There are limitations in this study. The use of ecological data rather than individual-level data increases the likelihood of erroneous findings, due to aggregations of population data that do not accurately reflect the characteristics or travel patterns of individuals in those localities. Yet, as all the data used in this study were publicly available, this study provides an example for population health programs wishing to conduct spatial analyses but unable to access individual-level data. Available data for CRC screening estimates were older than incidence and mortality data, causing potential errors in analyzing associations. The missing data, particularly for mortality and MIR, may have caused errors in analyses.

Research and practice implications

The spatial methods used in this study can be replicated in other geographic settings to elucidate some of the complex factors of CRC outcomes. Future spatial analysis using individual-level, rather than locality-level, CRC data can add to understanding of statewide populations and regions in Virginia at risk for cancer disparities. Future analyses should also include variables more proximal to cancer outcomes, such as late-stage diagnosis, statewide availability and accessibility of cancer screening, and specialist treatment resources. Further research is needed to understand the low incidence rates in the otherwise health disparate southwest Virginia region, especially whether this is due to under-detection, and to explore the lack of association between CRC screening rates and outcomes.

Practitioners can use this study’s findings to target programs and policies to specific regions and adapt interventions to better meet the needs of high-disparity population groups. For example, programs and policies to lower cost barriers to CRC screening and treatment should be targeted to low SES populations. Programs for CRC prevention and early screening should be tailored and targeted to effectively reach AA populations. Similarly, outreach programs should be tested and adapted for implementation in rural communities [60]. For areas with high MIR, programs and policies should target increased access to PCPs, screening, early detection, and CRC treatment specialists.

Conclusions

This study of the U.S. state of Virginia found that rurality, low SES index, proportion of AA population, and low density of PCPs were associated with CRC outcome disparities. The strength of associations varied spatially, indicating addressing risk factors in areas with stronger associations may yield greater impact on CRC outcomes. Spatial analyses such as hotspot clusters and non-stationarity are important for understanding nuanced attributes of risk factors for CRC.

Abbreviations

AA: African American; CRC: Colorectal cancer; FDR: False discovery rate; GI: Gastroenterologist; GWR: Geographically weighted regression; MIR: Mortality-incidence ratio; PCP: Primary care provider; SES: Socioeconomic status; U.S.: United States

Acknowledgements

Not applicable.

Authors' contributions

All authors made substantial contributions to the conception and design of the work. EJT and FC acquired and analyzed the data. EJT, FC, RA, LL, WC, PD, KP, and JZ interpreted the data, and drafted or substantively revised the work. All authors read and approved the final manuscript.

Funding

The authors received no financial support for the research, authorship, or publication of this article.

Availability of data and materials

Data used in this analysis are available for open public access at: State Cancer Profiles. Quick profiles: Virginia. 2019; Available from: www.statecancerprofiles.cancer.gov. University of Wisconsin Population Health Institute, *County Health Rankings & Roadmaps 2019*, <https://www.countyhealthrankings.org/>. United States Department of Agriculture Economic Research Service. *Rural-urban continuum codes*, <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/>. Appalachian Regional Commission. *County economic status in Appalachia, FY 2017*, <https://www.arc.gov/research/MapsofAppalachia.asp>. National Cancer Institute, *Small area estimates for cancer-related measures: CRC test ever*, <https://sae.cancer.gov/nhis-brfss/estimates/crc.html>. Health Resources & Services Administration, *Area health resources files*, <https://data.hrsa.gov/topics/health-workforce/ahrf>

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹University Hospitals, Cleveland, OH, USA. ²Department of Public Health Sciences, School of Medicine, University of Virginia, Charlottesville, USA. ³Department of Family Medicine, School of Medicine, University of Virginia, Charlottesville, USA. ⁴School of Nursing, University of Virginia, Charlottesville, USA.

Received: 29 April 2021 Accepted: 28 September 2021

Published online: 21 October 2021

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics 2019. *CA Cancer J Clin.* 2019; 69(1):7–34. ISSN 0007-9235. <https://doi.org/10.3322/caac.21551>.
2. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol.* 2019;16(12):713–32. ISSN 1759-5053. <https://doi.org/10.1038/s41575-019-0189-8>.
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–49. ISSN 0007-9235. <https://doi.org/10.3322/caac.21660>.
4. Musselwhite LW, May FP, Salem ME, Mitchell EP. Colorectal cancer: in the pursuit of health equity. *Am Soc Clin Oncol Edu Book.* 2021;41(41):108–17. https://doi.org/10.1200/edbk_321071.
5. Saldana-Ruiz N, Clouston SA, Rubin MS, Colen CG, Link BG. Fundamental causes of colorectal cancer mortality in the United States: understanding the importance of socioeconomic status in creating inequality in mortality. *Am J Public Health.* 2013;103(1):99–104. ISSN 1541-0048. <https://doi.org/10.2105/AJPH.2012.300743>.
6. Siegel RL, Sahar L, Robbins A, Jemal A. Where can colorectal cancer screening interventions have the most impact? *Cancer Epidemiol Biomark Prev.* 2015; 24(8):1151–6. ISSN 1055-9965. <https://doi.org/10.1158/1055-9965.EPI-15-0082>.
7. Rust G, Zhang S, Yu Z, Caplan L, Jain S, Ayer T, et al. Counties eliminating racial disparities in colorectal cancer mortality. *Cancer.* 2016;122(11):1735–48. ISSN 0008-543X. <https://doi.org/10.1002/cncr.29958>.
8. Zahnd WE, James AS, Jenkins WD, Izadi SR, Fogleman AJ, Steward DE, et al. Rural-urban differences in cancer incidence and trends in the United States. *Cancer Epidemiol Biomark Prev.* 2018;27(11):1265–74. <https://doi.org/10.1158/1055-9965.EPI-17-0430>.
9. Centers for Disease Control and Prevention. Vital signs: Colorectal cancer screening, incidence, and mortality—United States, 2002–2010. *MMWR Surveill Summ.* 2011;60(26):884. ISSN 1545-861X.
10. Aboagye JK, Kaiser HE, Hayanga AJ. Rural-urban differences in access to specialist providers of colorectal cancer care in the United States: a physician workforce issue. *JAMA Surg.* 2014;149(6):537–43. ISSN 2168-6254. <https://doi.org/10.1001/jamasurg.2013.5062>.
11. Davis MM, Freeman M, Shannon J, Coronado GD, Stange KC, Guise J-M, et al. A systematic review of clinic and community intervention to increase fecal testing for colorectal cancer in rural and low-income populations in the United States—how, what and when? *BMC Cancer.* 2018;18(1):40. ISSN 1471-2407. <https://doi.org/10.1186/s12885-017-3813-4>.
12. State Cancer Profiles: Quick profiles: Virginia. 2019. www.statecancerprofiles.cancer.gov. Accessed 11 June 2019.
13. Virginia Department of Health. Cancer in Virginia: overview and selected statistics. Richmond; 2016. https://www.vdh.virginia.gov/content/uploads/sites/71/2019/10/VCR_AnnualReport_final.pdf
14. Weldon Cooper Center for Public Service: Virginia's regions. 2014. <https://coopercenter.org/>. Accessed 31 Oct 2019
15. DeGuzman PB, Cohn WF, Camacho F, Edwards BL, Sturz VN, Schroen AT. Impact of urban neighborhood disadvantage on late stage breast cancer diagnosis in Virginia. *J Urban Health.* 2017;94(2):199–210. ISSN 1099-3460. <https://doi.org/10.1007/s11524-017-0142-5>.
16. Oliver MN, Smith E, Siadaty M, Hauck FR, Pickle LW. Spatial analysis of prostate cancer incidence and race in Virginia, 1990–1999. *Am J Prev Med.* 2006;30(2): S67–76. ISSN 0749-3797. <https://doi.org/10.1016/j.amepre.2005.09.008>.
17. Sunkara V, Hébert JR. The colorectal cancer mortality-to-incidence ratio as an indicator of global cancer screening and care. *Cancer.* 2015;121(10): 1563–9. ISSN 0008-543X. <https://doi.org/10.1002/cncr.29228>.
18. Ellis L, Belot A, Rached B, Coleman MP. The mortality-to-incidence ratio is not a valid proxy for cancer survival. *J Global Oncol.* 2019;5(5):1–9. ISSN 2378-9506. <https://doi.org/10.1200/JGO.19.00038>.
19. Buck KD. Modelling of geographic cancer risk factor disparities in US counties. *Appl Geogr.* 2016;75:28–35. ISSN 0143-6228. <https://doi.org/10.1016/j.apgeog.2016.08.001>.
20. Adams SA, Choi SK, Khang L, Campbell DA, Friedman DB, Eberth JM, et al. Decreased cancer mortality-to-incidence ratios with increased accessibility of federally qualified health centers. *J Community Health.* 2015;40(4):633–41. ISSN 0094-5145. <https://doi.org/10.1007/s10900-014-9978-8>.

21. Montazeri M, Hoseini B, Firouzaghi N, Kiani F, Raouf-Mobini H, Biabangard A, et al. Spatio-temporal mapping of breast and prostate cancers in South Iran from 2014 to 2017. *BMC Cancer*. 2020;20(1):1170. <https://doi.org/10.1186/s12885-020-07674-81471-2407>.
22. Moura AR, Marques AD, Dantas MS, de Abreu Costa Brito É, do Rosário Souza M, Lima MS, et al. Trends in the incidence and mortality of colorectal cancer in a Brazilian city. *BMC Res Notes*. 2020;13(1):560. <https://doi.org/10.1186/s13104-020-05411-91756-0500>.
23. Sherman RL, Henry KA, Tannenbaum SL, Feaster DJ, Kobetz E, Lee DJ. Applying spatial analysis tools in public health: an example using SaTScan to detect geographic targets for colorectal cancer screening interventions. *Prev Chronic Dis*. 2014;11:E41. ISSN 1545-1151. <https://doi.org/10.5888/pcd11.130264>.
24. Sui DZ. Tobler's first law of geography: a big idea for a small world? *Ann Assoc Am Geogr*. 2004;94(2):269–77. <https://doi.org/10.1111/j.1467-8306.2004.09402003.x0004-5608>.
25. Comber AJ, Brunsdon C, Radburn R. A spatial analysis of variations in health access: linking geography, socio-economic status and access perceptions. *Int J Health Geogr*. 2011;10(1):44. ISSN 1476-072X. <https://doi.org/10.1186/1476-072X-10-44>.
26. Goovaerts P, Xiao H, Adunlin G, Ali A, Tan F, Gwede CK, et al. Geographically-weighted regression analysis of percentage of late-stage prostate cancer diagnosis in Florida. *App Geog*. 2015;62:191–200 0143–6228.
27. Mansori K, Soleymani-Dodaran M, Mosavi-Jarrahi A, Motlagh AG, Salehi M, Delavari A, et al. Determination of effective factors on geographic distribution of the incidence of colorectal cancer in Tehran using geographically weighted poisson regression model. *Med J Islam Repub Iran*. 2019;33:23. <https://doi.org/10.34171/mjiri.33.231016-1430>.
28. Czaderny K. Increasing deaths from colorectal cancer in Poland - insights for optimising colorectal cancer screening in society and space. *Ann Agric Environ Med*. 2019;26(1):125–32. <https://doi.org/10.26444/aaem/992331898-2263>.
29. Goshayeshi L, Pourahmadi A, Ghayour-Mobarhan M, Hashtarkhani S, Karimian S, Shahhoseini Dastjerdi R, et al. Colorectal cancer risk factors in north-eastern Iran: A retrospective cross-sectional study based on geographical information systems, spatial autocorrelation and regression analysis. *Geospat Health*. 2019;14(2). <https://doi.org/10.4081/gh.2019.7931970-7096>.
30. Anderson RT, Yang TC, Matthews SA, Camacho F, Kern T, Mackley HB, et al. Breast cancer screening, area deprivation, and later-stage breast cancer in Appalachia: does geography matter? *Health Services Res*. 2014;49(2):546–67. ISSN 0017-9124. <https://doi.org/10.1111/1475-6773.12108>.
31. Virginia Association of Counties. Virginia counties. 2019. <https://www.vaco.org/county-profiles/>
32. University of Wisconsin Population Health Institute: County health rankings & roadmaps 2019. www.countyhealthrankings.org
33. United States Department of Agriculture Economic Service: Rural-urban continuum codes. 2019. <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/>
34. Appalachian Regional Commission: County economic status in Appalachia, FY 2017 <https://www.arc.gov/research/MapsofAppalachia.asp>
35. National Cancer Institute: Small area estimates for cancer-related measures: CRC test ever. 2019. <https://sae.cancer.gov/nhis-brfss/estimates/crc.html>
36. Health Resources & Services Administration: Area health resources files. 2019. <https://data.hrsa.gov/data/about>
37. Getis A, Ord JK. The analysis of spatial association by use of distance statistics. In: *Perspectives on spatial data analysis*. Berlin: Springer; 2010. p. 127–45.
38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Statistical Society: Series B (Methodological)*. 1995;57(1):289–300. ISSN 0035-9246. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
39. Caldas de Castro M, Singer BH. Controlling the false discovery rate: A new application to account for multiple and dependent tests in local statistics of spatial association. *Geog Anal*. 2006;38(2):180–208 ISSN 0016–7363.
40. Abu-Bader SH. Using statistical methods in social science research: with a complete SPSS guide: Oxford University press; 2011.
41. Xu Y, Kennedy E. An introduction to spatial analysis in social science research. *Quant Method Psych*. 2015;11(1):22–31. <https://doi.org/10.20982/tqmp.11.1.p022>.
42. Werneck GL, Maguire JH. Spatial modeling using mixed models: an ecologic study of visceral leishmaniasis in Teresina, Piauí state, Brazil. *Cadernos de Saúde Pública*. 2002;18(3):633–7. ISSN 0102-311X. <https://doi.org/10.1590/S0102-311X2002000300007>.
43. Nakaya T, Fotheringham AS, Brunsdon C, Charlton M. Geographically weighted poisson regression for disease association mapping. *Stat Med*. 2005;24(17):2695–717 0277–6715.
44. Wilson RJ, Ryerson AB, Singh SD, King JB. Cancer incidence in Appalachia, 2004–2011. *Cancer Epidemiol Biomark Prev*. 2016;25(2):250–8. <https://doi.org/10.1158/1055-9965.Epi-15-0946>.
45. Yao N, Alcalá HE, Anderson R, Balkrishnan R. Cancer disparities in rural Appalachia: incidence, early detection, and survivorship. *J Rural Health*. 2017;33(4):375–81. ISSN 0890-765X. <https://doi.org/10.1111/jrh.12213>.
46. Zahnd WE, McLafferty SL, Eberth JM. Multilevel analysis in rural cancer control: a conceptual framework and methodological implications. *Prev Med*. 2019;129:105835. ISSN 0091-7435. <https://doi.org/10.1016/j.jypmed.2019.105835>.
47. Hines R, Markossian T, Johnson A, Dong F, Bayakly R. Geographic residency status and census tract socioeconomic status as determinants of colorectal cancer outcomes. *Am J Public Health*. 2014;104(3):e63–71. ISSN 1541-0048. <https://doi.org/10.2105/AJPH.2013.301572>.
48. Baade PD, Dasgupta P, Aitken JF, Turrell G. Geographic remoteness, area-level socioeconomic disadvantage and inequalities in colorectal cancer survival in Queensland: a multilevel analysis. *BMC Cancer*. 2013;13(1):493. <https://doi.org/10.1186/1471-2407-13-4931471-2407>.
49. Adams LB, Richmond J, Corbie-Smith G, Powell W. Medical mistrust and colorectal cancer screening among African Americans. *J Community Health*. 2017;42(5):1044–61. <https://doi.org/10.1007/s10900-017-0339-21573-3610>.
50. Al-Mondhiry J, Wu J. Minding the gap: disparities and opportunities in colorectal cancer screening in the United States. *J Cancer Prev Curr Res*. 2016;5(6):00186. <https://doi.org/10.15406/jcpcr.2016.05.00186>.
51. Burnett-Hartman AN, Mehta SJ, Zheng Y, Ghai NR, McLerran DF, Chubak J, et al. Racial/ethnic disparities in colorectal cancer screening across healthcare systems. *Am J Prev Med*. 2016;51(4):e107–15. <https://doi.org/10.1016/j.amepre.2016.02.0250749-3797>.
52. Kimmick GG, Camacho F, Mackley HB, Kern T, Yao N, Matthews SA, et al. Individual, area, and provider characteristics associated with care received for stages i to iii breast cancer in a multistate region of Appalachia. *J Oncol Pract*. 2015;11(1):e9–e18. ISSN 1554-7477. <https://doi.org/10.1200/JOP.2014.001397>.
53. Martin AN, Hassinger TE, Kozower BD, Camacho F, Anderson RT, Yao N. Disparities in lung cancer screening availability: lessons from Southwest Virginia. *Ann Thorac Surg*. 2019;108(2):412–6. ISSN 0003-4975. <https://doi.org/10.1016/j.athoracsur.2019.03.003>.
54. Li C, Balluz L, Okoro C, Strint T, Lin J, Garvin W, et al. Surveillance of certain health behaviors and conditions among states and selected local areas — behavioral risk factor surveillance system, United States, 2009. *MMWR Surveill Summ*. 2011;60(9):1–250.
55. Levin TR, Corley DA, Jensen CD, Schottinger JE, Quinn VP, Zauberg AG, et al. Effects of organized colorectal cancer screening on cancer incidence and mortality in a large community-based population. *Gastroenterol*. 2018;155(5):1383–91. <https://doi.org/10.1053/j.gastro.2018.07.017>.
56. Cole AM, Jackson JE, Doescher M. Urban–rural disparities in colorectal cancer screening: cross-sectional analysis of 1998–2005 data from the centers for disease Control's behavioral risk factor surveillance study. *Cancer Med*. 2012;1(3):350–6. ISSN 2045-7634. <https://doi.org/10.1002/cam4.40>.
57. Pierannunzi C, Hu SS, 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(1):49. <https://doi.org/10.1186/1471-2288-13-491471-2288>.
58. Yao N, Foltz SM, Odisho AY, Wheeler DC. Geographic analysis of urologist density and prostate cancer mortality in the United States. *PLoS One*. 2015;10(6). <https://doi.org/10.1371/journal.pone.0131578>.
59. Cheng EM, Atkinson PM, Shahani AK. Elucidating the spatially varying relation between cervical cancer and socio-economic conditions in England. *Int J Health Geogr*. 2011;10(1):51. ISSN 1476-072X. <https://doi.org/10.1186/1476-072X-10-51>.
60. Wheeler SB, Davis MM. "Taking the bull by the horns": four principles to align public health, primary care, and community efforts to improve rural cancer control. *J Rural Health*. 2017;33(4):345–9. <https://doi.org/10.1111/jrh.12263>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.