Heterogeneity of Colon and Rectum Cancer Incidence Across 612 SEER Counties, 2000–2014

Pavel Chernyavskiy[1,2,3,*], **Victoria M. Kennerley**[1], **Ahmedin Jemal**[4], **Mark P. Little**[2], **Philip S. Rosenberg**[1]

[1]Biostatistics Branch, National Cancer Institute, DHHS, NIH, Division of Cancer Epidemiology and Genetics, Bethesda, Maryland 20892-9778, USA.

[2]Radiation Epidemiology Branch, National Cancer Institute, DHHS, NIH, Division of Cancer Epidemiology and Genetics, Bethesda, Maryland 20892-9778, USA.

[3]University of Wyoming, Department of Mathematics and Statistics, Laramie, Wyoming, USA.

[4]Surveillance and Health Services Research Program, American Cancer Society, Atlanta, GA.

Abstract

Recent analyses have suggested decreases over time in colorectal cancer incidence at older ages (55 years) but increases at younger ages (20–54 years). Understanding the geographic heterogeneity of incidence facilitates resource allocation for potential interventions and advances our knowledge of differential etiologies for these cancers. We performed age-period-cohort analysis using 2000–2014 county-level incidence from the Surveillance, Epidemiology, and End Results (SEER) database, estimating relative risk (RR) and age-adjusted annual percent change (Net Drifts) simultaneously for 612 counties via a hierarchical model, separately for colon and rectum cancer, stratified by age group (20–54 vs. 55–84). We also studied correlates of RR and Net Drift with various county-level characteristics. In all SEER counties, colon and rectum cancer incidence rates increased at ages 20–54, whereas rates decreased at ages 55–84. There was marked heterogeneity in both RR and Net Drift among states and counties for both cancer types. Maps of county RR and Net Drift revealed localized clusters in several states. For both cancer types, counties with high RR and unfavorable Net Drift tended to have higher prevalence of obesity and diabetes and to be of a lower socioeconomic status. Counties with higher overall screening rates tended to have lower Net Drifts for both cancer types. Increasing colorectal cancer incidence in the younger age group is geographically widespread, although there is significant heterogeneity in temporal trends and risk both within and between states. These geographic patterns correlate with different county-level characteristics depending on cancer type and age group.

Keywords

*Corresponding author contact information: Pavel Chernyavskiy, Department of Mathematics and Statistics, Laramie, WY 82071, pchern1@gmail.com, (240)276-7360.

Introduction

Colorectal cancer (CRC) is the fourth-most common malignancy in the United States (US) and the second leading cause of cancer death. Recent analyses[1–5] show that CRC incidence is decreasing in the older, heavily-screened age-group of 55 – 84 years, but increasing among 20–54-year-olds who have relatively little screening. Indeed, increasing incidence among young, predominantly unscreened men and women has prompted discussion about revising existent screening recommendations for those at average risk to start under the age of 50[6,7]. Developing a better understanding of the geographic heterogeneity of CRC incidence will help inform such discussions.

A key etiological question is whether the rising incidence in younger age groups and the declining incidence in older age groups is occurring to the same extent across the US. At the same time, the prevalence of several established CRC risk factors, such as smoking, obesity, unhealthy dietary patterns, and sedentary lifestyle are known to vary geographically. It is unknown whether variation in known risk factors can explain an appreciable proportion of the spatial heterogeneity of CRC incidence. Finally, learning about geographic disparities in incidence can inform resource allocation for potential local and national interventions.

In this study, we analyze CRC incidence in the 612 counties included in the Surveillance, Epidemiology, and End Results (SEER) database using hierarchical age-period-cohort (APC) models. This approach allows us for the first time to quantify spatial heterogeneity in CRC risk and trends over time between states and counties. We develop four separate models for strata defined by cancer type (colon vs. rectum) and age group (20–54 vs. 55–84) to stratify by screening guidelines. We also perform a correlation analysis between county-level risk factors and model-based estimates to ascertain which county characteristics can help explain the geographic patterns of risk and age-adjusted trends over time.

Materials and Methods

Study Design and Data Source

We conducted a retrospective cohort analysis of CRC patients aged 20–54 and 55–84, diagnosed from 2000 through 2014, ascertained using the 18-registry SEER database[8] that covers 28% of the US population. Excluding the Alaska Native Tumour Registry, we analyzed cases from the 17 remaining registries covering 612 counties in 12 states. SEER assigns cases to locations according to each patient's county of residence at the time of diagnosis. In each county, colon and rectum cases and person-years at risk were grouped into 5-year age and calendar period intervals, resulting in 7 age groups (20–24,25–29,…,50–54) for the younger population, and 6 age groups (55–59,60–64,…,80–84) for the older population, each with three 5-year calendar period groups (2000–2004,2005–2009,2010–2014). Thus, incidence rates for the younger population consisted of 21 data cells per county, and rates for the older population consisted of 18 data cells per county. In all, for both colon and rectum cancer, we had 12,852 observations ($612 \times 7 \times 3$) for the younger population and 11,016 observations ($612 \times 6 \times 3$) for the older population.

Statistical Methods

We modeled incidence rates separately for the two types of cancer (colon and rectum) in each age group (ages 20–54 and 55–84) using APC models. These models were fit to all counties simultaneously using a Generalized Linear Mixed Models approach[9] with state, and county-within-state random effects to account for the hierarchical structure of the data. This modeling approach stabilizes estimates in sparsely-populated areas by "borrowing information" across SEER states and across counties within each state. By using random effects, we allow each county to have its own mean rate for the reference age and birth cohort, its own longitudinal age trend (LAT) to measure how quickly rates change with age, and its own age-adjusted annual percent change (Net Drift) to measure how quickly rates change over time. Using the county-specific mean incidence rates, we computed county-level incidence relative risk (RR) for the reference age and birth cohort, relative to the SEER average. We tested each random effects standard deviation using the Bayes Factor[10] to further quantify evidence in favor of geographic heterogeneity. Overall and within-state goodness-of-fit was assessed using standard posterior distribution predictive checks[13] (Supplementary Figures 1–8). All models were estimated using the brms package[11] in R 3.4.0[12]; further details are given in Supplementary Methods.

Ecological Correlation Analysis—To better understand potential correlates of county-level RR and Net Drift, we performed a post-estimation correlation analysis (Table 2). County characteristics included: screening rates, racial/ethnic composition, measures of socioeconomic status, measures of the food environment, and prevalence of obesity and diabetes. Variables and data sources are described in detail in Supplementary Table 1. To arrive at the correct inference about the strength and direction of correlation, we did not use standard correlation coefficients, but instead employed a partial R^2 approach[14] that accounts for the nested structure of counties within states. This provides an interpretable SEER-wide measure of association (R_β^2) for each county characteristic with a minimum value of 0 and a maximum value of 1. Further details are given in Supplementary Methods.

Results

SEER-wide CRC incidence increased significantly in the younger population and decreased significantly in the older population, with similar estimated annual percent changes for colon and rectum cancer (Table 1). These trends were geographically widespread: in all SEER counties, incidence rates increased in the younger population (Figure 1) and decreased in the older population (Figure 2). For both cancer types, there was generally strong evidence of heterogeneity in Net Drift, with standard deviations ranging from 0.73% (0.55%, 0.93%) to 0.28% (0.02%, 0.60%) within states, and from 1.03% (0.63%, 1.67%) to 0.34% (0.01%, 0.93%) between states (Table 1). There was more Net Drift heterogeneity in the older population than in the younger population, with greater differences for rectum cancer. Rectum cancer in the younger population had the most geographically homogeneous Net Drift – a finding also illustrated by the corresponding choropleth map (Figure 1).

Regardless of cancer type or age group, California, Washington, Utah, New Mexico, and Connecticut generally had lower than average, or average risk; Kentucky and Louisiana had

higher than average risk; levels of risk in the remaining SEER states (Hawaii, Iowa, Georgia, and New Jersey) depended on age group and cancer type (Figures 3 and 4). For both cancer types and in both age groups there was very strong evidence of RR heterogeneity both between and within states, with standard deviations ranging from 17.00% (10.73%, 27.86%) to 15.20% (8.74%, 23.94%) at the state level, and from 11.09% (8.31%, 13.97%) and 8.11% (7.22%, 9.10%) at the county level.

Choropleth maps reveal several localized geographic patterns of RR and Net Drift for both cancer types and in both age groups. For example, RR of rectum cancer in the younger population was below the SEER average in the Atlanta metropolitan area, and at-or-above the SEER average in the rest of Georgia (Figure 3). Additionally, in the younger population, colon cancer incidence appeared to be increasing faster in the northern Atlanta suburbs than in the southern suburbs (Figure 1). A cluster of counties in northwest New Mexico had colon cancer incidence increasing faster than average in the younger population, with McKinley county also exhibiting elevated RR (Figure 3). Localized geographic heterogeneity was not limited to the mainland US: RR and Net Drift on the island of Oahu in Hawaii exceeded that of the outer Hawaiian Islands and the overall SEER average; this was especially so for rectum cancer.

Our ecological correlation analysis suggests that rates of obesity and diabetes, followed by measures of socio-economic status (income per capita, poverty rate, and median rent) were the strongest correlates of RR and Net Drift, regardless of age group or cancer type (Table 2). County-level screening rates were negatively associated with Net Drift for colon cancer in the younger population and both RR and Net Drift for rectum cancer in the older population. County racial and ethnic composition generally did not correlate with model parameters, with two exceptions: % Hispanic in the younger population and % Asian for rectum cancer incidence in the older population. The food environment – measured by prevalence of US Department of Agriculture-designated food deserts – did not correlate significantly with model parameters. The one characteristic that suggested a potential correlation with diet quality was average Supplemental Nutritional Assistance Program (SNAP) benefit, especially for colon cancer in the younger population.

Discussion

We analyzed CRC incidence in 612 counties within 12 SEER states and found marked geographic heterogeneity between and within states, for both relative risk and estimated annual percent change (Net Drift). In line with previous work[1,5], we found increasing incidence in the younger, less screened population and decreasing incidence in the older, heavily screened population. Our work suggests that increasing incidence in the younger population is geographically widespread, at least across the US states covered by SEER. We identified geographic clusters of counties with elevated relative risk and higher-than-average secular trend, and showed that these clusters correlate with county-level prevalence of established CRC risk factors[15]. Thus, just as there are higher risk and lower risk individuals, our models allow us to identify higher risk and lower risk counties.

Specifically, we found that counties with high rates of diabetes and obesity and low socioeconomic status tend to have the greatest relative risk of CRC. In these same counties, the rates are also increasing more quickly over time than in all SEER areas combined. The negative correlation with socioeconomic status may reflect that the latter can serve as a proxy for a number of lifestyle risk factors, including higher smoking prevalence, lower levels of physical activity, and greater lack of access and lower utilization of appropriate screening[16]. Generally, we found that counties with high screening rates tend to have more favorable trends and lower relative risk, with statistically significant associations for colon cancer in the younger population and rectum cancer in the older population. Taken together, this suggests that adherence to screening may serve as an informative marker of long-term cancer control. The strong correlation between screening and rectum cancer – but not colon cancer – incidence in the same population is noteworthy because screening for both types of cancer typically occurs concurrently. It is possible that patients may be more aware of rectal polyps and seek treatment more frequently, or certain types of screening (e.g., fecal occult blood test) are more sensitive to rectal polyps.

Dietary intake has been implicated as an important factor in CRC development: pro-inflammatory diets are associated with higher risk[17–19], and "high quality" diets, which more closely follow established dietary guidelines, have been shown to be protective[20–22]. We assessed the county food environment using the prevalence of food deserts, because these can limit access to fresh fruits and vegetables, and average SNAP benefit amounts, because SNAP recipients are less likely to meet dietary guidelines[23].

Although we did not find a significant association between CRC and prevalence of food deserts, we did find higher relative risk and estimated annual percent changes in counties with higher average SNAP benefits, especially for colon cancer occurring in the younger population. It is difficult to disentangle the effects of poor diet, obesity, and poverty using an ecological study such as ours because these risk factors are all linked, and all correlate with CRC incidence. Indeed, there is evidence that income and access to fresh food must be considered jointly: whereas studies have shown that proximity to supermarkets is associated with lower obesity[24,25], this association disappears when accounting for food prices[26]. At the county level, SNAP benefits may capture the combination of low income and lower quality diet, due to an increased difficulty of purchasing fresh food. Additional studies are needed to formally assess this hypothesis.

Somewhat surprisingly, we found few associations between model parameters and county racial and ethnic composition. Negative associations between model parameters and % Hispanic and % Asian populations are likely a byproduct of the fact that California has lower-than-average risk and more favorable trends for both cancer types. Cases in California account for a large proportion of all SEER cases, and thus will greatly influence SEER-wide associations. American Indians were not explicitly captured in our correlation analysis, but several geographic clusters of elevated risk and unfavorable trends correspond closely to the locations of major Native American reservations. Examples include counties in Northwest New Mexico and Northeast Utah, which cover large portions of the Navajo, Pueblo, Zuni, and Apache Reservations (New Mexico), and the Uintah and Ouray Reservations (Utah).

Elevated risk on several islands of Hawaii (e.g., Island of Oahu) may reflect concentrations of Native Hawaiians, who also were not explicitly represented in the correlation analysis.

Our study has several limitations. Using registry data as we do, our study is ecological in nature and thus omits important information about individual risk factors, for example: family history, and clinical history of inflammatory bowel disease and/or colorectal polyps. Furthermore, SEER links incidence to counties based on patients' residential location at time of diagnosis. Especially in the younger population, residential mobility may bias the results of our analysis and post-estimation correlations with county risk factors. The risk factors themselves were ascertained at the county level using multiple data sources for different time periods, and may be subject to measurement error. Finally, several assumptions were made to make our modelling approach tenable: 1) age-period-cohort curvature parameters were effectively averaged across counties, regardless of state; 2) distributional assumptions were made about state and county-within-state random effects, as well as incident case counts themselves; and 3) temporal incidence patterns were assumed to be well-characterized using 5-year age and period intervals. Although these are simplifying assumptions, there was no indication that states had substantially different curvature parameters and there was no evidence that distributional assumptions were violated, due to satisfactory goodness-of-fit, as shown in the supplementary figures. Five-year intervals - although standard in epidemiological analyses - remain a limitation, and suggests that our results should be replicated at a finer temporal resolution (e.g., 1-year data), if computationally feasible.

Despite these limitations, our analysis demonstrates that high-risk-for-cancer counties can be identified using a spatial analysis of cancer registry data and correlational analysis using increasingly rich spatial datasets that describe the prevalence and distribution of established and potential cancer risk factors. These analyses have the potential to help target interventions in light of small-scale variations in need.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal Cancer Incidence Patterns in the United States, 1974–2013. J Natl Cancer Inst. 2017;109(8):1–6.

2. Austin H, Henley SJ, King J, Richardson LC, Eheman C. Changes in colorectal cancer incidence rates in young and older adults in the United States: what does it tell us about screening. Cancer causes & control. 2014;25(2):191–201. [PubMed: 24249437]

3. Bailey CE, Hu C-Y, You YN, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975–2010. JAMA surgery. 2015;150(1):17–22. [PubMed: 25372703]

4. Singh KE, Taylor TH, Pan C-JG, Stamos MJ, Zell JA. Colorectal cancer incidence among young adults in California. Journal of adolescent and young adult oncology. 2014;3(4):176–184. [PubMed: 25538862]

5. Brenner DR, Ruan Y, Shaw E, De P, Heitman SJ, Hilsden RJ. Increasing colorectal cancer incidence trends among younger adults in Canada. Preventive Medicine. 2017.

6. Molenaar RJ, Radivoyevitch T, Wilmink JW. RE: Colorectal Cancer Incidence Patterns in the United States, 1974–2013. JNCI: Journal of the National Cancer Institute. 2017;109(8):djx103–djx103.

7. Murphy CC, Sanoff HK, Stitzenberg KB, et al. RE: Colorectal Cancer Incidence Patterns in the United States, 1974–2013. JNCI: Journal of the National Cancer Institute. 2017;109(8):djx104. [PubMed: 29117391]

8. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov): Incidence - SEER 18 Regs Research Data, Nov 2015 sub, National Cancer Institute, DCCPS, Surveillance Research Program. 2017.

9. Chernyavskiy P, Little MP, Rosenberg PS. A unified approach for assessing heterogeneity in age-period-cohort model parameters using random effects. Stat Methods Med Res. 2017;10.1177/0962280217713033.

10. Kass RE, Raftery AE. Bayes factors. Journal of the american statistical association. 1995;90(430):773–795.

11. Buerkner P-C. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software. 2016.

12. R: A Language and Environment for Statistical Computing [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2015.

13. Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A. Visualization in Bayesian workflow. arXiv preprint arXiv:1709.01449. 2017.

14. Edwards LJ, Muller KE, Wolfinger RD, Qaqish BF, Schabenberger O. An R2 statistic for fixed effects in the linear mixed model. Statistics in medicine. 2008;27(29):6137–6157. [PubMed: 18816511]

15. Colorectal Cancer Risk Factors. 2017; https://www.cancer.org/cancer/colon-rectal-cancer/causes-risks-prevention/risk-factors.html. Accessed 10/16/2017, 2017.

16. Ward E, Jemal A, Cokkinides V, et al. Cancer disparities by race/ethnicity and socioeconomic status. CA: a cancer journal for clinicians. 2004;54(2):78–93. [PubMed: 15061598]

17. Tabung FK, Steck SE, Ma Y, et al. The association between dietary inflammatory index and risk of colorectal cancer among postmenopausal women: results from the Women's Health Initiative. Cancer causes & control. 2015;26(3):399–408. [PubMed: 25549833]

18. Harmon BE, Wirth MD, Boushey CJ, et al. The Dietary Inflammatory Index Is Associated with Colorectal Cancer Risk in the Multiethnic Cohort. The Journal of nutrition. 2017;147(3):430–438. [PubMed: 28179489]

19. Shivappa N, Prizment AE, Blair CK, Jacobs DR, Steck SE, Hébert JR. Dietary inflammatory index and risk of colorectal cancer in the Iowa Women's Health Study. Cancer Epidemiology and Prevention Biomarkers. 2014.

20. Vargas AJ, Neuhouser ML, George SM, et al. Diet Quality and Colorectal Cancer Risk in the Women's Health Initiative Observational Study. American Journal of Epidemiology. 2016;184(1):23–32. [PubMed: 27267948]

21. Park S-Y, Boushey CJ, Wilkens LR, Haiman CA, Le Marchand L. High-Quality Diets Associate With Reduced Risk of Colorectal Cancer: Analyses of Diet Quality Indexes in the Multiethnic Cohort. Gastroenterology. 2017;153(2):386–394.e382. [PubMed: 28428143]

22. Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: A systematic review and meta-analysis of observational studies. International journal of cancer. 2014;135(8):1884–1897. [PubMed: 24599882]

23. Andreyeva T, Tripp AS, Schwartz MB. Dietary quality of Americans by Supplemental Nutrition Assistance Program participation status: a systematic review. American journal of preventive medicine. 2015;49(4):594–604. [PubMed: 26238602]

24. Morland K, Roux AVD, Wing S. Supermarkets, other food stores, and obesity: the atherosclerosis risk in communities study. American journal of preventive medicine. 2006;30(4):333–339. [PubMed: 16530621]

25. Dubowitz T, Ghosh-Dastidar M, Eibner C, et al. The Women's Health Initiative: the food environment, neighborhood socioeconomic status, BMI, and blood pressure. Obesity. 2012;20(4):862–871. [PubMed: 21660076]

26. Ghosh-Dastidar B, Cohen D, Hunter G, et al. Distance to store, food prices, and obesity in urban food deserts. American journal of preventive medicine. 2014;47(5):587–595. [PubMed: 25217097]

Novelty and Impact:

Understanding the geographic heterogeneity of incidence facilitates resource allocation for potential interventions and can advance our knowledge of cancer etiology. For the first time, we quantify the spatial heterogeneity of colorectal cancer incidence in the United States among SEER states and counties. We show that increasing incidence in the younger, generally unscreened population (< age 55) is geographically widespread with clusters that correlate to well-established county-level risk factors (e.g., obesity).
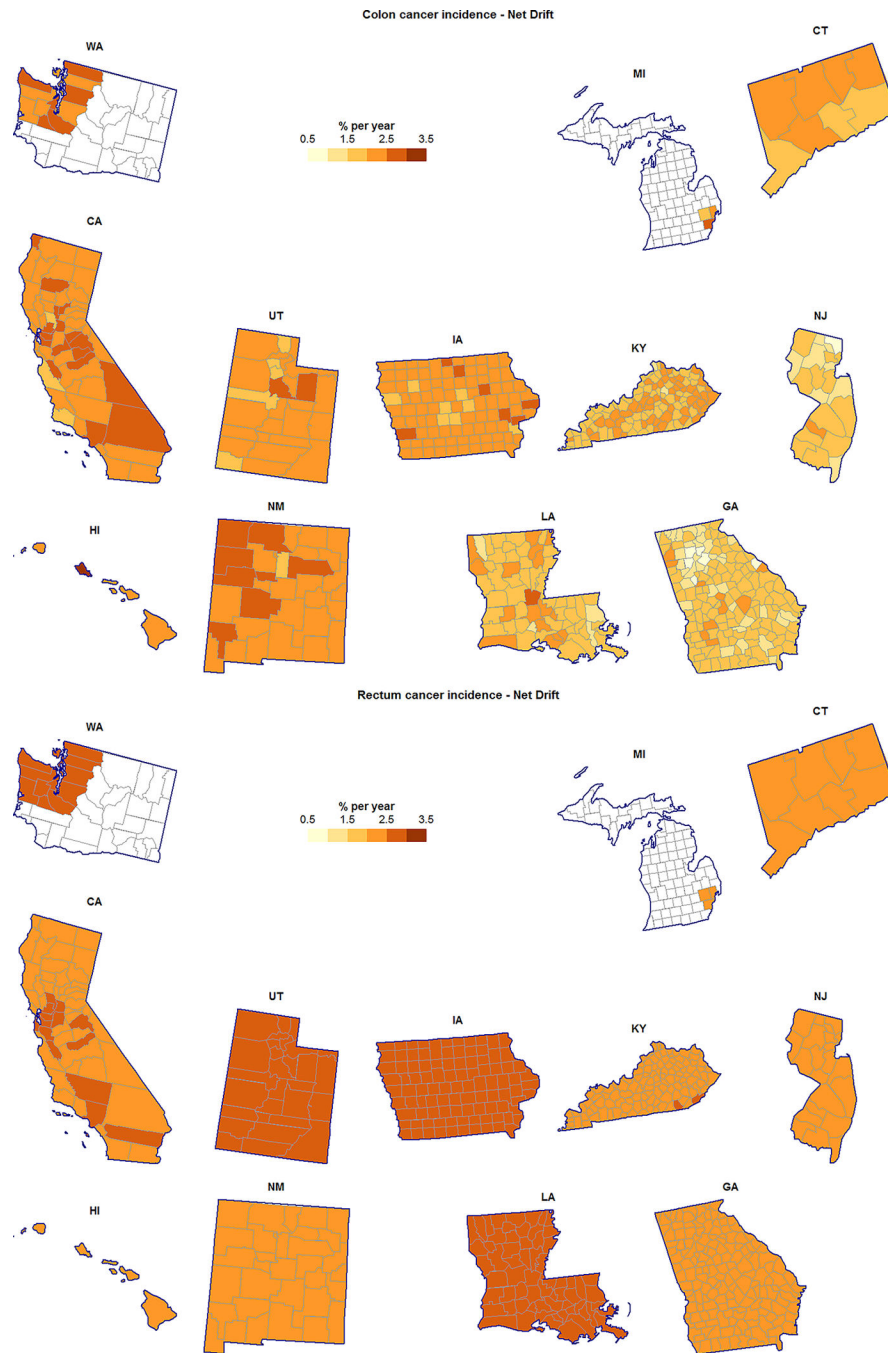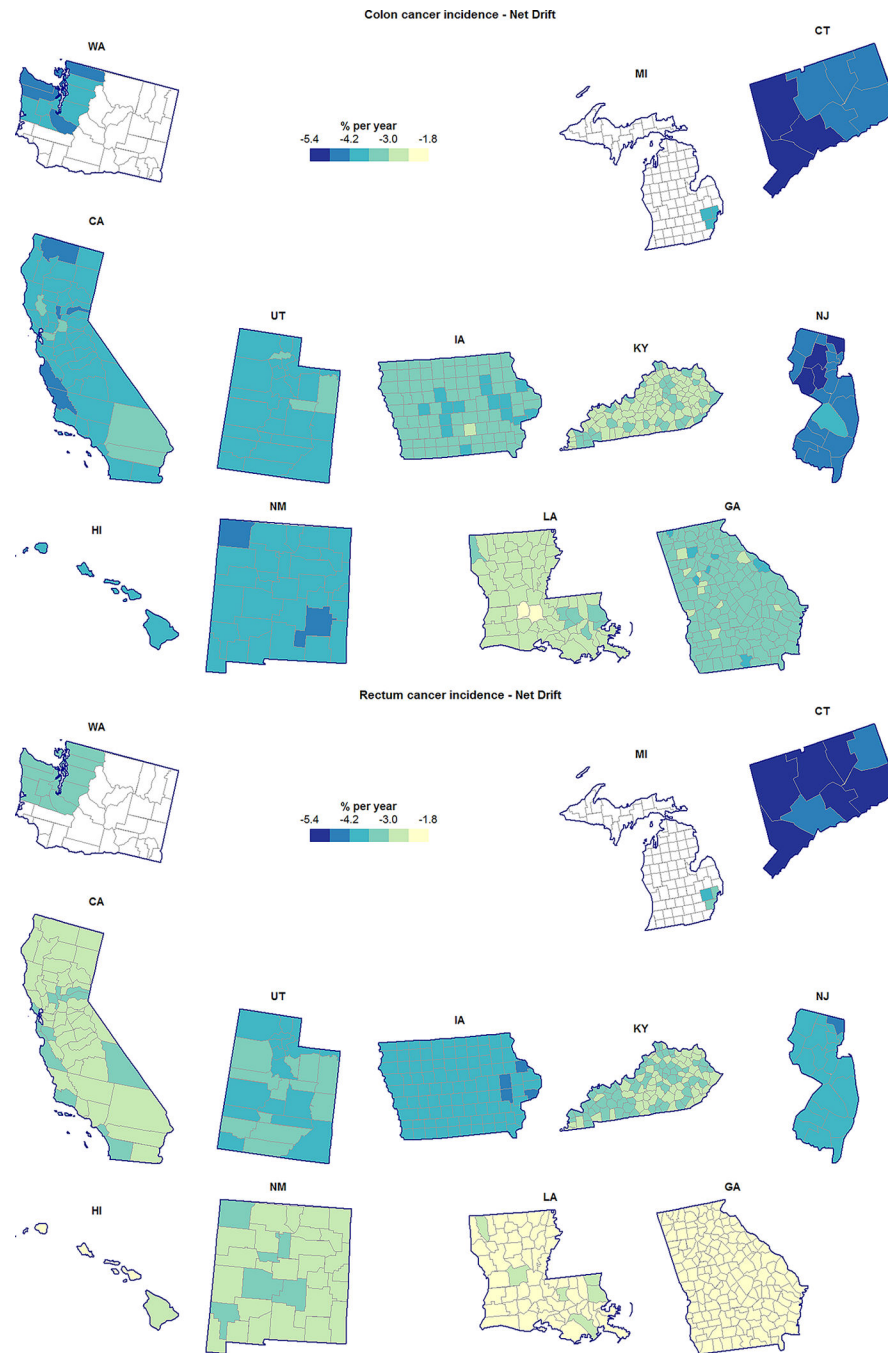
Figure 1.
County-specific age-adjusted annual percent change (Net Drift) for colon cancer incidence (*top*) and rectum cancer incidence (*bottom*), unscreened/lightly-screened men and women aged 20–54 years.
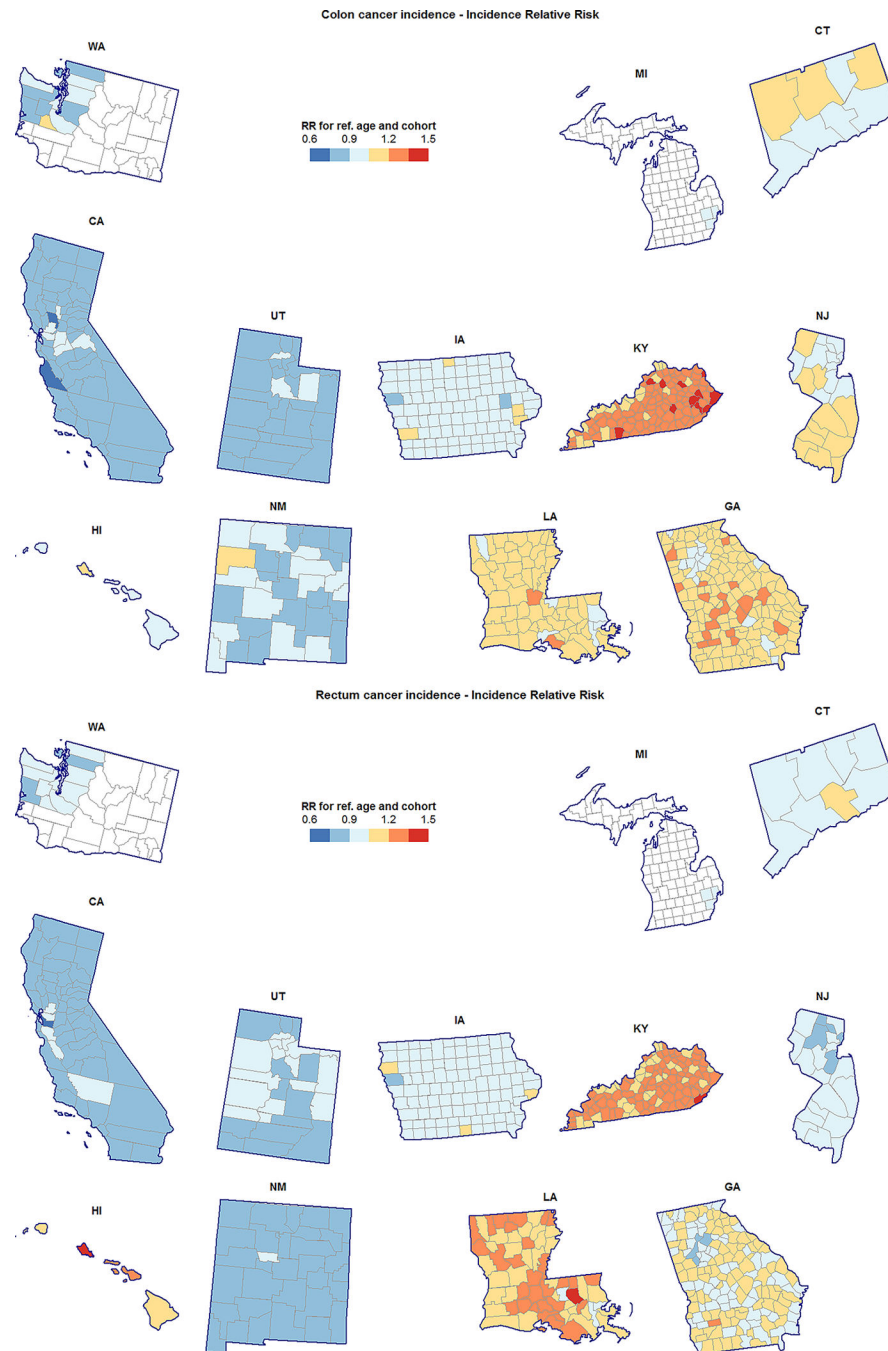
Figure 2.
County-specific age-adjusted annual percent change (Net Drift) for colon cancer incidence (*top*) and rectum cancer incidence (*bottom*), heavily-screened men and women aged 55–84 years.

Figure 3.
County-specific incidence relative risk (RR) for colon cancer (*top*) and rectum cancer (*bottom*), unscreened/lightly-screened men and women aged 20–54 years. RR shown are for the reference age (37) and birth cohort (1970) relative to the SEER average in this population.
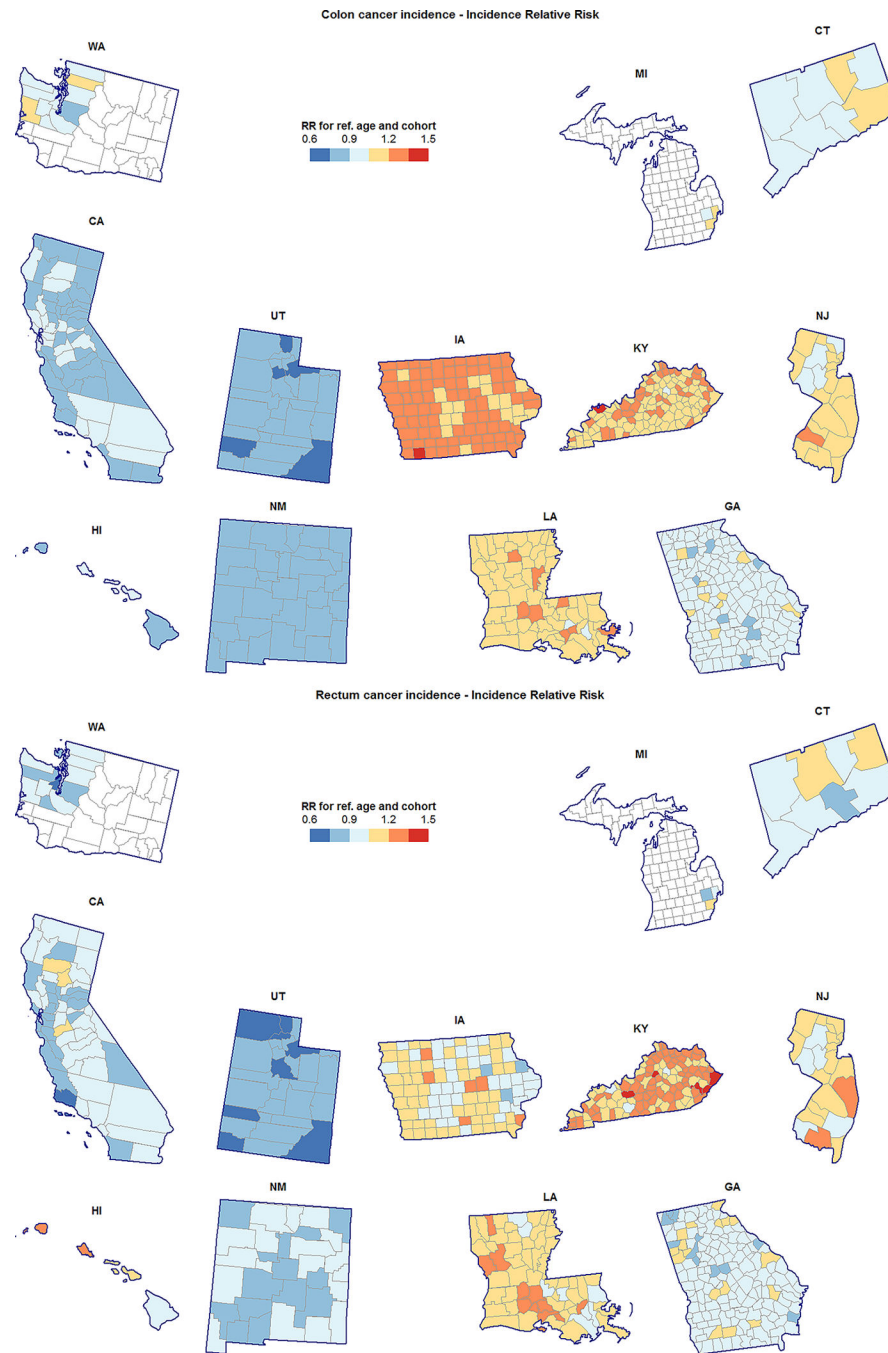
Figure 4.
County-specific incidence relative risk (RR) for colon cancer (*top*) and rectum cancer (*bottom*), heavily-screened men and women aged 55–84 years. RR shown are for the reference age (69.5) and birth cohort (1938) relative to the SEER average in this population.

Table 1.

Age-period-cohort model estimates for colon and rectum cancer incidence in men and women aged 20–54 (unscreened/lightly-screened) and 55–84 (heavily-screened), 2000–2014. Posterior medians are reported with 95% Credible Intervals in parentheses. Mean incidence rates are shown for the reference (mean) age group and birth cohort per 10^5 person-years. Evidence of heterogeneity is based on a Bayes Factor test of each standard deviation; we provide a verbal summary of the strength of evidence following Kass and Raftery (1995).

Model Parameter	Unscreened/lightly-screened	Heavily-screened
Colon cancer		
Mean incidence rate (per 10^5 person-years)	5.23 (4.74, 5.70)	120.39 (109.17, 134.09)
Between-state SD (%)	15.20 (8.74, 23.94)	17.00 (10.73, 27.86)
Evidence of heterogeneity	Very strong	Very strong
Within-state SD (%)	11.09 (8.31, 13.97)	8.11 (7.22, 9.10)
Evidence of heterogeneity	Very strong	Very strong
Age-adjusted trend (Net Drift) (% / yr)	2.11 (1.66, 2.63)	−3.58 (−4.15, −3.01)
Between-state SD (%)	0.44 (0.09, 0.91)	0.90 (0.54, 1.52)
Evidence of heterogeneity	Strong	Very strong
Within-state SD (%)	0.67 (0.30, 1.08)	0.73 (0.55, 0.93)
Evidence of heterogeneity	Very strong	Very strong
Longitudinal Age Trend (LAT) (% / yr)	15.12 (14.47, 15.76)	2.88 (2.41, 3.35)
Between-state SD (%)	0.70 (0.32, 1.23)	0.70 (0.39, 1.20)
Evidence of heterogeneity	Very strong	Very strong
Within-state SD (%)	0.67 (0.24, 1.11)	0.58 (0.33, 0.83)
Evidence of heterogeneity	Very strong	Very strong
Rectum cancer		
Mean incidence rate (per 10^5 person-years)	3.33 (3.06, 3.66)	46.12 (41.87, 50.86)
Between-state SD (%)	15.81 (9.45, 25.42)	15.64 (9.68, 25.91)
Evidence of heterogeneity	Very strong	Very strong
Within-state SD (%)	10.44 (7.67, 13.41)	10.72 (9.38, 12.15)
Evidence of heterogeneity	Very strong	Very strong
Age-adjusted trend (Net Drift) (% / yr)	2.44 (1.94, 2.98)	−3.29 (−3.96, −2.64)
Between-state SD (%)	0.34 (0.01, 0.93)	1.03 (0.63, 1.67)
Evidence of heterogeneity	Positive	Very strong
Within-state SD (%)	0.28 (0.02, 0.60)	0.50 (0.20, 0.85)
Evidence of heterogeneity	Positive	Strong
Longitudinal Age Trend (LAT) (% / yr)	16.07 (15.44, 16.70)	0.07 (−0.36, 0.44)
Between-state SD (%)	0.41 (0.03, 1.02)	0.53 (0.22, 0.96)
Evidence of heterogeneity	Positive	Very strong
Within-state SD (%)	0.23 (0.01, 0.65)	0.36 (0.03, 0.75)
Evidence of heterogeneity	Weak	Positive

Table 2.

County level association between estimated age-period-cohort model parameters (RR and Net Drift) and county characteristics. We report measures of overall linear association (R_β^2) for each characteristic that account for the hierarchical data structure and have a maximum of 1 and a minimum of 0.

County characteristics	Unscreened/lightly-screened		Heavily-screened	
	Relative Risk	Net Drift	Relative Risk	Net Drift
Colon cancer				
% screened	0.37 (−)	0.59 *(−)	0.43 (−)	0.29 (−)
% White	0.00 (−)	0.13 (−)	0.04 (+)	0.07 (−)
% Black	0.35 (+)	0.35 (+)	0.09 (+)	0.00 (+)
% Asian	0.35 (−)	0.36 (−)	0.35 (−)	0.35 (−)
% Hispanic	0.62 *(−)	0.56 (−)	0.46 (−)	0.01 (−)
Income per capita	0.65 **(−)	0.65 **(−)	0.47 *(−)	0.55 *(−)
Poverty rate	0.57 *(+)	0.75 *(+)	0.21 (+)	0.60 (+)
Avg SNAP benefit	0.80 *(+)	0.66 *(+)	0.51 (+)	0.74 *(+)
% Low income and food desert	0.38 (+)	0.39 (+)	0.17 (+)	0.14 (+)
% food desert	0.00 (+)	0.01 (−)	0.21 (−)	0.00 (−)
% diabetic	0.65 **(+)	0.68 **(+)	0.45 *(+)	0.66 **(+)
% obese	0.71 **(+)	0.57 **(+)	0.53 *(+)	0.65 **(+)
% urban	0.14 (−)	0.03 (−)	0.09 (−)	0.00 (+)
Median rent	0.68 **(−)	0.65 *(−)	0.34 (−)	0.44 *(−)
Population density	0.26 (−)	0.03 (−)	0.03 (−)	0.23 (−)
Rectum cancer				
% screened	0.36 (−)	0.43 (−)	0.82 **(−)	0.85 **(−)
% White	0.07 (+)	0.37 (−)	0.02 (−)	0.03 (−)
% Black	0.40 (+)	0.21 (+)	0.14 (+)	0.06 (+)
% Asian	0.25 (−)	0.26 (−)	0.39 (−)	0.43 *(−)
% Hispanic	0.61 **(−)	0.03 (−)	0.01 (−)	0.00 (−)
Income per capita	0.48 *(−)	0.16 (−)	0.87 **(−)	0.90 **(−)
Poverty rate	0.31 (+)	0.09 (+)	0.43 *(+)	0.54 *(+)
Avg SNAP benefit	0.29 (+)	0.15 (+)	0.59 *(+)	0.57 *(+)
% Low income and food desert	0.02 (+)	0.02 (+)	0.17 (+)	0.24 (+)
% food desert	0.02 (−)	0.22 (−)	0.11 (−)	0.05 (−)
% diabetic	0.46 (+)	0.38 (+)	0.71 **(+)	0.81 **(+)
% obese	0.63 *(+)	0.21 (+)	0.80 **(+)	0.84 **(+)
% urban	0.12 (−)	0.05 (−)	0.21 (−)	0.13 (−)
Median rent	0.51 *(−)	0.08 (−)	0.78 **(−)	0.84 **(−)
Population density	0.06 (−)	0.03 (+)	0.08 (−)	0.14 (−)

Symbols (+) and (−) indicate the direction of association

*
indicates the association is significant at ($\alpha = 0.05$)

**
indicates the association is significant at ($\alpha = 0.01$).