Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Ronak D. Ghiya Date

Approval Sheet

Sociodemographic Characteristics of Younger Adults Diagnosed with Colorectal Cancer (CRC)

By

Ronak D. Ghiya, MD

MPH

Executive Master of Public Health in Applied Epidemiology

_____

Michael Goodman, MD, MPH

Committee Chair

_____

Alex Edwards, MPH

Committee Member

Sociodemographic Characteristics of Younger Adults Diagnosed with Colorectal Cancer (CRC)

By

Ronak D. Ghiya, MD

Wright State University

2017

Thesis Committee Chair: Michael Goodman MD, MPH

An abstract of

A thesis submitted to the Faculty of the Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Executive Master of Public Health in Applied Epidemiology

2022

Abstract

Sociodemographic Characteristics of Younger Adults Diagnosed with Colorectal Cancer (CRC)

By: Ronak D. Ghiya, MD

Background: Colorectal cancer (CRC) is the third most common invasive malignant tumor diagnosed in males and second most common in females. Relatively little is known about the typical CRC presentation in pre-screening populations. Approximately 10% of new cases are in adults under the age of 50 years. CRC incidence among adults between 40 and 49 years increased by almost 15% from 2000-2002 to 2014-2016. The purpose of this thesis is to examine factors associated with stage of CRC diagnosis among persons between ages 18 and 44 years. Understanding sociodemographic characteristics of persons who present with advanced disease may offer important opportunities for reducing disparities.

Methods: Data from the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI) were utilized to obtain research data from 2009-2018. Patients were selected based on primary disease site and morphology codes. All CRC cases newly diagnosed among adult patients of prescreening age (18-44 years) and reported to one of the 21 SEER registries were categorized according to age, race/ethnicity, gender, registry/region, and year of diagnosis. These variables were included in a series of logistic regression models that used various stage-at-diagnosis characteristics as outcomes of interest.

Results: Among 28,828 age-eligible CRC patients reported to SEER during the study period 27,607 (95.6%) had known stage disease and among those 17,784 (64.4%) were diagnosed with advanced CRC. The crude and multivariable analyses of factors associated with advanced stage CRC among person with known stage disease were quite similar. Advanced CRC was more likely in persons under age 35 years and persons with undefined race/ethnicity. Similar results were found in the secondary analysis. Outside of the East Coast/Eastern region, patients were less likely to have unknown disease stage.

Conclusions: These analyses demonstrate that clinicians may need to pay closer attention to abdominal symptoms in younger patients and especially young females. Future research studies may consider capturing data from all 50 states and incorporate additional markers of social determinants of health as well as examine outcomes of CRC in younger age groups with a specific focus on survival and risk of recurrence.

Sociodemographic Characteristics of Younger Adults Diagnosed with Colorectal Cancer (CRC)

By

Ronak D. Ghiya, MD

Wright State University

2017

Thesis Committee Chair: Michael Goodman MD, MPH

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Executive Master of Public Health in Applied Epidemiology

2022

Acknowledgements

- Thank you to the Emory University Rollins School of Public Health faculty, staff, and student colleagues. It has been a pleasure learning and working together with all of you.

- Thank you to the Executive MPH Program and Thesis Committee for their time, feedback, and continued support. Special thanks to Dr. Michael Goodman for serving as my Thesis Committee Chair, Mr. Alex Edwards for serving as my Thesis Field Advisor, and Ms. Zelda Ray for serving as my Associate Director of Academic Programs.

- Thank you to the Emory University School of Medicine Preventive Medicine Residency program faculty and co-residents for their time, feedback, and continued support for as well.

- Acknowledgement to Surveillance, Epidemiology, and End Results (SEER) Data Program of the National Cancer Institute (NCI) for allowing us to utilize and analyze the colorectal cancer database.

Table of Contents

Overview

Colorectal cancer (CRC) is the third most common malignancy both in terms of incidence of cancer-specific mortality among men and women in the United States. (Islami, Ward et al. 2021) The lifetime risk for developing CRC is about 1 in 23 (4.3%) for men and 1 in 25 (4.0%) for women. (Siegel, Miller et al. 2022) Currently, the United States Preventive Task Force (USPSTF) guidelines recommends that men and women ages 45-75 years should be screened for CRC and that decisions about whether or not to screen for adults ages 76 to 85 years should be made in consultation with their health care providers. (Davidson, Barry et al. 2021) For these reasons, many newly diagnosed CRC patients are identified in the relatively early stages of the disease. By contrast, relatively little is known about the typical CRC presentation in the younger population groups that are not eligible for screenings. (Davidson, Barry et al. 2021) The purpose of this analysis is to examine incidence and prognostic characteristics of CRC in persons of ages 18-44 years with a primary focus on stage at diagnosis (early vs. advanced/late). The independent variables of interest include demographic and geographic factors such as race (White, Black, Asian/Pacific Islander, American Indian/Alaskan Native) and ethnicity (Hispanic and Non-Hispanic), gender (male and female), location (Eastern, Central/Midwest, Mountain, Western/Pacific Coast), and year of diagnosis (2009-2013 and 2014-2018). All analyses will be performed using logistic regression models.

Background

According to the World Health Organization, colorectal cancer (CRC) is the third most common invasive malignant tumor diagnosed in males and second most common cancer diagnosed in females. (Islami, Ward et al. 2021) In the United States, the lifetime incidence of CRC is approximately 4%. (Siegel, Miller et al. 2022) CRC incidence is approximately 25% higher in males compared to females. (Schoenfeld, Cash et al. 2005) In the United States, African Americans have the highest rates of CRC; approximately 20% higher than in Non-Hispanic Whites. (Jemal, Siegel et al. 2010)

Prior to May 2021, the United States Preventive Services Task Force (USPSTF) recommended CRC screening for men and women ages 50 to 75 years. In May 2021, the USPSTF revised its recommendations to indicate that CRC screening should begin at 45 years of age. This recommendation is for asymptomatic adults of average risk, with no prior diagnosis of CRC, no history of colon or rectal polyps, and no personal or family history of genetic disorders or mutations that would place them at higher risk for CRC. USPSTF recommends selective screening for adults ages 76-85 in consultation with their health care providers based on overall health, co-morbidities, prior screening history, and preferences. Among screened population groups, many newly diagnosed CRC patients are identified in the relatively early stages of the disease. By contrast, relatively little is known about the typical CRC presentation in the younger population groups that are not eligible for screening. (Davidson, Barry et al. 2021)

Approximately 10% of new colorectal cancer cases are in adults under the age of 50 years. (Islami, Ward et al. 2021) CRC incidence among adults between 40 and 49 years increased by almost 15% from 2000-2002 to 2014-2016. (Abualkhair, Zhou et al. 2020) The reasons for this increase are unclear and have not been systematically investigated. (Siegel, Miller et al. 2022)

In 2018, the American Cancer Society (ACS) updated its guidance to lower the threshold for screening adults who are at average risk for CRC from age 50 to 45. (Wolf, Fontham et al. 2018) In 2017, the United States Multi-Society Task Force on Colorectal Cancer recommended screening begin at age 45 in African Americans and age 50 in other racial groups. (Kwaan and Jones-Webb 2018) By contrast, the USPSTF did not make a separate and specific screening recommendation regarding African Americans in its 2021 update. (Davidson, Barry et al. 2021)

The purpose of this thesis is to examine factors associated with stage of CRC diagnosis among persons between ages 18 and 44 years. As stage of diagnosis is an important determinant of CRC survival, understanding sociodemographic characteristics of persons who present with advanced disease may offer important opportunities for reducing disparities. (Lansdorp-Vogelaar, Van Ballegooijen et al. 2009) This task can be accomplished by analyzing the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI). The SEER Program is an authoritative source of information on cancer incidence and survival in the United States. SEER currently collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 48% of the U.S.

population. SEER coverage includes 42% of the total population for Whites, 44.7% for African Americans, 66.3% for Hispanics, 59.9% for American Indians and Alaska Natives, 70.7% for Asians, and 70.3% for Hawaiian/Pacific Islanders. (Institute 2022)

The SEER registries routinely collect data on patient demographics, primary tumor site, tumor morphology and stage at diagnosis, first course of treatment, and follow-up for vital status. The SEER Program is the only comprehensive source of population-based information in the United States that includes stage of cancer at the time of diagnosis and patient survival data. The mortality data reported by SEER are provided by the National Center for Health Statistics, and the population data used in calculating cancer rates is obtained periodically from the Census Bureau. SEER data are updated annually and are available for public access. (Institute 2022)

Problem Statement

CRC is the third most common cancer in incidence and death among men and women in the United States. (Islami, Ward et al. 2021) There is currently insufficient evidence regarding typical presentation of CRC among adults younger than 45 years old. (Davidson, Barry et al. 2021)

Purpose Statement

Relatively little is known about CRC and prognosis among adults of pre-screening age which currently includes a range of 18-44 years. (Davidson, Barry et al. 2021) For this reason, the purpose of this analysis is to examine incidence and prognostic characteristics of CRC in this age group with the primary focus on stage of diagnosis (early vs. advanced/late). The independent variables of interest will include demographic and geographic factors such as race (White, Black, Asian/Pacific Islander, American Indian/Alaskan Native) and ethnicity (Hispanic and Non-Hispanic), gender (male and female), and location (by SEER registry).

Methods

The research data obtained from the Surveillance Epidemiology and End Results (SEER) Research Plus Program covered the period from 2009-2018. Patients were selected based on primary disease site and morphology codes according to the International Classification of Diseases for Oncology, third edition (ICD-O-3). The inclusion criteria were primary site of 1) carcinoma of the colon (ICD-O-3 codes 18.0-18.9), 2) carcinoma of the rectosigmoid junction (ICD-O-3 code C19.9), 3) carcinoma of the rectum (ICD-O-3 code C20.9), 4) patients in a research database ages 18-44 and 5) year of diagnosis was between 2009-2018 (inclusive). Patients were excluded from all analyses if their registry was listed as "Not Coded." Age was separated into three categories: 18-24 years, 25-34 years, 35-44 years. Race/ethnicity was separated by Non-Hispanic Whites, Non-Hispanic Blacks, American Indian/Alaskan Native, Asian/Pacific Islanders, Hispanics, and Other/Unknown. Gender was dichotomized as male and female. SEER 21 covers approximately 36.7% of the U.S. population, based on the 2010 Census. This is the largest geographic coverage for delay-adjusted statistics. Registry data included: 1) Eastern: Massachusetts, Connecticut, Detroit, New Jersey, New York, Kentucky, Georgia, 2) Central: Iowa and Louisiana, 3) Mountain: Idaho, Utah, New Mexico and 4) Western/Pacific: Alaska, California, Hawaii, Seattle (Puget Sound). The years of diagnosis were dichotomized by 2009-2013 and 2014-2018. We decided to use 2014 as a cutoff because this was the year of Medicaid expansion in some states.

TNM categories, stage groups, and definitions used by SEER are based on the *TNM Classification of Malignant Tumours, 7th Edition*. If the patient's diagnosis was classified as "In Situ" or "Localized" it was recoded as "Early Stage" and if diagnosis was classified as "Regional" or "Distant" it was recoded as "Late Stage." In addition, all newly diagnosed CRC cases were classified as "Known Stage" if relevant information was available or "Unknown Stage," if information was missing.

Data from 21 SEER Registries was used first to calculate the incidence rate of colorectal cancer for patients stratified according to age, race/ethnicity, gender, registry/region, and year of diagnosis. When performing logistic regression models, we analyzed each of the covariates age, race/ethnicity, gender, registry/region, and year of diagnosis both as univariable analysis and then in multivariable analysis will all the covariates included. The results of both univariable and multivariable regression analyses are expressed as crude and adjusted odds ratios (ORs) accompanied by 95% Cis, respectively. The reference categories were Age: 35-44 Years, Race: Non-Hispanic Whites, Gender: Male, Registry/Region: East Coast/Eastern, Year of Diagnosis: 2009-2013. All analyses were performed using the statistical software packages SAS (version 9.4; SAS Institute, Inc., Cary, NC) and SEER*Stat (version 8.3.9.2; National Cancer Institute, Calverton, MD).

Results

Among 28,828 age-eligible CRC patients reported to SEER during the study period 27,607 (95.6%) had known stage disease and the remaining 1221 (4.2%) had cancer of unknown stage. As shown in **Table 1**, the proportion of adults under the age of 25 was greater among CRC patients diagnosed with early-stage disease (8.6%) compared to advanced stage cancer patients (3.2%). Relative to the early-stage group, the advanced-stage CRC patients included somewhat higher proportions of men (52.7% vs. 47.7%), cases reported to the West Coast registries (39.7% vs. 36.7%), Hispanics (20.7% vs. 18.8%), and persons diagnosed prior to 2014 (46.3 vs. 43.7%). A corresponding comparison of CRC patients diagnosed with known stage disease to those who cancer stage was not known (**Table 2**) demonstrated that the former included greater proportions of Non-Hispanic Whites (55.7% vs. 47.3%), persons between ages 35 and 44 years (72.1 vs. 68.6%), and West coast residents (38.7 vs. 34.0%).

The crude and multivariable analyses of factors associated with advanced stage CRC among person with known stage disease are presented in **Table 3**. In general, the unadjusted and adjusted associations were quite similar. Using ages 35-44 years as the reference category, the adjusted ORs for age groups 18-24 years and 25-34 years were 2.92 (95% CI 2.61-3.26) and 1.24 (95% CI 1.17-1.32), respectively. In the analyses by race/ethnicity, using Non-Hispanic Whites as the reference group, the adjusted OR (95% CI) estimates were 0.89 (0.83-0.97) for Non-Hispanic Blacks and 0.90 (0.84-0.96) for Hispanics and 4.55 (3.45-6.00) for persons whose race ethnicity was classified as Other/Unknown. Other notable results included elevated OR for female vs.

male patients (OR=1.22; 95% CI 1.16-1.28) and for cases reported in the last 5 years of the study (2014-2018) relative to pre-2014 period (OR=1.08; 95% CI 1.03-1.14). Conversely, CRC cases reported to West Coast cases reported in the last 5 years of the study registries were less likely to be diagnosed with advanced disease (OR=0.88; 95% CI: 0.83-0.93) compared to their East Coast counterparts.

The corresponding analyses of factors associated with unknown stage CRC are presented in **Table 4.** Using ages 35-44 years as the reference category, the adjusted ORs for ages 18-24 years and 25-34 years were 1.33 (95% CI 1.05-1.68) and 1.13 (95% CI 0.99-1.30), respectively. Relative to Non-Hispanic white patients, elevated odds of unknown stage CRC were observed among Non-Hispanic Blacks (OR=1.31; 95% CI 1.11-1.55), Asian/Pacific Islanders (OR=1.28; 95% CI 1.03-1.57), Hispanics (OR=1.24; 95% CI 1.06-1.45), and person with Other/Unknown race/ethnicity (OR=8.97; 95% CI 6.91-11.65). A comparison by geographic area demonstrated that CRC patients diagnosed and treated in the Midwest/Central region, the Mountain region and the West Coast/Pacific region were all significantly less likely to have unknown disease stage compared to patients from the East Coast/Eastern region with OR (95% CI) estimates of 0.61 (95% CI 0.47-0.79), 0.49 (0.35-0.68), and 0.69 (95% CI 0.61-0.79), respectively.

Discussion

The present analysis, which focused on characteristics of CRC diagnosed among individuals not yet eligible for screening, produced several noteworthy findings. First, advanced CRC was more likely in persons under the age of 35 and especially those who received a CRC diagnosis in their late teens and early twenties. Second, the odds of advanced stage of diagnosis appear to differ relatively modestly by year and geographic region and across racial and ethnic groups, although persons with undefined race/ethnicity were significantly more likely to receive a late-stage CRC diagnosis. Finally, there was relatively small but appreciable gender disparity in stage of diagnosis with female CRC patients having higher odds of advanced disease compared to their male counterparts.

Younger individuals may be more likely to be diagnosed with advanced stage CRC because clinicians may not think about this type of cancer when encountering younger ages. It is also possible that CRC is more aggressive in younger age groups and the disease is more likely to rapidly advance from asymptomatic to symptomatic phase. Once patients become symptomatic, the diagnosis may be too delayed for successful treatment (Rodriguez, Brennan et al. 2018). Distinct histopathologic features resulting in increased aggressiveness are found in younger individuals less than age 50 years. These features indicate a higher proportion of locally advanced tumors invading adjacent structures (pT4). Furthermore, lymph node metastases were significantly more frequent in young patients. More frequent distant metastases were significantly diagnosed in patients younger than 40 years. Younger patients

more frequently suffered from advanced stages at time of diagnosis. Patients below age 50 years showed poor differentiation of tumor grading (G3) compared to those above age 50 years. Signet ring cell differentiation was significantly more frequent in patients below age 40 years. (Mueller, Schneider et al. 2021).

Individuals with undefined race/ethnicity may be more likely to present with late-stage CRC diagnosis because of factors pertaining to their social determinants of health. For example, this population may have a lack of health insurance, and inadequate access to medical care, and incomplete medical records with missing demographic information. Undefined race/ethnicity may serve as an indicator for poor medical care in general due to health inequity. (Zavala, Bracci et al. 2021)

Females have a higher odds of advanced stage CRC disease because they may be more likely to delay medical care in believing their symptoms do not warrant immediate medical attention. (Kim 2015) While men are more vocal and persistent with their concerns, females may not be as comfortable advocating for themselves and pushing back. Colon cancer symptoms are so wide and variable. Women may perceive symptoms such as abdominal bloating and discomfort perhaps as related to their menstrual cycle instead. Even when females present to a health care setting, their concerns are likely to be misdiagnosed or even dismissed by clinicians as a less critical issue. (Sebring 2021) For decades, the male body was the standard for medical research but there still has not been nearly enough progress made regarding health differences and outcomes by gender. While men and women may present with different symptoms, clinicians

are more accustomed to recognizing the male version and a “one size fits all approach.” (Regitz-Zagrosek 2012)

The secondary analysis revealed additional findings. For instance, individuals under the age of 35 years were more likely to have an unknown CRC stage. In addition, the likelihood of those with an unknown CRC stage was higher across racial and ethnic minority groups, and again especially in individuals with undefined race/ethnicity. Compared to patients residing in the East Coast of the United States individuals in most other regions had a lower odds of unknown CRC stage.

While younger individuals may be more likely to be diagnosed with CRC, the exact stage of their diagnosis may be unknown due to the medical complexity of their case. Individuals with undefined race/ethnicity are at higher risk of unknown stage due to social determinants of health as reflected in inadequate access to care and insufficient resources (Zavala, Bracci et al. 2021).

Our analysis contributed to the existing literature in several ways. CRC incidence is approximately 25% higher in males compared to females in all age groups (Schoenfeld, Cash et al. 2005); however, our analysis found that among young CRC patients female gender was associated with greater likelihood of advanced disease. In the United States, African Americans have the highest rates of CRC in all age groups (Jemal, Siegel et al. 2010), but our analysis showed that young Non-Hispanic Black patients were less likely present with advanced CRC

compared to their Non-Hispanic White counterparts. Similar discrepancy between the results for stage of diagnosis in the younger patients and the reported results for incidence rates across all ages was observed among Hispanics. (Garcia, Pruitt et al. 2018)

In the US, rates of early-onset CRC vary widely by state. The lowest incidence rates of CRC are found in the Mississippi Delta Region and Appalachia whereas Mississippi and Kentucky have the highest incidence rates of CRC. (Stoffel and Murphy 2020) Our analysis focused on odds of advanced stage and unknown stage of CRC, but we found that individuals living in the West Coast were less likely to be diagnosed with advanced CRC disease compared to their East Coast counterparts.

In general, younger patients are more likely to be treated aggressively with surgery, chemotherapy, and/or radiotherapy; however, most studies report no difference in CRC survival between younger and older patients. (Kneuertz, Chang et al. 2015) The distinguishing feature of our analysis is the focus on CRC characteristics among patients of pre-screening ages. This was accomplished by conducting a multivariable analysis that used a national cancer database compromised of 16 geographical areas.

The present analysis has several notable limitations. While the SEER program represents nearly one-half of the US population, it does capture data from all 50 states. Perhaps more importantly the SEER registries do not collection information on a number of important variables. The variables that would have strengthened our analysis include measures of

socioeconomic status, insurance status and type, rural/urban geography and measured of access to care. Future studies should also examine outcomes of CRC in the younger age groups with specific focus on survival and risk of recurrence.

Conclusions

These analyses demonstrate that clinicians may need to pay closer attention to abdominal symptoms in younger patients and especially young females. Future research studies may consider capturing data from all 50 states and incorporate additional markers of social determinants of health as well as examine outcomes of CRC in younger age groups with a specific focus on survival and risk of recurrence.

References

1. Abualkhair, W. H., et al. (2020). "Trends in Incidence of Early-Onset Colorectal Cancer in the United States Among Those Approaching Screening Age." JAMA Network Open **3**(1): e1920407.

2. Davidson, K. W., et al. (2021). "Screening for Colorectal Cancer." JAMA **325**(19): 1965.

3. Garcia, S., et al. (2018). "Colorectal cancer incidence among Hispanics and non-Hispanic Whites in the United States." Cancer Causes & Control **29**(11): 1039-1046.

4. Institute, N. C. (2022). "About the SEER Program." Retrieved January 24, 2022, from https://seer.cancer.gov/about/overview.html.

5. Islami, F., et al. (2021). "Annual Report to the Nation on the Status of Cancer, Part 1: National Cancer Statistics." JNCI: Journal of the National Cancer Institute **113**(12): 1648-1669.

6. Jemal, A., et al. (2010). "Cancer Statistics, 2010." CA: A Cancer Journal for Clinicians **60**(5): 277-300.

7. Kim, S.-E. (2015). "Sex- and gender-specific disparities in colorectal cancer risk." World Journal of Gastroenterology **21**(17): 5167.

8. Kneuertz, P. J., et al. (2015). "Overtreatment of Young Adults With Colon Cancer." JAMA Surgery **150**(5): 402.

9. Kwaan, M. R. and R. Jones-Webb (2018). "Colorectal Cancer Screening in Black Men: Recommendations for Best Practices." American Journal of Preventive Medicine **55**(5): S95-S102.

10. Lansdorp-Vogelaar, I., et al. (2009). "Individualizing colonoscopy screening by sex and race." Gastrointestinal Endoscopy **70**(1): 96-108.e124.

11. Mueller, M., et al. (2021). "Colorectal cancer of the young displays distinct features of aggressive tumor biology: A single-center cohort study." World Journal of Gastrointestinal Surgery **13**(2): 164-175.

12. Regitz-Zagrosek, V. (2012). "Sex and gender differences in health. Science & Society Series on Sex and Science." EMBO Rep **13**(7): 596-603.

13. The new concept of evidence-based sex and gender medicine—which includes the fundamental differences of biology and behaviour between women and men—should improve health care for both sexes.

14. Rodriguez, L., et al. (2018). "Disease Characteristics, Clinical Management, and Outcomes of Young Patients With Colon Cancer: A Population-based Study." Clinical Colorectal Cancer **17**(4): e651-e661.

15. Schoenfeld, P., et al. (2005). "Colonoscopic Screening of Average-Risk Women for Colorectal Neoplasia." New England Journal of Medicine **352**(20): 2061-2068.

16. Sebring, J. C. H. (2021). "Towards a sociological understanding of medical gaslighting in western health care." Sociol Health Illn **43**(9): 1951-1964.

17. Siegel, R. L., et al. (2022). "Cancer statistics, 2022." CA: A Cancer Journal for Clinicians **72**(1): 7-33.

18. Stoffel, E. M. and C. C. Murphy (2020). "Epidemiology and Mechanisms of the Increasing Incidence of Colon and Rectal Cancers in Young Adults." Gastroenterology **158**(2): 341-353.

19. Wolf, A. M. D., et al. (2018). "Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society." CA: A Cancer Journal for Clinicians **68**(4): 250-281.

20. Zavala, V. A., et al. (2021). "Cancer health disparities in racial/ethnic minorities in the United States." British Journal of Cancer **124**(2): 315-332.

Tables

Table 1: Descriptive characteristics of the study population by CRC stage

Participant characteristics	Early stage		Advanced stage	
	N	%	N	%
Age (years)				
18-24	849	8.6	576	3.2
25-34	2390	24.3	3878	21.8
35-44	6584	67.0	13,330	75.0
Race/ethnicity				
Non-Hispanic Whites	5595	57.0	9778	55.0
Non-Hispanic Blacks	1256	12.8	2423	13.6
American Indian/Alaska Native	98	1.0	146	0.8
Asian/Pacific Islanders	838	8.5	1691	9.5
Hispanics	1850	18.8	3674	20.7
Other/unknown	186	1.9	72	0.4
Gender				
Male	4681	47.7	9375	52.7
Female	5142	52.4	8409	47.3
Registry/region				
East Coast/Eastern	4884	49.7	8416	47.3
Midwest/Central	794	8.1	1317	7.4
Mountain	540	5.5	987	5.6
West Coast/Pacific	3605	36.7	7064	39.7
Year of diagnosis				
2009-2013	4290	43.7	8236	46.3
2014-2018	5533	56.3	9548	53.7
Total	9823	100.0	17784	100.0

Table 2: Descriptive characteristics of the study population by availability of stage information

Participant characteristics	Known stage		Unknown stage	
	N	%	N	%
Age (years)				
18-24	1425	5.2	83	6.8
25-34	6268	22.7	301	24.7
35-44	19914	72.1	837	68.6
Race/ethnicity				
Non-Hispanic Whites	15373	55.7	578	47.3
Non-Hispanic Blacks	3679	13.3	193	15.8
American Indian/Alaska Native	244	0.9	8	0.7
Asian/Pacific Islanders	2529	9.2	113	9.3
Hispanics	5524	20.0	243	19.9
Other/unknown	258	0.9	86	7.0
Gender				
Male	14056	50.9	636	52.1
Female	13551	49.1	585	47.9
Registry/region				
East Coast/Eastern	13300	48.2	704	57.7
Midwest/Central	2111	7.7	65	5.3
Mountain	1527	5.5	37	3.0
West Coast/Pacific	10669	38.7	415	34.0
Year of diagnosis				
2009-2013	12526	45.4	527	43.2
2014-2018	15081	54.6	694	56.8
Total	27607	100.0	1221	100.0

Table 3: Logistic regression analysis assessing factors associated with advanced CRC stage

Participant characteristics	Crude analysis		Adjusted analysis	
	OR	95% CI	OR	95% CI
Age (years)				
35-44	1.00	Reference	1.00	Reference
25-34	1.25	(1.18,1.32)	1.24	(1.17,1.32)
18-24	2.98	(2.67, 3.33)	2.92	(2.61,3.26)
Race/ethnicity				
Non-Hispanic Whites	1.00	Reference	1.00	Reference
Non-Hispanic Blacks	0.91	(0.84,0.98)	0.89	(0.83,0.97)
American Indian/Alaska Native	1.17	(0.91,1.52)	1.27	(0.97,1.64)
Asian/Pacific Islanders	0.87	(0.79,0.95)	0.92	(0.84,1.01)
Hispanics	0.88	(0.83,0.94)	0.90	(0.84,0.96)
Other/unknown	4.51	(3.43,5.93)	4.55	(3.45,6.00)
Gender				
Male	1.00	Reference	1.00	Reference
Female	1.23	(1.17,1.29)	1.22	(1.16,1.28)
Registry/region				
East Coast/Eastern	1.00	Reference	1.00	Reference
Midwest/Central	1.04	(0.95,1.14)	1.03	(0.94,1.14)
Mountain	0.94	(0.84,1.05)	0.90	(0.81,1.01)
West Coast/Pacific	0.88	(0.83,0.93)	0.88	(0.83,0.93)
Year of diagnosis				
2009-2013	1.00	Reference	1.00	Reference
2014-2018	1.11	(1.06,1.17)	1.08	(1.03,1.14)

Table 4: Logistic regression analysis assessing factors associated with unknown CRC stage

Participant characteristics	Crude analysis		Adjusted analysis	
	OR	95% CI	OR	95% CI
Age (years)				
35-44	1.00	Reference	1.00	Reference
25-34	1.14	(1.00,1.31)	1.13	(0.99, 1.30)
18-24	1.39	(1.10,1.75)	1.33	(1.05,1.68)
Race/ethnicity				
Non-Hispanic Whites	1.00	Reference	1.00	Reference
Non-Hispanic Blacks	1.40	(1.18,1.65)	1.31	(1.11,1.55)
American Indian/Alaska Native	0.87	(0.43,1.77)	1.09	(0.54,2.23)
Asian/Pacific Islanders	1.19	(0.97,1.46)	1.28	(1.03,1.57)
Hispanics	1.17	(1.00,1.36)	1.24	(1.06,1.45)
Other/unknown	8.87	(6.85,11.47)	8.97	(6.91,11.65)
Gender				
Male	1.00	Reference	1.00	Reference
Female	0.95	(0.85,1.07)	0.95	(0.84,1.06)
Registry/region				
East Coast/Eastern	1.00	Reference	1.00	Reference
Midwest/Central	0.58	(0.45,0.75)	0.61	(0.47,0.79)
Mountain	0.46	(0.33,0.64)	0.49	(0.35,0.68)
West Coast/Pacific	0.74	(0.65,0.83)	0.69	(0.61,0.79)
Year of diagnosis				
2009-2013	1.00	Reference	1.00	Reference
2014-2018	1.09	(0.97,1.23)	1.08	(0.96,1.21)