

# Risk Factors Associated With Early-Onset Colorectal Cancer

Valerie Gausman,\* David Dornblaser,\* Sanya Anand,\* Richard B. Hayes,†  
Kelli O’Connell,§ Mengmeng Du,§ and Peter S. Liang\*,||



\*Department of Medicine, NYU Langone Health, New York, New York; †Department of Population Health, NYU Langone Health, New York, New York; §Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York; and ||Department of Medicine, VA New York Harbor Health Care System, New York, New York

**BACKGROUND & AIMS:** The incidence of colorectal cancer (CRC) is increasing in individuals younger than 50 years, who do not usually undergo screening if they are of average risk. We sought to identify risk factors for CRC in this population.

**METHODS:** We compared sociodemographic and medical characteristics of patients who received a diagnosis of CRC at an age of 18–49 years (early-onset) with patients who received a diagnosis of CRC at an age of 50 years or older (late-onset) and with age-matched, cancer-free individuals (controls) at a tertiary academic hospital. We collected data from all adult patients with a diagnosis of CRC from January 1, 2011 through April 3, 2017 from electronic health records. Associations with risk factors were assessed using univariable and multivariable logistic regression models.

**RESULTS:** We identified 269 patients with early-onset CRC, 2802 with late-onset CRC, and 1122 controls. Compared with controls, patients with early-onset CRC were more likely to be male (odds ratio [OR], 1.87; 95% CI, 1.39–2.51), have inflammatory bowel disease (IBD) (3% vs 0.4% for controls; univariable  $P < .01$ ), and have a family history of CRC (OR, 8.61; CI, 4.83–15.75). Prevalence values of well-established modifiable CRC risk factors, including obesity, smoking, and diabetes, were similar. Compared to patients with late-onset CRC, patients with early-onset CRC were more likely to be male (OR, 1.44; 95% CI, 1.11–1.87), black (OR, 1.73; 95% CI, 1.08–2.65) or Asian (OR, 2.60; 95% CI, 1.57–4.15), and have IBD (OR, 2.97; 95% CI, 1.16–6.63) or a family history of CRC (OR, 2.87; 95% CI, 1.89–4.25). Sensitivity analyses excluding IBD and family history of CRC showed comparable results. Early-onset CRC was more likely than late-onset disease to be detected in the left colon or rectum (75% vs 59%,  $P = .02$ ) and at a late stage of tumor development (77% vs 62%,  $P = .01$ ).

**CONCLUSIONS:** In a retrospective study of patients with early-onset CRC vs late-onset CRC or no cancer, we identified non-modifiable risk factors, including sex, race, IBD, and family history of CRC, to be associated with early-onset CRC.

*Keywords:* Colon Cancer; Young Onset; Screening; BMI.

See editorial on page 2686.

Colorectal cancer (CRC) is the third leading cause of cancer for women and men in the United States.<sup>1</sup> CRC incidence has declined overall, and this has been attributed to population-level reductions in modifiable risk factors as well as increased participation in screening over the past 3 decades.<sup>2</sup> In contrast, CRC incidence among individuals younger than 50 years of age is on the rise, and at the current rate it is estimated to double by 2030.<sup>3</sup>

The reasons for this trend are unclear but likely include a combination of nonmodifiable and modifiable risk factors. Current guidelines recommend screening at an earlier age for individuals with a family history of CRC

and personal history of inflammatory bowel disease (IBD),<sup>4,5</sup> whereas specific recommendations for individuals with obesity or history of smoking do not exist. The American Cancer Society recently advocated lowering the age of screening to 45 years for the general population.<sup>6</sup> This approach has been shown to be cost-effective, but it would still add substantial cost to the healthcare system, and there are concerns that it may

*Abbreviations used in the paper:* BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; HANES, Health and Nutrition Examination Study; IBD, inflammatory bowel disease; NYC, New York City; OR, odds ratio.

Most current article

© 2020 by the AGA Institute  
1542-3565/\$36.00

<https://doi.org/10.1016/j.cgh.2019.10.009>

divert resources away from older individuals who have a higher absolute risk of cancer.<sup>7,8</sup> Alternatively, identifying specific risk factors in patients with early-onset CRC may permit risk stratification and targeted screening. We performed a single-institution study in a large, diverse metropolitan center to identify sociodemographic, medical, and histologic predictors of early-onset CRC.

## Materials and Methods

We conducted a retrospective chart review of patients who were seen at NYU Langone Health, a tertiary academic medical center in New York City. The protocol was approved by the New York University School of Medicine Institutional Review Board (Study #17-00077).

### Patient Selection

We identified all patients 18 years of age and older with a diagnosis of CRC who received care at our institution between January 1, 2011, and April 3, 2017, using the electronic health record. Patients were identified as having a history of CRC based on International Classification of Diseases–Ninth Revision (153.0–154.9) and International Classification of Diseases–Tenth Revision (C18–C20) codes. Based on recommendations from the U.S. Multi-Society Task Force on Colorectal Cancer and most other organizations to begin average-risk screening at 50 years of age, we defined early-onset CRC patients as those diagnosed between 18 and 49 years of age and created 2 comparison groups.<sup>4,9</sup> The first group comprised late-onset CRC patients diagnosed at 50 years of age or older. The second group comprised control subjects without cancer who were age-matched to early-onset cases in a 4:1 ratio. Control participants were randomly selected from patients who received care at our institution during the study period between 2011 and 2017, without matching for sex. Our initial automatic query identified 297 early-onset CRC cases, 2864 late-onset CRC cases, and 1188 age-matched control subjects. After manual review, we reclassified 11 late-onset cases as early onset and excluded 8 control subjects due to an incorrect birth year, leaving a total of 4341 patients (308 early-onset cases, 2853 late-onset cases, and 1180 age-matched control subjects). We excluded cases with no personal history of CRC (incorrect International Classification of Diseases coding), a personal history of a hereditary CRC syndrome, and substantial missing data. A hereditary CRC syndrome was defined as documentation of Lynch syndrome, familial adenomatous polyposis or *MUTYH*-associated polyposis, or pathogenic germline mutations in the mismatch repair or epithelial cellular adhesion molecule (*EPCAM*) genes. Control subjects with a prior history of cancer were also excluded. All patients received care at our institution, but some were diagnosed elsewhere.

## What You Need to Know

### Background

The incidence of early-onset colorectal cancer (CRC) (younger than 50 years of age) is increasing, but associated risk factors are not well understood.

### Findings

In a retrospective study of patients with early-onset CRC vs late-onset CRC or no cancer, we identified nonmodifiable risk factors, including male sex, black or Asian race, inflammatory bowel disease, and family history of CRC, to be associated with early-onset CRC. We did not identify an association with modifiable risk factors, including obesity, smoking, and diabetes.

### Implications for patient care

Risk-stratification efforts for early-onset CRC should include these nonmodifiable risk factors.

### Data Collection

We collected data on demographics, clinical history, and CRC outcomes through a combination of automated extraction and manual chart review. Clinical variables of interest included personal history of known CRC risk factors such as IBD, obesity, smoking, and diabetes, as well as a family history of CRC. These data were collected before diagnosis for most cases and within 6 months of diagnosis when prior assessment was unavailable. Area-level socioeconomic status including income and education were obtained using a crosswalk of residential zip codes and Primary Care Service Areas from the Dartmouth Atlas.<sup>10</sup> Tumor data (location, stage, grade, and molecular testing) were available for a minority of cases (~17%) from our tumor registry and were supplemented by manual chart review. Missing data was excluded from frequency calculations. When data from the automated extraction conflicted with results of the manual chart review, the latter was considered more accurate and used for analysis.

### Study Aims

Our primary aim was to identify sociodemographic and clinical risk factors for early-onset CRC by comparing early-onset cases with age-matched control subjects. Our secondary aim was to compare the early- and late-onset groups to identify differences in sociodemographic factors and clinical risk factors, as well as tumor characteristics. Age-related medical conditions such as hypertension, hyperlipidemia, coronary artery disease, stroke, and diabetes were only compared between early-onset cases and control subjects. We also performed sensitivity analyses of our primary and secondary outcomes excluding early-onset CRC patients

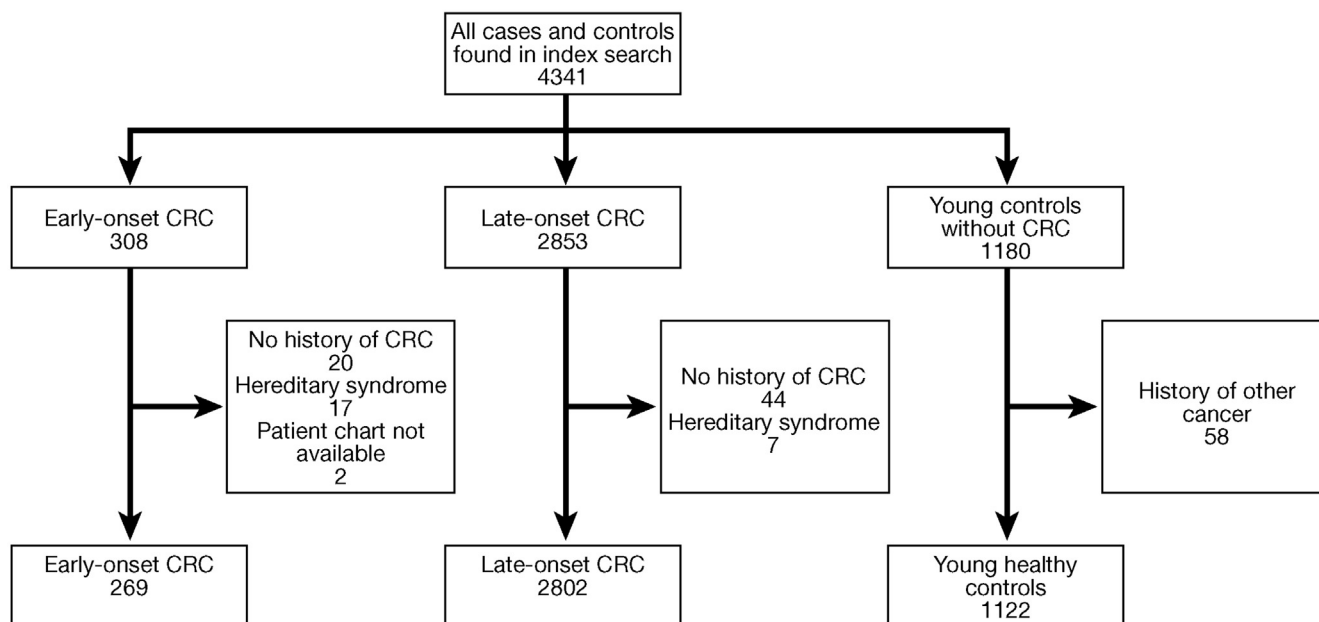


Figure 1. Study cohort flowchart.

with any family history of CRC or personal history of IBD. To assess the generalizability of our results to the broader New York City population, we also performed an analysis comparing the demographics of our hospital-based control subjects to a similarly-aged community cohort from the 2013 New York City Health and Nutrition Examination Study (NYC HANES), a survey and physical exam study representative of the entire adult noninstitutionalized NYC population.<sup>11</sup>

### Statistical Analysis

For the univariable analysis, we used the chi-square and Fisher exact tests for categorical variables and Student *t* test and Mann-Whitney *U* test for continuous variables. A 2-sided  $P < .05$  was considered statistically significant. Variables with  $P < .20$  were carried forward to the multivariable logistic regression model, in which we obtained adjusted odds ratios (ORs) and 95% confidence intervals (CIs). Some variables had either insufficient sample size or too much missing data to carry forward to multivariable analysis. Thus, they were left as univariable comparisons for the sake of hypothesis generation. Analysis was performed using R Statistical Software Version 1.0.143 (R Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics for NYC HANES were generated using sampling weights to account for the complex survey design.

### Results

A total of 4193 individuals met the inclusion and exclusion criteria (Figure 1), of whom 47% were men and 65% were non-Hispanic white. In addition, 0.9% had a personal history of IBD and 5% had a family history of CRC.

### Sociodemographic and Clinical Risk Factors of Early-Onset CRC

Compared with age-matched control subjects, patients with early-onset CRC were more likely to be men (OR, 1.87; 95% CI, 1.39–2.51) and have a family history of CRC (OR, 8.61; 95% CI, 4.83–15.75) or a personal history of IBD (Table 1). Of note, history of IBD was excluded from the multivariable logistic regression model because of the small sample size in the control group. Race as well as area-level mean household income and education level were not significantly different between these 2 groups. With respect to common age-related comorbidities, hyperlipidemia was more prevalent in the control group, and there was no difference in body mass index (BMI), smoking, coronary artery disease, hypertension, stroke, or diabetes between the 2 groups (Table 1).

Compared with patients with late-onset CRC, those with early-onset CRC were more likely to be men (OR, 1.44; 95% CI, 1.11–1.87), be black (OR, 1.73; 95% CI, 1.08–2.65), be Asian (OR, 2.60; 95% CI, 1.57–4.15), have a family history of CRC (OR, 2.87; 95% CI, 1.89–4.25), or have IBD (OR, 2.97; 95% CI, 1.16–6.63) (Table 2). Additional subanalyses on sex and race are shown in the Supplementary Materials. Mean household income and education level were similar between the 2 groups. The prevalence of common comorbidities such as obesity and diabetes were not compared because these comparisons in older vs younger CRC patients are intractably confounded by age.

Because patients with IBD or a family history of CRC are a high-risk population, we performed a sensitivity analysis that excluded early-onset CRC patients with any family history of CRC or personal

**Table 1.** Comparison of Individuals With Early-Onset CRC vs Young Control Subjects Without Cancer

Variable	Early-Onset group (n = 269)	Control group (n = 1122)	Univariable <i>P</i> <sup>a</sup>	Multivariable OR (95% CI)	Multivariable <i>P</i>
<b>Sociodemographic</b>					
Age, y	43 ± 6	45 ± 6			
Male	146 (54)	501 (45)	<.01	1.87 (1.39–2.51)	<.01
Race/ethnicity			<.26		
Non-Hispanic white	154 (57)	617 (55)			
Non-Hispanic black	26 (10)	108 (10)			
Non-Hispanic Asian	24 (9)	67 (6)			
Hispanic	10 (4)	45 (4)			
Other	55 (20)	285 (25)			
Income, US dollars	72,325 ± 27,127	72,703 ± 27,543	.84		
High school education, %	85 ± 9	85 ± 9	.77		
<b>Medical</b>					
BMI, kg/m <sup>2</sup>	27 ± 6	28 ± 6	.02	0.98 (0.95–1.00)	.06
Family history of CRC, n (%)	34 (13)	21 (2)	<.01	8.61 (4.83–15.75)	<.01
Inflammatory bowel disease, n (%)	7 (3)	5 (0.4)	<.01		
Crohn's disease, n (%)	4 (57)	4 (80)			
Ulcerative colitis, n (%)	3 (43)	1 (20)			
Smoker	70 (27)	312 (29)	.53		
Coronary artery disease	20 (4)	43 (4)	>.99		
Hypertension	50 (19)	221 (20)	.85		
Hyperlipidemia	41 (16)	259 (23)	.01	0.57 (0.38–0.83)	<.01
Stroke	1 (0.4)	8 (0.7)	>.99		
Diabetes	19 (7)	64 (6)	.48		

Values are mean ± SD or n (%).

BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; OR, odds ratio.

<sup>a</sup>Chi-square and Fisher exact tests were used for categorical variables and Student *t* test for continuous variables.

history of IBD and compared them with both control subjects and late-onset CRC cases ([Supplementary Tables 1 and 2](#)). The sensitivity analysis did not substantively alter our outcomes and therefore we have reported the main findings including these individuals.

### Comparison of Control Group With the 2013 NYC HANES Cohort

To assess whether our hospital-based control group is representative of the local population, we compared demographics and medical comorbidities with similarly aged participants (27–56 years of age) from the 2013 NYC HANES. Comparing frequencies in the control group with the weighted frequencies in NYC HANES, our cohort was older, contained a higher proportion of non-Hispanic whites and persons with coronary artery disease, and contained a lower proportion of smokers ([Table 3](#)). The sex distribution, BMI, and prevalence of other common comorbidities such as congestive heart failure, hypertension, hyperlipidemia, and diabetes were similar in the 2 groups. These data suggest that, for the most part, our hospital-based control group is representative of the NYC population with respect to most key conditions and risk factors linked to CRC.

### Tumor Characteristics of Early-Onset vs Late-Onset CRC

We also compared tumor characteristics of early-onset vs late-onset CRC in the subset of cases with available data ([Table 4](#)). Due to the limited available data, these analyses are exploratory and aimed at hypothesis generation. Early-onset CRC, compared with late-onset CRC, was associated with distal tumors (left colon or rectum, 75% vs 59%; *P* = .02) and late-stage disease at diagnosis (stage III/IV, 77% vs 62%; *P* = .01). Tumor grade was similar in both groups, with approximately 80% classified as low grade or well differentiated. Early-onset CRC had a lower prevalence of microsatellite instability than did late-onset CRC (6% vs 18%; *P* = .03), but the prevalence of *KRAS* mutations was similar (48% vs 41%; *P* = .52) in the 2 groups.

### Age of Diagnosis of Early-Onset CRC Patients and Family History of CRC

Among patients with early-onset CRC, 12 (4%) individuals were diagnosed at 20–29 years of age, 60 (22%) at 30–39 years of age, and 197 (73%) at 40–49 years of age. Thirty-four of the 269 (13%) patients in this group had a documented family history of CRC, of whom

**Table 2.** Comparison of Individuals With Early-Onset CRC vs Late-Onset CRC

Variable	Early-Onset group (n = 269)	Late-Onset group (n = 2802)	Univariable <i>P</i> <sup>a</sup>	Multivariable OR (95% CI)	Multivariable <i>P</i>
<b>Sociodemographic</b>					
Age at CRC diagnosis, y	43 ± 6	71 ± 11	.04	1.44 (1.11–1.87)	<.01
Male	146 (54)	1335 (48)	<.01		
<b>Race/ethnicity</b>					
Non-Hispanic white	154 (57)	1970 (70)			
Non-Hispanic black	26 (10)	219 (8)		1.73 (1.08–2.65)	.02
Non-Hispanic Asian	24 (9)	126 (5)		2.60 (1.57–4.15)	<.01
Hispanic	10 (4)	88 (3)		1.64 (0.78–3.09)	.16
Other	55 (20)	399 (14)		1.89 (1.34–2.65)	<.01
Income, US dollars	72,325 ± 27,127	72,307 ± 26,135	.99		
High school education, %	85 ± 9	85 ± 9	.38		
<b>Medical</b>					
BMI, kg/m <sup>2</sup>	27 ± 6	28 ± 6	.02	0.98 (0.95–0.99)	.04
Family history of CRC	(13)	(5)	<.01	2.87 (1.89–4.25)	<.01
Inflammatory bowel disease	(3)	(1)	.03	2.97 (1.16–6.63)	.01
Crohn's disease	(57)	(78)			
Ulcerative colitis	(43)	(22)			
Duration of IBD before CRC diagnosis, y	22 (17–25)	13 (0–36)	.71		

Values are mean ± SD, n (%), or median (interquartile range).

BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; IBD, inflammatory bowel disease; OR, odds ratio.

<sup>a</sup>Chi-square and Fisher exact tests were used for categorical variables and Student's *t* test and Mann-Whitney *U* test for continuous variables.

17 had a first-degree relative with CRC and therefore should have received screening before 50 years of age. A slight majority (n = 19 of 34, 56%) of the early-onset CRC patients with a family history of CRC were diagnosed before current U.S. Multi-Society Task Force on Colorectal Cancer guidelines would have recommended screening.<sup>4</sup> They were diagnosed a median 6 (interquartile range, 1–29) years before the recommended screening age. The remaining 15 (44%) patients, all of whom had indications for screening before 50 years of age, were diagnosed a median 6 (interquartile range, 0–19) years after the recommended screening age.

The vast majority of early-onset CRC patients (n = 252 of 269, 94%) were considered average-risk for screening. Of these 252 individuals, 101 (40%) were diagnosed between 46 and 49 years of age, 15 (6%) at 45 years of age, and the other 136 (54%) between 21 and 44 years of age.

## Discussion

In this large retrospective study of early-onset CRC, we identified male sex, family history of CRC, and personal history of IBD as predictors of early-onset CRC compared with both age-matched control subjects and late-onset CRC cases. Additionally, early-onset CRC patients were more likely to be black or Asian compared with late-onset CRC patients. Common age-related comorbidities were not more prevalent in early-onset cases than age-matched control subjects, and

socioeconomic factors were not significant risk factors. These data show certain nonmodifiable risk factors contribute to early-onset CRC.

After excluding all known cases of hereditary cancer syndromes, we found early-onset CRC patients were more than 8 times as likely to have a family history of CRC compared with control subjects and nearly 3 times more likely than late-onset CRC patients. CRC diagnosed before 50 years of age has been more strongly associated with family history of CRC or probable hereditary syndrome than has CRC diagnosed later in life. Chen et al<sup>12</sup> found an 8% higher absolute prevalence of family history in young-onset CRC cases compared with cases diagnosed at 50 years of age or older (25% vs 17%; *P* = .03). Although our study had a lower overall prevalence of positive family history due to the exclusion of hereditary syndromes, we also found an 8% absolute difference in prevalence of family history in the early-onset and late-onset CRC groups (13% vs 5%; *P* < .01). The prevalence of hereditary syndromes is significantly higher in early-onset CRC cases than in healthy control subjects, but to our knowledge, the association with family history of CRC in the absence of hereditary syndromes has not previously been reported. This could be attributed to weaker genetic risk factors for early-onset CRC including intermediate penetrant genes, low-risk genetic variations with additive effect, and genetic variants that modify the expression of known CRC susceptibility genes.<sup>13</sup> Our results also confirm that IBD is a risk factor for early-onset CRC.<sup>14,15</sup> The prevalence of IBD in early-onset CRC patients compared with control subjects has not

**Table 3.** Comparison of Control Subjects vs NYC HANES 2013

Variable	Control subjects (n = 908)	NYC HANES (n = 1122)	NYC HANES (Weighted %) (95% CI)
<b>Sociodemographic</b>			
Age at search/ survey (by 10 y)			
27–36 y	131 (12)	397	38.9 (35.3–42.7)
37–46 y	435 (39)	247	30.5 (26.9–34.4)
47–56 y	556 (50)	264	30.5 (26.9–34.5)
Age at search/ survey (by 5 y)			
27–31 y	41 (4)	216	21.2 (18.5–24.1)
32–36 y	90 (8)	181	17.7 (14.9–21.1)
37–41 y	117 (10)	114	13.6 (11.1–16.6)
42–46 y	318 (28)	133	16.9 (14.1–20.0)
47–51 y	376 (34)	126	14.7 (12.1–17.7)
52–56 y	180 (16)	138	15.9 (12.9–19.5)
Male	501 (45)	392	47.3 (44.0–50.5)
Race/ethnicity			
Non-Hispanic white	617 (55)	294	33.2 (27.7–39.1)
Non-Hispanic black	108 (10)	208	22.4 (17.3–28.6)
Non-Hispanic Asian	67 (6)	134	16.2 (12.1–21.3)
Hispanic	45 (4)	218	25.2 (20.8–30.3)
Other	285 (25)	54	3.0 (2.2–4.2)
<b>Medical</b>			
BMI, kg/m <sup>2</sup>	28 ± 6	28 ± 0.25	27.5–28.5 <sup>a</sup>
Smoker	312 (29)	367	42.1 (38.1–46.1)
Coronary artery disease	43 (4)	8	0.95 (0.47–1.9)
Congestive heart failure	6 (0.5)	9	1.0 (0.5–2.0)
Hypertension	221 (20)	195	22.0 (18.8–25.5)
Hyperlipidemia	259 (23)	218	25.2 (22.0–28.6)
Diabetes	64 (6)	64	7.3 (5.6–9.4)

NOTE. Values are n (%) or mean ± SD, unless otherwise indicated. BMI, body mass index; CI, confidence interval; NYC HANES, New York City Health and Nutrition Examination Study.  
<sup>a</sup>95% CI

been previously examined, but a recent Swedish study found that patients diagnosed with IBD as children were more likely to be diagnosed with early-onset CRC than individuals without IBD.<sup>16</sup>

The finding that Asians and blacks have higher rates of early-onset CRC than do non-Hispanic whites is supported by studies using the Surveillance, Epidemiology, and End Results registry.<sup>17,18</sup> A 1973–2009 Surveillance, Epidemiology, and End Results registry analysis found that the proportion of CRC patients who were diagnosed under 50 years of age was 1.8-fold higher in Asians/Pacific Islanders than in non-Hispanic whites; however, overall CRC incidence in Asians was lower.<sup>17</sup> It has been hypothesized that the rising incidence of CRC in Asians may be attributed to the adoption of the Western diet.<sup>19</sup> Asians who are born in the United States and exposed to

a lifetime Western diet may carry a higher risk for early-onset and overall CRC. On the other hand, the majority of Asians in the United States are immigrants who may have less exposure to the Western diet. Thus, lifestyle differences between Asians born in the United States and those who immigrated may explain the discrepancy between overall and early-onset CRC risk. With respect to black patients, a study of rectal and rectosigmoid cancers in individuals under 40 years of age found a higher absolute incidence in blacks compared with whites (0.67 per 100,000 [95% CI, 0.60–0.74] vs 0.51 per 100,000 [95% CI, 0.48–0.53]).<sup>18</sup>

In contrast to a large prospective cohort study that found higher risk for early-onset CRC in obese women,<sup>20</sup> we did not observe an association between obesity and early-onset CRC. A large retrospective study comparing young CRC cases and control subjects found specific dietary components, but not obesity or diabetes, to be risk factors for early-onset CRC.<sup>21</sup> Therefore, the relationship between metabolic conditions such as obesity or diabetes and early-onset CRC may be confounded by unmeasured dietary or environmental factors. Additional prospective studies with dietary and environmental exposure data are needed to further investigate this relationship.

Socioeconomic disparities in CRC incidence are well documented and have been attributed to a higher burden of predisposing comorbidities and lower rates of screening in groups with lower socioeconomic status.<sup>22</sup> That socioeconomic status was not a risk factor for early-onset CRC in our study may be explained by our patient demographics. The relative affluence of our New York City population (area-level mean household income >\$72,000) and abundant local resources for CRC screening may have reduced our ability to detect disparities related to healthcare access. Moreover, it is not clear that socioeconomic disparities observed for CRC overall are applicable for early-onset disease, as biological mechanisms may be distinct and access to screening may be less relevant for early-onset cases.

The finding that early-onset CRC presents at a more advanced stage than late-onset disease is consistent with the literature.<sup>12,23,24</sup> Ample evidence suggests the existence of biological differences in early-onset vs late-onset CRC. First, the rising incidence of early-onset CRC is predominantly due to left-sided disease, in contrast to a higher proportion of right-sided tumors in late-onset CRC.<sup>24–27</sup> Second, early-onset CRC exhibits a higher prevalence of mucinous or signet-ring histology with poor differentiation than that of older adults.<sup>27</sup> Last, as our study demonstrated, late-onset CRC also has a higher prevalence of microsatellite instability, which is associated with early-stage<sup>23</sup> and right-sided disease.<sup>28</sup> Prior studies have shown that approximately 12%–17% of all CRC is positive for microsatellite instability, and the majority of these are sporadic.<sup>29</sup>

Forty percent of the average-risk early-onset cases in our population were diagnosed between 46 and 49 years

**Table 4.** Comparison of Tumor Characteristics of Early-Onset CRC vs Late-Onset CRC

Variable	Early onset (known)	Early onset	Late onset (known)	Late onset	Univariable $P^a$
Site	79		463		.02
Right colon		20 (25)		190 (41)	
Left colon		32 (41)		162 (35)	
Rectum		27 (34)		111 (24)	
Stage	83		460		.01
Early stage		19 (23)		173 (38)	
Late stage		64 (77)		287 (62)	
Grade	73		412		.53
Low grade		58 (80)		343 (83)	
High grade		15 (21)		69 (17)	
KRAS mutation	98	47 (48)	49	20 (41)	.52
MSI	117	7 (6)	56	10 (18)	.03

NOTE. Values are n or n (%).

CRC, colorectal cancer; MSI, microsatellite instability

<sup>a</sup>Chi-square and Fisher exact tests were used for analysis.

of age, which may support the American Cancer Society recommendation to start screening at 45 years of age. However, it is unclear what proportion of these cases would have been prevented or detected at an earlier stage. Additionally, as a recent Markov model analysis demonstrated, targeted screening in high-risk individuals may be a more cost-effective approach.<sup>8</sup> Consequently, simple risk scores that can identify individuals at highest risk for early-onset CRC are needed. Our data highlight that efforts to construct these scores should include nonmodifiable risk factors.

The strengths of our study include its large sample size, the availability of sociodemographic, medical, and tumor data, as well as a hospital-based control group that is representative of the local population. However, several limitations should be noted. First, certain established risk factors—such as dietary history, physical activity, and aspirin use—were not included in the analysis because they were either unavailable or could not be easily extracted from the medical record. Second, because data was collected retrospectively, it is possible that certain medical data (eg, BMI) for cases may have been influenced by cancer itself. We tried to minimize this bias by including medical data obtained before or within 6 months of CRC diagnosis for cases diagnosed at our institution. Third, although we conducted a manual chart review to identify and exclude patients with hereditary syndromes, it is possible that some genetic testing was missing and a small number of patients with hereditary syndromes were included in the analysis. Fourth, because not all patients underwent surgery and some received resection outside of our institution, only a minority of patients had tumor data available. Therefore, the tumor analysis should be considered exploratory and the results interpreted with caution. Finally, there may be selection bias in that the hospital-based control cohort may not reflect a truly healthy population. However, our comparison to NYC HANES showed that cancer-free

control subjects were comparable to the NYC population with respect to important predictors of CRC.

In summary, the majority of early-onset CRC cases in our single-center study were sporadic. Patients with early-onset CRC were more likely to be men and have a family history of CRC or personal history of IBD. In addition, they were more likely to be black or Asian compared with individuals with late-onset CRC. We did not observe associations with well-established modifiable risk factors such as obesity, smoking, and diabetes. These data suggest that nonmodifiable factors should be included in risk prediction models to facilitate targeted screening in individuals under 50 years of age. Such efforts can be refined as more granular predictors of early-onset CRC, especially early-life exposures that are measurable and readily available in the clinical setting, are identified from future prospective studies.

## Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at [www.cghjournal.org](http://www.cghjournal.org), and at <https://doi.org/10.1016/j.cgh.2019.10.009>.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics 2018. *CA Cancer J Clin* 2018;68:7–30.
2. Siegel RL, Ward EM, Jemal A. Trends in colorectal cancer incidence rates in the United States by tumor location and stage, 1992–2008. *Cancer Epidemiol Biomarkers Prev* 2012; 21:411–416.
3. Bailey CE, H CY, You YN, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975–2010. *JAMA Surg* 2015;150:17–22.
4. Rex DK, Bolan CR, D'Onofrio JA, et al. Colorectal Cancer Screening: recommendations for physicians and patients from the U.S. Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol* 2017;112:1016–1030.

5. American Society for Gastrointestinal Endoscopy Standards of Practice Committee, Shergill AK, Lightdale JR, et al. The role of endoscopy in inflammatory bowel disease. *Gastrointest Endosc* 2015;81:1101–1121.e1–13.
6. Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin* 2018;68:250–281.
7. Liang PS, Allison J, Ladabaum U, et al. Potential intended and unintended consequences of recommending initiation of colorectal cancer screening at age 45 years. *Gastroenterology* 2018;155:950–954.
8. Ladabaum U, Mannalithara A, Meester RGS, Gupta S, Schoen RE. Cost-effectiveness and national effects of initiating colorectal cancer screening for average-risk persons at age 45 years instead of 50 years. *Gastroenterology* 2019;157:137–148.
9. U.S. Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, et al. Screening for colorectal cancer: U.S. Preventive Services Task Force Recommendation Statement. *JAMA* 2016;315:2564–2575.
10. Dartmouth Atlas Project. The Dartmouth Atlas of Healthcare; 2018. Available at: <https://www.dartmouthatlas.org/>. Accessed December 2, 2017.
11. Thorpe LE, Greene C, Freeman A, et al. Rationale, design and respondent characteristics of the 2013–2014 New York City Health and Nutrition Examination Survey (NYC HANES 2013–2014). *Prev Med Rep* 2015;2:580–585.
12. Chen FW, Sundaram V, Chew TA, Ladabaum U. Advanced-stage colorectal cancer in persons younger than 50 years not associated with longer duration of symptoms or time to diagnosis. *Clin Gastroenterol Hepatol* 2017;15:728–737.e3.
13. Jaspersen KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology* 2010;138:2044–2058.
14. Eaden JA, Abrams KR, Mayberry JF. The risk of colorectal cancer in ulcerative colitis: a meta-analysis. *Gut* 2001;48:526–535.
15. Bernstein CN, Blanchard JF, Kliwer E, Wajda A. Cancer risk in patients with inflammatory bowel disease: a population-based study. *Cancer* 2001;91:854–862.
16. Olen O, Askling J, Sachs MC, et al. Childhood onset inflammatory bowel disease and risk of cancer: a Swedish nationwide cohort study 1964–2014. *BMJ* 2017;358:j3951.
17. Rahman R, Schmaltz C, Jackson CS, Simoes EJ, Jackson-Thompson J, Ibdah JA. Increased risk for colorectal cancer under age 50 in racial and ethnic minorities living in the United States. *Cancer Med* 2015;4:1863–1870.
18. Meyer JE, Narang T, Schnoll-Sussman FH, Pochapin MB, Christos PJ, Sherr DL. Increasing incidence of rectal cancer in patients aged younger than 40 years: an analysis of the Surveillance, Epidemiology, and End Results database. *Cancer* 2010;116:4354–4359.
19. Deng Y. Rectal Cancer in Asian vs Western Countries: Why the Variation in Incidence? *Curr Treat Options Oncol* 2017;18:64.
20. Liu PH, Wu K, Ng K, et al. Association of obesity with risk of early-onset colorectal cancer among women. *JAMA Oncol* 2019;5:37–44.
21. Rosato V, Bosetti C, Levi F, et al. Risk factors for young-onset colorectal cancer. *Cancer Causes Control* 2013;24:335–341.
22. Manser CN, Bauerfeind P. Impact of socioeconomic status on incidence, mortality, and survival of colorectal cancer patients: a systematic review. *Gastrointest Endosc* 2014;80:42–60.e9.
23. Kim TJ, Kim ER, Hong SN, Chang DK, Kim YH. Long-term outcome and prognostic factors of sporadic colorectal cancer in young patients: a large institutional-based retrospective study. *Medicine (Baltimore)* 2016;95:e3641.
24. Myers EA, Feingol DL, Forde KA, Arnell T, Jang JH, Whelan RL. Colorectal cancer in patients under 50 years of age: a retrospective analysis of two institutions' experience. *World J Gastroenterol* 2013;19:5651–5657.
25. Abdelsattar ZM, Wong SL, Regenbogen SE, Jomaa DM, Hardiman KM, Hendren S. Colorectal cancer outcomes and treatment patterns in patients too young for average-risk screening. *Cancer* 2016;122:929–934.
26. Teng A, Lee DY, Cai J, Patel SS, Bilchik AJ, Goldfarb MR. Patterns and outcomes of colorectal cancer in adolescents and young adults. *J Surg Res* 2016;205:19–27.
27. You YN, Xing Y, Feig BW, Chang GJ, Cormier JN. Young-onset colorectal cancer: is it time to pay attention? *Arch Intern Med* 2012;172:287–289.
28. Pillozzi E, Maresca C, Duranti E, et al. Left-sided early-onset vs late-onset colorectal carcinoma: histologic, clinical, and molecular differences. *Am J Clin Pathol* 2015;143:374–384.
29. Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterology* 2010;138:2073–2087.e3.

---

#### Reprint requests

Address requests for reprints to: Peter S. Liang, MD, MPH, VA New York Harbor Health Care System, 423 East 23rd Street, 11 North, Gastroenterology, New York, New York 10010. e-mail: [peter.liang@nyumc.org](mailto:peter.liang@nyumc.org); fax: (212) 951-3481.

#### Conflicts of interest

The authors disclose no conflicts.

#### Funding

This work was supported by the ReMission Foundation, National Cancer Institute grant nos. K08CA230162 (to Peter S. Liang) and P30 CA008748 (to Mengmeng Du and Kelli O'Connell), and National Center for Advancing Translational Sciences grant no. UL1 TR002384 (to Mengmeng Du and Kelli O'Connell). This material is the result of work supported in part by resources from the Veterans Health Administration. The views expressed in this article are those of the authors and do not represent the views of the Department of Veterans Affairs.

## Supplementary Materials

### Supplementary Subanalysis 1: Sex and Age

As life expectancy is greater for women than for men, it is possible that the association between early-onset colorectal cancer (CRC) and male sex is due to female predominance in older age groups. To explore this, we performed a subanalysis of sex comparing early-onset CRC patients to a younger subset of late-onset CRC patients who were diagnosed between 50–65 years of age. On univariable analysis, male sex had a *P* value of .14. However, on multivariable logistic regression, male sex was significantly associated with early-onset CRC (odds ratio [OR], 1.40; 95% confidence interval [CI],

1.05–1.86), and the risk estimate was similar to what was observed in the full analysis (OR, 1.44; 95% CI, 1.11–1.87).

### Supplementary Subanalysis 2: Black Race

Several organizations recommend earlier screening for blacks. In 2 subanalyses, we performed multivariable logistic regression comparing black vs non-black race in early-onset CRC vs both control subjects and late-onset CRC. There was no statistically significant association between black race and early-onset CRC compared with either control subjects (OR, 1.42; 95% CI, 0.90–2.15) or late-onset CRC (OR, 1.28; 95% CI, 0.78–2.03).

**Supplementary Table 1.** Sensitivity Analysis Comparing Individuals With Early-Onset CRC vs Young Control Subjects Without Cancer (Excluding Family History of CRC or Personal History of IBD)

Variable	Early onset group (n = 269)	Control group (n = 1122)	Univariable <i>P</i> <sup>a</sup>	Multivariable OR (95% CI)	Multivariable <i>P</i>
<b>Sociodemographic</b>					
Age, y <sup>b</sup>	43 ± 6	45 ± 6			
Male	124 (55)	492 (45)	<.01	1.79 (1.32–2.44)	<.01
Race/ethnicity			.13		
Non-Hispanic white	117 (52)	601 (55)			
Non-Hispanic black	25 (11)	105 (10)			
Non-Hispanic Asian	23 (10)	63 (6)			
Hispanic	9 (4)	44 (4)			
Other	50 (22)	283 (26)			
Income, US dollars	70,785 ± 27,176	72,674 ± 27,619	.35		
High school education, %	84 (10)	85 (9)	.22		
<b>Medical</b>					
BMI, kg/m <sup>2</sup>	27 ± 6	28 ± 6	.07	0.98 (0.95–1.00)	.07
Smoker	58 (27)	302 (29)	.62	8.61 (4.83–15.75)	
Coronary artery disease	8 (4)	40 (4)	>.99		
Hypertension	44 (20)	216 (20)	>.99		
Hyperlipidemia	36 (16)	250 (23)	.03	0.62 (0.41–0.92)	<.02
Stroke	0	7 (0.6)	.61		
Diabetes	17 (8)	64 (6)	.40		

NOTE. Values are mean ± SD or n (%). Variables with *P* < .20 were carried forward to the multivariable logistic regression model.

BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; IBD, inflammatory bowel disease; OR, odds ratio.

<sup>a</sup>Chi-square and Fisher exact tests were used for categorical variables and Student *t* test for continuous variables.

<sup>b</sup>Age at CRC diagnosis in early-onset cohort, at time of search in control subjects (cases and control subjects matched by birth year).

**Supplementary Table 2.** Sensitivity Analysis Comparing Individuals With Early-Onset vs Late-Onset CRC (Excluding Family History of CRC or Personal History of IBD)

Variable	Early onset group (n = 269)	Late-Onset group (n = 2802)	Univariable <i>P</i> <sup>a</sup>	Multivariable OR (95% CI)	Multivariable <i>P</i>
Age at CRC diagnosis, y	43 ± 6	71 ± 11			
Male	124 (55)	1271 (48)	.04	1.44 (1.09–1.93)	.01
Race/ethnicity			<.01 <sup>a</sup>		
Non-Hispanic white	117 (52)	1851 (70)			
Non-Hispanic black	25 (11)	207 (8)		2.15 (1.33–3.36)	<.01
Non-Hispanic Asian	23 (10)	121 (5)		3.12 (1.86–5.04)	<.01
Hispanic	9 (4)	82 (3)		1.94 (0.88–3.78)	.07
Other	50 (22)	383 (15)		2.13 (1.47–3.04)	<.01
Income, US dollars	70,785 ± 27,176	72,270 ± 26,207	.43		
High school education, %	84 ± 10	86 ± 9	.07	0.53 (0.11–2.61)	.43
BMI, kg/m <sup>2</sup>	27 ± 6	28 ± 6	.07	0.98 (0.95–1.00)	.11

NOTE. Values are n (%) or mean ± SD. Variables with *P* < .20 were carried forward to the multivariable logistic regression model.

BMI, body mass index; CI, confidence interval; CRC, colorectal cancer; IBD, inflammatory bowel disease; OR, odds ratio.

<sup>a</sup>Chi-square and Fisher exact tests were used for categorical variables and Student *t* test for continuous variables.