

# Impact of Sociodemographic Disparities and Insurance Status on Survival of Patients with Early-Onset Colorectal Cancer

MOHAMED E. SALEM,<sup>a</sup> ALBERTO PUCCINI,<sup>b,c</sup> SALLY J. TRUFAN,<sup>b</sup> WEI SHA,<sup>b</sup> KUNAL C. KADAKIA,<sup>a</sup> MARION L. HARTLEY,<sup>d</sup> LAURA W. MUSSELWHITE,<sup>a</sup> JAMES T. SYMANOWSKI,<sup>b</sup> JIMMY J. HWANG,<sup>a</sup> DEREK RAGHAVAN<sup>a</sup>

<sup>a</sup>Departments of Solid Tumor Oncology and <sup>b</sup>Cancer Biostatistics, Levine Cancer Institute, Charlotte North Carolina, USA; <sup>c</sup>University of Genova, Ospedale Policlinico San Martino IRCCS, Genoa, Italy; <sup>d</sup>The Ruesch Center for the Cure of Gastrointestinal Cancers at Georgetown Lombardi Comprehensive Cancer Center, Washington DC, USA

Disclosures of potential conflicts of interest may be found at the end of this article.

**Key Words.** Colorectal Cancer • Early-onset • Young adult • Sociodemographic disparity • Socioeconomic status • Survival

## ABSTRACT

**Background.** Low socioeconomic status (SES) has been linked to worse survival in patients with colorectal cancer (CRC); however, the impact of SES on early-onset CRC remains undescribed.

**Materials and Methods.** Retrospective analysis of data from the National Cancer Database (NCDB) between 2004 and 2016 was conducted. We combined income and education to form a composite measure of SES. Logistic regression and  $\chi^2$  testing were used to examine early-onset CRC according to SES group. Survival rates and Cox proportional hazards models compared stage-specific overall survival (OS) between the SES groups.

**Results.** In total, 30,903 patients with early-onset CRC were identified, of whom 78.7% were White; 14.5% were Black. Low SES compared with high SES patients were more likely to be Black (26.3% vs. 6.1%) or Hispanic (25.3% vs. 10.5%), have T4 tumors (21.3% vs. 17.8%) and/or N2 disease (13%

vs. 11.1%), and present with stage IV disease (32.8% vs. 27.7%) at diagnosis ( $p < .0001$ , all comparisons). OS gradually improved with increasing SES at all disease stages ( $p < .001$ ). In stage IV, the 5-year survival rate was 13.9% vs. 21.7% for patients with low compared with high SES. In multivariable analysis, SES (low vs. high group; adjusted hazard ratio [HR<sub>adj</sub>], 1.35; 95% confidence interval [CI], 1.26–1.46) was found to have a significant effect on survival ( $p < .0001$ ) when all of the confounding variables were adjusted. Insurance (not private vs. private; HR<sub>adj</sub>, 1.38; 95% CI, 1.31–1.44) mediates 31% of the SES effect on survival.

**Conclusion.** Patients with early-onset CRC with low SES had the worst outcomes. Our data suggest that SES should be considered when implementing programs to improve the early detection and treatment of patients with early-onset CRC. *The Oncologist* 2021;26:e1730–e1741

**Implications for Practice:** Low socioeconomic status (SES) has been linked to worse survival in patients with colorectal cancer (CRC); however, the impact of SES on early-onset CRC remains undescribed. In this retrospective study of 30,903 patients with early-onset CRC in the National Cancer Database, a steady increase in the yearly rate of stage IV diagnosis at presentation was observed. The risk of death increased as socioeconomic status decreased. Race and insurance status were independent predictors for survival. Implementation of programs to improve access to care and early diagnostic strategies among younger adults, especially those with low SES, is warranted.

## INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer diagnosed in both men and women and the second leading cause of cancer-related death in the U.S. [1]. Over the last

few decades, incidence and mortality of CRC have declined by over 50%, largely because of population-based CRC screening and therapy improvements. In contrast, the

Correspondence: Mohamed E. Salem, M.D., Department of Solid Tumor Oncology, Levine Cancer Institute, 1021 Morehead Medical Drive, Charlotte, North Carolina 28204, USA. Telephone: (980)-442-6410; e-mail: mohamed.salem@atriumhealth.org. Received May 22, 2021; accepted for publication July 13, 2021; published Online First on August 5, 2021. <http://dx.doi.org/10.1002/onco.13908>

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

incidence of CRC, particularly left-sided tumors, in adolescents and young adults (defined as early-onset CRC) has been steadily increasing [2–5], with the sharpest increase among those aged 20 to 34 years [6, 7]. Not only is the incidence of early-onset CRC rising, but the related mortality rates in certain subgroups is increasing [7]. This is notable as patients with early-onset CRC tend to present with advanced-stage disease and more histologically aggressive tumors, such as those with mucinous or signet ring features and poor differentiation [8]. Advanced and aggressive disease at presentation is, of course, associated with worse outcomes [8–11], which highlights the critical need to identify and evaluate symptomatic individuals earlier.

In the U.S., CRC incidence and survival differ not only by patient age but also by gender, race, and geography [12], and many of these disparities are quite complex in nature. Black patients have worse survival rates than White patients, even when diagnosed early [13]. Additionally, CRC incidence is higher, and mortality rates are 40% greater among men than women [14].

In addition to the disease's clinicopathologic features, socioeconomic status (SES), such as education level, income, insurance status, and access to health care, can also impact outcomes [15, 16]. In the US, young adults are less likely to have health insurance and promptly seek medical care [15]. They are also more likely to have lower incomes. All of these factors can result in worse survival.

Socioeconomic determinants of health are well established as significant prognostic factors for patients with CRC. However, the degree of socioeconomic-factor impact on disease features and outcomes in young adults with early-onset CRC have not been well characterized. Furthermore, delineating potential socioeconomic disparities and the risk they pose regarding incidence and survival in early-onset CRC will allow us to address the growing burden of early-onset CRC and improve access to care and outcomes.

Herein, we report on our National Cancer Database (NCDB) analysis to determine the impact of socioeconomic determinants of health on the clinicopathological features of early-onset CRC and patient survival.

---

## SUBJECTS, MATERIALS, AND METHODS

### Data Source and Database

The NCDB contains patient data from 2004 to 2016. The NCDB is a joint project of the Commission on Cancer (CoC) of the American College of Surgeons and the American Cancer Society. NCDB data is derived from the cancer registries of more than 1,500 CoC accredited facilities, and it represents approximately 70% of all new cancer cases in the U.S. [17].

### Patients

Patients aged 18–40 years at time of diagnosis with colorectal cancer (colon, rectosigmoid junction, and rectal cancers) were included in the current study. Patients with appendiceal cancers were excluded.

Of note, the age definition of the group referred to as “adolescent and young adults” is still under debate. Age cutoffs vary widely among published studies: some authors recommend an upper limit of 50 years of age, based on historically recommended CRC screening guidelines in the average-risk population, whereas others select patients below the age of 40 years, based on physiological and pathological variables.

### Patient Characteristics

Patient and tumor characteristics (e.g., age, race, ethnicity, insurance type, tumor grade, clinical American Joint Committee on Cancer (AJCC) stage, pathological AJCC stage, Charlson-Deyo comorbidity index, year of diagnosis, and area of living) were recorded and compared.

### Income Level

Income level, as specified by the NCDB, was determined by matching each patient's ZIP code at the time of diagnosis with data derived from the 2012 American Community Survey on median household income and was adjusted for 2012 inflation. Income categories were based on equally proportioned quartiles. The quartiles of median household income were defined as (a) median household income less than \$38,000 (lowest income level), (b) median household income between \$38,000 to \$47,999, (c) median household income \$48,000 to \$62,999, and (d) median household income greater than or equal to \$63,000 (highest income level).

### Education Level

Education level was determined by matching each patient's ZIP code at the time of diagnosis with data derived from the 2012 American Community Survey on the percentage of people aged 25 years and older who had not graduated high school (earned a high school diploma). Education categories were based on equally proportioned quartiles. The quartiles were defined as (a) 21% or more had not graduated high school (lowest education level), (b) 13%–20.9%, (c) 7%–12.9%, and (d) less than 7% (highest education level).

### Composite SES Measure

To determine the impact of socioeconomic determinants of health on the clinicopathological features and outcomes of individuals with early-onset CRC, we combined two socioeconomic variables—income and education—to create a composite measure of socioeconomic status (SES composite) [18, 19]. The quartile assignments (1, 2, 3, 4) of the income and education measures were added together, and new composite SES groups were created for a combined score of 2–3, 4–5, 6–7, and 8 (supplemental online Table 1).

### Area of Residence

Metropolitan, urban, and rural population size designations are assigned in NCDB using data from the U.S. Department of Agriculture Economic Research Service. However, the number of patients in our early-onset CRC cohort who resided in rural areas was small; therefore, we combined

rural and suburban designations into one category and compared this with the metropolitan population.

### Overall Survival

Overall survival was determined from “any-cause” mortality, as reported by the NCDB. Survival time was defined as the number of months from the date of initial CRC diagnosis to the date of death or last reported follow-up.

### Statistical Analysis

Demographic, clinical, and treatment characteristics were analyzed using descriptive statistics, and differences between comparison groups were assessed using  $\chi^2$  and Kruskal-Wallis tests, as appropriate. Kaplan-Meier survival curves comparing composite SES groups were fitted for the entire cohort and stratified by disease stage. The Cochran-Armitage test was used to examine trends over time for stage at presentation and insurance type. Univariable and multivariable Cox Proportional Hazard (CPH) models were fitted for SES and the potential confounders that were selected a priori to assess their associations with survival. Backward elimination with cutoff  $p < .2$  was used to obtain a final multivariable model. For mediation analysis, insurance status was added to the final multivariable model to determine if adjusting the potential mediator attenuates the effect of SES on survival. Causal mediation analysis with counterfactual framework [20, 21] was performed to determine the proportion of SES effect on survival that is mediated by insurance status. Benjamini-Hochberg procedure [22] was used to calculate FDR adjusted  $p$  value (FDR- $p$ ), with FDR- $p < .05$  as the cutoff for statistical significance. All analyses were completed with SAS 9.4 (Cary, NC).

### Ethics Statement

This study was exempt from review by an institutional review board as the data from the NCDB is deidentified prior to distribution.

## RESULTS

### Patient Characteristics

A total of 30,903 patients between the ages of 18–40 years at CRC diagnosis were identified in the NCDB between 2004 and 2016. Descriptive demographic and disease data are summarized in Table 1. The median patient age was 36 years (interquartile range [IQR], 32–39) for the overall population. Most patients (25,081, 81.2%) were aged between 31 and 40 years, with the remaining 5,822 patients (18.8%) being between 18 and 30 years of age. Fifty-two percent were male; 78.7% were White, 16.3% were Hispanic, and 14.5% were Black. Most patients (69.4%) had left-sided tumors, 19.1% had right-sided tumors, and 11.4% had primary tumors in the transverse colon.

In the overall population, 14.4% had stage I CRC at presentation; 19.4%, stage II; 36.3%, stage III; and 29.9%, stage IV (Table 1). However, for the entire study cohort over the 12-year study period, a steady increase in the yearly rate of

stage IV diagnosis at presentation was observed (from 28.9% in 2004 to 33.6% in 2016,  $p < .0001$ ; supplemental online Table 2). Additionally, we observed a gradual decline in the rate of patients with private insurance over the same 12-year study period (from 74.5% in 2004 to 68.6% in 2016,  $p < .0001$ ; supplemental online Table 3).

### Patient Distribution in the SES Composites

The distributions of patients and their characteristics within the four SES composite groups are shown in Table 1 and supplemental online Table 4. There were 7,044 (22.8%), 8,877 (28.7%), 9,452 (30.6%), and 5,530 (17.9%) in the low, mid-low, mid-high, and high SES groups, respectively. Significant differences were seen in the distribution of race, ethnicity, T and N stage, clinical stage at presentation, presence of comorbidities, rehospitalization within 30 days of surgery, area of residence, and insurance type (all comparisons FDR- $p < .001$ ). Specifically, compared with patients in the high SES group, those in the low SES group were more likely to be Black (26.3% vs. 6.1%; odds ratio [OR], 5.36, 95% confidence interval [CI], 4.74–6.06) or Hispanic (25.3% vs. 10.5%; OR, 2.87; 95% CI, 2.59–3.18); have T4 tumors (21.3% vs. 17.8%; OR, 1.78; 95% CI, 1.49–2.13) and/or N2 disease (13% vs. 11.1%; OR, 1.27; 95% CI, 1.10–1.46); present with stage IV disease (32.8% vs. 27.7%; OR, 1.54; 95% CI, 1.37–1.73); have comorbidities (9.2% vs. 6.4%; OR, 1.47; 95% CI, 1.29–1.69); and be rehospitalized within 30 days of surgery (8.7% vs. 6.9%; OR, 1.30; 95% CI, 1.13–1.49).

Additionally, patients in the low SES group were more likely live in suburban/rural areas (27% vs. 1.9%; OR, 19.6; 95% CI, 15.9–24.0; FDR- $p < .0001$ ), more likely to have no insurance (17.0% vs. 5.6%; OR, 4.93; 95% CI, 4.31–5.63; FDR- $p < .0001$ ), and less likely to have private insurance (52.2% vs. 85.6%; OR, 0.20; 95% CI, 0.18–0.23, FDR- $p < .0001$ ) compared with patients in high SES group.

### Overall Survival

We examined the entire cohort's overall survival according to the SES groups and then stratified by CRC stage at presentation. For the overall cohort, significant differences in OS between the four SES composite groups were observed ( $p < .001$ ), with a 5-year OS rate of 55.8% for the low SES group, 59.3% for the mid-low SES group, 64.3% for the mid-high SES group, and 67.9% in the high SES group (Fig. 1).

We then examined survival differences between SES groups according to the CRC stage at presentation. At all CRC stages, survival positively correlated with SES, where OS gradually improved with improving SES (Fig. 2; Table 2).

For example, median OS and 5-year survival rates for patient with stage IV CRC in the low, mid-low, mid-high, and high SES groups were as follows: 20.9 months (hazard ratio [HR], 1.33; 95% CI, 1.22–1.44) and 13.9%; 21.6 months (HR, 1.23; 95% CI, 1.13–1.33) and 16.5%; 25.0 months (HR, 1.06; 95% CI, 0.98–1.15) and 19.6%; and 25.4 months and 21.7% (reference group; overall FDR- $p < .0001$ ); respectively (Table 2). It can be seen that patients with early-onset CRC with lower SES had the

**Table 1.** Patient, sociodemographic, and clinicopathologic characteristics by socioeconomic status

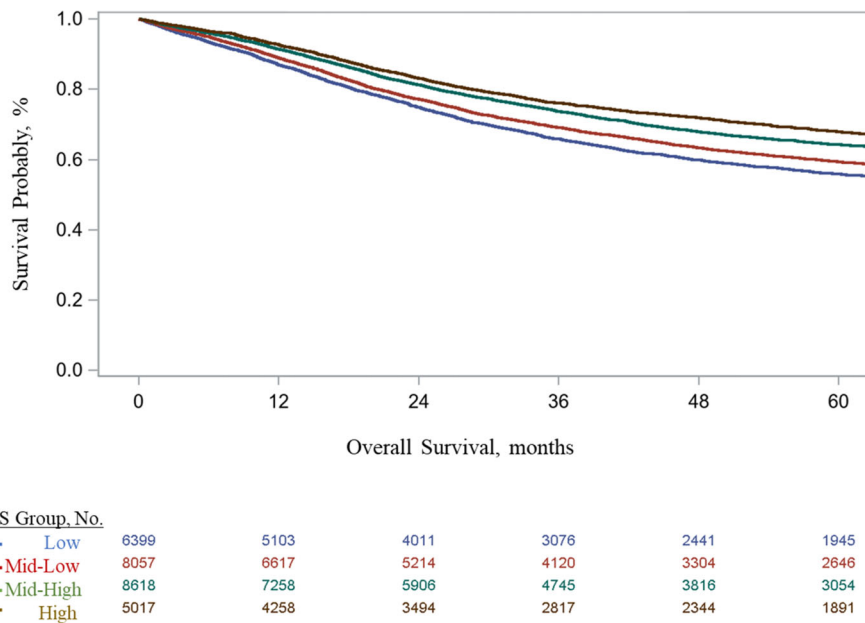
| Characteristics                 | Total<br>(n = 30,903) | Low SES<br>(n = 7,044;<br>22.8%) | Mid-low SES<br>(n = 8,877;<br>28.7%) | Mid-high SES<br>(n = 9,452;<br>30.6%) | High SES<br>(n = 5,530;<br>17.9%) | FDR-p  |
|---------------------------------|-----------------------|----------------------------------|--------------------------------------|---------------------------------------|-----------------------------------|--------|
| Age at diagnosis, yr            |                       |                                  |                                      |                                       |                                   | <.0001 |
| Mean (SD)                       | 34.7 (4.8)            | 34.5 (5.0)                       | 34.7 (4.9)                           | 34.7 (4.7)                            | 35.0 (4.7)                        |        |
| Median                          | 36.0                  | 36.0                             | 36.0                                 | 36.0                                  | 36.0                              |        |
| Q1, Q3                          | 32.0, 39.0            | 32.0, 39.0                       | 32.0, 39.0                           | 32.0, 39.0                            | 33.0, 39.0                        |        |
| Range                           | (18.0–40.0)           | (18.0–40.0)                      | (18.0–40.0)                          | (18.0–40.0)                           | (18.0–40.0)                       |        |
| Age group, yr n (%)             |                       |                                  |                                      |                                       |                                   | <.0001 |
| 18–20                           | 328 (1.1)             | 97 (1.4)                         | 93 (1.1)                             | 80 (0.8)                              | 58 (1.1)                          |        |
| 21–30                           | 5,494 (17.8)          | 1,351 (19.2)                     | 1,583 (17.8)                         | 1,694 (17.9)                          | 866 (15.7)                        |        |
| 31–40                           | 25,081 (81.2)         | 5,596 (79.4)                     | 7,201 (81.1)                         | 7,678 (81.2)                          | 4,606 (83.3)                      |        |
| Sex, n (%)                      |                       |                                  |                                      |                                       |                                   | .6033  |
| Female                          | 14,812 (47.9)         | 3,402 (48.3)                     | 4,239 (47.8)                         | 4,490 (47.5)                          | 2,681 (48.5)                      |        |
| Male                            | 16,091 (52.1)         | 3,642 (51.7)                     | 4,638 (52.2)                         | 4,962 (52.5)                          | 2,849 (51.5)                      |        |
| Race, n (%)                     |                       |                                  |                                      |                                       |                                   | <.0001 |
| White                           | 24,039 (78.7)         | 4,770 (68.2)                     | 7,001 (79.7)                         | 7,631 (81.9)                          | 4,637 (85.2)                      |        |
| Black                           | 4,437 (14.5)          | 1,841 (26.3)                     | 1,267 (14.4)                         | 995 (10.7)                            | 334 (6.1)                         |        |
| Other                           | 2,064 (6.8)           | 382 (5.5)                        | 519 (5.9)                            | 694 (7.4)                             | 469 (8.6)                         |        |
| Missing                         | 363 (1.2)             | 51 (0.7)                         | 90 (1.0)                             | 132 (1.4)                             | 90 (1.6)                          |        |
| Ethnicity, n (%)                |                       |                                  |                                      |                                       |                                   | <.0001 |
| Hispanic                        | 5,029 (16.3)          | 1,780 (25.3)                     | 1,477 (16.6)                         | 1,189 (12.6)                          | 583 (10.5)                        |        |
| Non-Hispanic                    | 25,874 (83.7)         | 5,264 (74.7)                     | 7,400 (83.4)                         | 8,263 (87.4)                          | 4,947 (89.5)                      |        |
| Comorbidities, n (%)            |                       |                                  |                                      |                                       |                                   | <.0001 |
| None                            | 28,375 (91.8)         | 6,399 (90.8)                     | 8,082 (91.0)                         | 8,718 (92.2)                          | 5,176 (93.6)                      |        |
| 1 or more                       | 2,528 (8.2)           | 645 (9.2)                        | 795 (9.0)                            | 734 (7.8)                             | 354 (6.4)                         |        |
| Insurance, n (%)                |                       |                                  |                                      |                                       |                                   | <.0001 |
| Medicare                        | 1,051 (3.5)           | 364 (5.4)                        | 333 (3.8)                            | 238 (2.6)                             | 116 (2.1)                         |        |
| Medicaid                        | 4,787 (15.9)          | 1,731 (25.2)                     | 1,620 (18.7)                         | 1,075 (11.6)                          | 361 (6.7)                         |        |
| None or unknown                 | 3,239 (10.8)          | 1,151 (17.0)                     | 983 (11.3)                           | 799 (8.7)                             | 306 (5.6)                         |        |
| Private                         | 21,024 (69.8)         | 3,541 (52.2)                     | 5,726 (66.1)                         | 7,120 (77.1)                          | 4,637 (85.6)                      |        |
| Missing                         | 802 (0.03)            | 257 (0.04)                       | 215 (0.02)                           | 220 (0.02)                            | 110 (0.02)                        |        |
| Area of living, n (%)           |                       |                                  |                                      |                                       |                                   | <.0001 |
| Metropolitan                    | 25,571 (84.8)         | 5,065 (73.0)                     | 6,836 (78.8)                         | 8,447 (91.8)                          | 5,223 (98.1)                      |        |
| Suburban/rural                  | 4,574 (15.2)          | 1,878 (27.0)                     | 1,842 (21.2)                         | 755 (8.2)                             | 99 (1.9)                          |        |
| Missing                         | 758 (0.02)            | 101 (0.01)                       | 199 (0.02)                           | 250 (0.03)                            | 208 (0.04)                        |        |
| Rehospitalized within 30, n (%) |                       |                                  |                                      |                                       |                                   | .0003  |
| No                              | 27,649 (92.4)         | 6,224 (91.3)                     | 7,991 (92.9)                         | 8,467 (92.3)                          | 4,967 (93.1)                      |        |
| Yes                             | 2,286 (7.6)           | 596 (8.7)                        | 615 (7.1)                            | 709 (7.7)                             | 366 (6.9)                         |        |
| Stage, n (%)                    |                       |                                  |                                      |                                       |                                   | <.0001 |
| I                               | 4,125 (14.4)          | 830 (12.7)                       | 1,113 (13.5)                         | 1,337 (15.2)                          | 845 (16.5)                        |        |
| II                              | 5,570 (19.4)          | 1,238 (18.9)                     | 1,622 (19.7)                         | 1,743 (19.8)                          | 967 (18.9)                        |        |
| III                             | 10,435 (36.3)         | 2,332 (35.6)                     | 2,977 (36.1)                         | 3,232 (36.7)                          | 1,894 (37.0)                      |        |
| IV                              | 8,591 (29.9)          | 2,144 (32.8)                     | 2,530 (30.7)                         | 2,498 (28.4)                          | 1,419 (27.7)                      |        |
| Missing                         | 2,182 (0.07)          | 500 (0.07)                       | 635 (0.07)                           | 642 (0.07)                            | 405 (0.07)                        |        |
| Pathologic T, n (%)             |                       |                                  |                                      |                                       |                                   | <.0001 |
| pTx                             | 556 (2.7)             | 85 (1.9)                         | 163 (2.8)                            | 187 (2.9)                             | 121 (3.2)                         |        |
| pT1                             | 1,986 (9.7)           | 398 (8.8)                        | 507 (8.6)                            | 640 (10.0)                            | 441 (11.6)                        |        |
| pT2                             | 2,798 (13.6)          | 548 (12.1)                       | 780 (13.3)                           | 903 (14.1)                            | 567 (15.0)                        |        |
| pT3                             | 11,181 (54.3)         | 2,521 (55.8)                     | 3,238 (55.1)                         | 3,435 (53.7)                          | 1,987 (52.4)                      |        |
| pT4                             | 4,054 (19.7)          | 963 (21.3)                       | 1,188 (20.2)                         | 1,228 (19.2)                          | 675 (17.8)                        |        |
| Missing                         | 10,328 (33.4)         | 2,529 (35.9)                     | 3,001 (33.8)                         | 3,059 (32.3)                          | 1,739 (31.4)                      |        |

(continued)

**Table 1.** (continued)

| Characteristics                                | Total<br>(n = 30,903) | Low SES<br>(n = 7,044;<br>(22.8%)) | Mid-low SES<br>(n = 8,877;<br>(28.7%)) | Mid-high SES<br>(n = 9,452;<br>(30.6%)) | High SES<br>(n = 5,530;<br>(17.9%)) | FDR-p  |
|--|-----------------------|------------------------------------|--|---|-------------------------------------|--------|
| Pathologic N, n (%)                            |                       |                                    |  |   |                                     | .0267  |
| pN0  | 8,537 (41.7)          | 1,794 (40.0)                       | 2,413 (41.3)                           | 2,695 (42.4)                            | 1,635 (43.3)                        |        |
| pN1  | 9,418 (46.0)          | 2,109 (47.0)                       | 2,710 (46.4)                           | 2,878 (45.3)                            | 1,721 (45.6)                        |        |
| pN2  | 2,503 (12.2)          | 585 (13.0)                         | 717 (12.3)                             | 781 (12.3)                              | 420 (11.1)                          |        |
| Missing  | 10,445 (33.7)         | 2,556 (36.2)                       | 3,037 (34.2)                           | 3,098 (32.7)                            | 1,754 (31.7)                        |        |
| Side of tumor, n (%)                           |                       |                                    |  |   |                                     | <.0001 |
| Left   | 22,210 (71.9)         | 4,891 (69.4)                       | 6,403 (72.1)                           | 6,863 (72.6)                            | 4,053 (73.3)                        |        |
| Right  | 5,644 (18.3)          | 1,347 (19.1)                       | 1,611 (18.1)                           | 1,717 (18.2)                            | 969 (17.5)                          |        |
| Other  | 3,049 (9.9)           | 806 (11.4)                         | 863 (9.7)                              | 872 (9.2)                               | 508 (9.2)                           |        |
| Tumor grade, n (%)                             |                       |                                    |  |   |                                     | .6033  |
| Well differentiated                            | 2,164 (8.0)           | 478 (7.9)                          | 609 (7.9)                              | 652 (7.8)                               | 425 (8.8)                           |        |
| Intermediate                                   | 17,917 (66.5)         | 4,037 (66.6)                       | 5,124 (66.4)                           | 5,575 (67.0)                            | 3,181 (65.7)                        |        |
| Poorly   | 6,017 (22.3)          | 1,373 (22.6)                       | 1,726 (22.4)                           | 1,831 (22.0)                            | 1,087 (22.5)                        |        |
| Undifferentiated                               | 837 (3.1)             | 174 (2.9)                          | 255 (3.3)                              | 260 (3.1)                               | 148 (3.1)                           |        |
| Missing  | 3,968 (12.8)          | 982 (13.9)                         | 1,163 (13.1)                           | 1,134 (11.9)                            | 689 (12.4)                          |        |
| Receipt of surgery of the primary tumor, n (%) |                       |                                    |  |   |                                     | <.0001 |
| No   | 5,410 (17.5)          | 1,452 (20.6)                       | 1,587 (17.9)                           | 1,522 (16.1)                            | 849 (15.4)                          |        |
| Yes  | 25,493 (82.5)         | 5,592 (79.4)                       | 7,290 (82.1)                           | 7,930 (83.9)                            | 4,681 (84.6)                        |        |

Abbreviations: N, nodal involvement per TNM staging; Q, quartile; SES, socioeconomic status; T, tumor size per TNM staging.



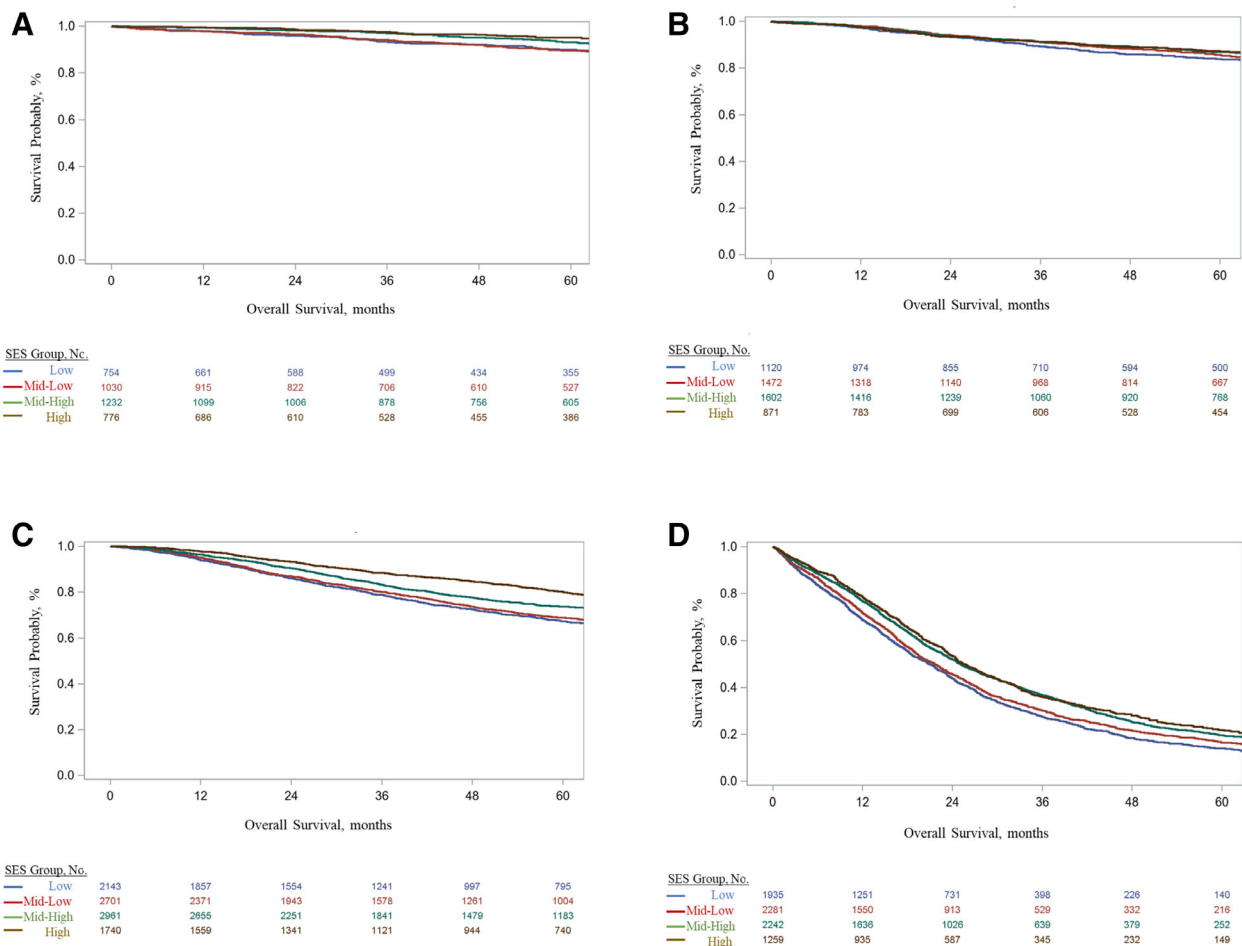
**Figure 1.** Overall survival for the entire cohort according to the SES groups. Overall survival for the entire cohort by socioeconomic composite group. Log-rank *p* value: <.0001  
Abbreviation: SES, socioeconomic status.

worst outcomes, after which OS gradually improved with improving SES.

Similarly, as shown in Table 2, in stage I to III patients, worse survival was seen in the low SES group, which improved incrementally with improving SES.

**Effect of SES on Survival**

In the univariate and multivariable Cox proportional hazard analyses for survival in the entire population, low SES was independently associated with increased risk of death after adjusting for all other covariates (Table 3). Compared with



**Figure 2.** Overall survival in the four SES groups by disease stage. Overall survival by SES composite group with **(A)**: Stage I CRC ( $p$  value < .0001). **(B)**: Stage II CRC ( $p$  value = .0325). **(C)**: Stage III CRC ( $p$  value < .0001). **(D)**: Stage IV CRC ( $p$  value < .0001). Abbreviation: SES, socioeconomic status

the high SES group, the risk of death progressively increased as SES decreased ( $p$  < .0001): low SES group, adjusted HR ( $HR_{adj}$ ), 1.35 (95% CI, 1.26–1.46); mid-low SES group,  $HR_{adj}$ , 1.29 (95% CI, 1.20–1.38); and mid-high SES group,  $HR_{adj}$  = 1.15 (95% CI, 1.0–1.23).

Confounding factors were determined a priori and included age, sex, race, ethnicity, area of living, stage of diagnosis, grade, side of tumor, surgery of the primary tumor, chemotherapy, comorbidity, and the interaction between race and SES. All of them were included in the initial model. Using backward variable selection at cutoff  $p$  < .2, the only term that was eliminated was race  $\times$  SES interaction ( $p$  = .63). Therefore, 11 confounders were adjusted in the final multivariable model.

### Insurance Status Mediates the Effect of SES on Survival

When insurance status was added to the final multivariable model, the effect of SES was reduced to the following:  $HR_{adj}$ , 1.27 (95% CI 1.18–1.36) for the low SES group,  $HR_{adj}$ , 1.24 (95% CI 1.16–1.33) for the mid-low SES group, and  $HR_{adj}$ , 1.13 (95% CI 1.05–1.20) for the mid-high SES group, which indicates insurance may mediate a portion of the SES effect on survival. Because the current causal mediation

analysis procedure cannot analyze a mediator with more than two categories, insurance status was grouped into two categories: uninsured (including Medicaid) versus insured. Patients with cancer diagnosis can receive Medicaid if they are uninsured, but the duration of Medicaid coverage is not available; therefore, Medicaid was grouped with uninsured. Insurance status was found to be significantly associated with survival (FDR- $p$  < .0001; Table 4). Compared with the insured population, patients without insurance were associated with a 28% increased risk of death ( $HR_{adj}$ , 1.28; 95% CI, 1.22–1.34). Causal mediation analysis found that insurance status was a significant mediator (FDR- $p$  < .0001), and it explains 19.45% of the SES effect on survival status (Table 5).

For a thorough investigation, we also used two other ways to categorize insurance status: (a) uninsured (not including Medicaid) versus insured and (b) without private insurance versus with private insurance, and repeated the mediation analysis described above. Both analyses (a) and (b) also found significant mediation effect in insurance status (Table 5). Compared with patients with private insurance, patients without private insurance were associated with a 38% increased risk of death ( $HR_{adj}$ , 1.38, 95% CI, 1.31–1.44). Private insurance status was found to mediate 31.19% of the SES effect on survival (FDR- $p$  < .0001).

**Table 2.** Overall survival by stage and socioeconomic group

| Variable                      | <i>n</i> | Events,<br><i>n</i> (%) | 5-yr survival rate<br>(95% CI), % | Cox univariate hazard<br>ratio (95% CI) | Cox univariate<br>score, FDR- <i>p</i> |
|-------------------------------|----------|-------------------------|-----------------------------------|---|--|
| Stage I ( <i>n</i> = 3,803)   |          |                         |                                   |   | <.0001                                 |
| Low SES                       | 760      | 85 (11)                 | 89.9 (87.4–92.5)                  | 2.21 (1.53–3.20)                        |  |
| Mid-low SES                   | 1,032    | 114 (11)                | 89.7 (87.6–91.9)                  | 2.03 (1.42–2.89)                        |  |
| Mid-high SES                  | 1,235    | 88 (7)                  | 93.3 (91.6–95.0)                  | 1.31 (0.91–1.89)                        |  |
| High SES (referenced)         | 776      | 42 (5)                  | 95.1 (93.3–96.9)                  |   |  |
| Stage II ( <i>n</i> = 5,066)  |          |                         |                                   |   | .0325                                  |
| Low SES                       | 1,120    | 182 (16)                | 83.8 (81.3–86.3)                  | 1.32 (1.05–1.67)                        |  |
| Mid-low SES                   | 1,472    | 206 (14)                | 85.6 (83.5–87.7)                  | 1.10 (0.87–1.38)                        |  |
| Mid-high SES                  | 1,603    | 210 (13)                | 86.9 (85.0–88.8)                  | 1.01 (0.81–1.27)                        |  |
| High SES (referenced)         | 871      | 118 (14)                | 87.0 (84.5–89.5)                  |   |  |
| Stage III ( <i>n</i> = 9,545) |          |                         |                                   |   | <.0001                                 |
| Low SES                       | 2,143    | 655 (30)                | 67.3 (65.0–69.6)                  | 1.76 (1.55–2.01)                        |  |
| Mid-low SES                   | 2,701    | 776 (29)                | 68.9 (66.9–70.9)                  | 1.65 (1.45–1.87)                        |  |
| Mid-high SES                  | 2,961    | 765 (26)                | 73.7 (71.9–75.5)                  | 1.41 (1.24–1.60)                        |  |
| High SES (referenced)         | 1,740    | 332 (19)                | 80.0 (77.8–82.3)                  |   |  |
| Stage IV ( <i>n</i> = 7,721)  |          |                         |                                   |   | <.0001                                 |
| Low SES                       | 1,937    | 1,502 (78)              | 13.9 (12.1–15.8)                  | 1.33 (1.22–1.44)                        |  |
| Mid-low SES                   | 2,281    | 1,765 (77)              | 16.5 (14.7–18.2)                  | 1.23 (1.13–1.33)                        |  |
| Mid-high SES                  | 2,244    | 1,665 (74)              | 19.6 (17.7–21.5)                  | 1.06 (0.98–1.15)                        |  |
| High SES (referenced)         | 1,259    | 900 (71)                | 21.7 (19.1–24.4)                  |   |  |

Abbreviations: CI, confidence interval; SES, socioeconomic status.

Young patients who receive Medicare usually have some type of disability, and the disability data were not available. To remove the potential bias caused by disability on survival, as a sensitivity analysis, we removed patients with Medicare from the mediation analysis. This analysis still found significant mediation effect in insurance status (supplemental online Tables 5, 6).

The interaction effect between insurance and SES was initially included in the Cox regression analysis but was not found to be significant in any of the analyses described above. Therefore, it was removed from the mediation analysis.

### Other Predictors of Survival

Furthermore, race was also found to be a significant predictor of survival (FDR-*p* < .0001), with Black patients having a 20% increased risk of death relative to White patients (HR<sub>adj</sub>, 1.20; 95% CI, 1.13–1.27).

Finally, after adjusting for the impact of several other factors, including our composite SES variable and insurance status, the risk of death was significantly higher in patients living in suburban/rural areas (HR<sub>adj</sub>, 1.12; 95% CI, 1.05–1.19; FDR-*p* < .001) compared with those living in metropolitan areas.

### Association between Income and Education Status in Patients with Early-Onset CRC

The distributions of education groups within each income quartile are shown in supplemental online Figure 1. It can be seen that the high school diploma rate was moderately correlated with income (Spearman correlation = 0.681).

The lowest quartile of education, representing a > 21% high school dropout rate, was inversely correlated with income, whereas the highest quartile, representing only a 7% dropout rate, was positively correlated with income.

### DISCUSSION

Increasing rates of early-onset CRC poses a global health and economic problem and a significant burden on patients, families, and health care systems. During the U.S.'s last 5-year assessment, annual incidence rates in adults aged <50 years increased by 2.2% [13]. Siegel et al. (2020) reported that the observed overall 2% annual increase in CRC incidence is driven by trends in non-Hispanic Whites aged younger than 50 years [1]. Furthermore, Bailey et al. predicted that, by 2030, CRC incidence rates will have increased by 90.0% for patients aged 20 to 34 years and 27.7% for patients aged 35 to 49 years [6].

In 2018, the sharp increase in early-onset CRC incidence and mortality prompted the American Cancer Society to publish a qualified recommendation to screen beginning at 45 years of age instead of 50 years [23].

Furthermore, on May 18, 2021, the U.S. Preventive Services Task Force published a final recommendation statement on screening for colorectal cancer and now recommends that screening start at age 45 (B grade recommendation; [https://uspreventiveservicestaskforce.org/uspstf/sites/default/files/file/supporting\\_documents/colorectal-cancer-screening-final-rec-bulletin.pdf](https://uspreventiveservicestaskforce.org/uspstf/sites/default/files/file/supporting_documents/colorectal-cancer-screening-final-rec-bulletin.pdf)). Lowering the recommended initial screening age from 50 to 45 years is a definite

**Table 3.** Univariable and multivariable Cox proportional hazard models for survival in the overall cohort

| Variable                      | n      | Events,<br>n (%) | Median,<br>mo | 5-yr<br>survival %<br>(95% CI) | Univariable model                              |                              | Multivariable model   |  |
|-------------------------------|--------|------------------|---------------|--------------------------------|--|------------------------------|---|--|
|                               |        |                  |               |                                | Cox<br>univariable<br>hazard ratio<br>(95% CI) | Cox<br>univariable,<br>FDR-p | Cox<br>multivariable<br>hazard ratio,<br>adjusted<br>(95% CI) | Cox<br>multivariable<br>likelihood ratio,<br>FDR-p<br>(n = 24,493) |
| SES group                     |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| Low                           | 7,044  | 2,588 (37)       | 90.6          | 55.8 (54.4–57.2)               | 1.53 (1.43–1.63)                               |                              | 1.35 (1.26–1.46)  |  |
| Mid-low                       | 8,877  | 3,025 (34)       | 117.5         | 59.3 (58.1–60.5)               | 1.36 (1.28–1.45)                               |                              | 1.29 (1.20–1.38)  |  |
| Mid-high                      | 9,452  | 2,890 (31)       | NR            | 64.3 (63.1–65.4)               | 1.16 (1.09–1.23)                               |                              | 1.15 (1.07–1.23)  |  |
| High                          | 5,530  | 1,490 (27)       | NR            | 67.9 (66.4–69.3)               |  |                              |   |  |
| Age group, yr                 |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| 18–30                         | 5,822  | 2,052 (35)       | 103.2         | 57.2 (55.7–58.8)               | 1.19 (1.14–1.25)                               |                              | 1.15 (1.09–1.21)  |  |
| 31–40                         | 25,081 | 7,941 (32)       | 155.8         | 62.6 (61.9–63.3)               |  |                              |   |  |
| Sex                           |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| Female                        | 14,812 | 4,567 (31)       | NR            | 63.5 (62.6–64.4)               |  |                              |   |  |
| Male                          | 16,091 | 5,426 (34)       | 120.1         | 59.8 (58.8–60.7)               | 1.15 (1.11–1.20)                               |                              | 1.18 (1.13–1.23)  |  |
| Race                          |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| White                         | 24,039 | 7,549 (31)       | NR            | 62.9 (62.2–63.7)               |  |                              |   |  |
| Black                         | 4,437  | 1,737 (39)       | 71.9          | 53.1 (51.3–54.8)               | 1.38 (1.31–1.45)                               |                              | 1.20 (1.13–1.27)  |  |
| Other                         | 2,064  | 593 (29)         | NR            | 64.0 (61.4–66.5)               | 0.97 (0.89–1.05)                               |                              | 0.93 (0.85–1.02)  |  |
| Ethnicity                     |        |                  |               |                                |  | .222                         |   | .0044  |
| Non-Hispanic                  | 23,534 |                  |               |                                |  |                              |   |  |
| Hispanic                      | 4,642  |                  |               |                                | 0.97 (0.92–1.02)                               |                              | 0.92 (0.86–0.97)  |  |
| Stage at diagnosis            |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| I                             | 4,125  | 329 (8)          | NR            | 92.0 (91.0–93.0)               |  |                              |   |  |
| II                            | 5,570  | 716 (13)         | NR            | 85.9 (84.8–87.0)               | 1.68 (1.48–1.92)                               |                              | 1.63 (1.43–1.87)  |  |
| III                           | 10,435 | 2,528 (24)       | NR            | 72.0 (71.0–73.1)               | 3.46 (3.09–3.88)                               |                              | 3.54 (3.14–4.01)  |  |
| IV                            | 8,591  | 5,832 (68)       | 23.2          | 17.6 (16.6–18.6)               | 18.69 (16.72–20.90)                            |                              | 15.52 (13.76–17.52)   |  |
| Grade                         |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| Well/mod/none                 | 24,886 | 7,291 (29)       | NR            | 64.9 (64.1–65.6)               |  |                              |   |  |
| Poorly diff                   | 6,017  | 2,702 (45)       | 52.6          | 48.5 (47.1–50.0)               | 1.75 (1.67–1.82)                               |                              | 1.67 (1.59–1.75)  |  |
| Side of tumor                 |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| Left                          | 22,210 | 6,922 (31)       | 146.3         | 63.0 (62.2–63.8)               |  |                              |   |  |
| Right                         | 5,644  | 1,831 (32)       | NR            | 60.6 (59.1–62.1)               | 1.10 (1.05–1.16)                               |                              | 1.23 (1.16–1.30)  |  |
| Other                         | 3,049  | 1,240 (41)       | 71.1          | 52.8 (50.8–54.9)               | 1.49 (1.40–1.58)                               |                              | 1.12 (1.05–1.20)  |  |
| Surgery of the primary tumor  |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| No                            | 5,410  | 3,210 (59)       | 19.9          | 21.8 (20.4–23.2)               | 4.56 (4.37–4.76)                               |                              | 2.42 (2.29–2.54)  |  |
| Yes                           | 25,493 | 6,783 (27)       | NR            | 69.3 (68.6–70.0)               |  |                              |   |  |
| Chemotherapy                  |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| Had chemo                     | 20,931 | 7,599 (36)       | 93.2          | 56.9 (56.1–57.7)               | 1.58 (1.51–1.66)                               |                              | 0.83 (0.79–0.88)  |  |
| No chemo                      | 9,880  | 2,355 (24)       | NR            | 72.0 (71.0–73.0)               |  |                              |   |  |
| Comorbidity                   |        |                  |               |                                |  | <.0001                       |   | <.0001   |
| None                          | 28,375 | 9,115 (32)       | 149.7         | 61.9 (61.2–62.5)               |  |                              |   |  |
| 1 or more                     | 2,528  | 878 (35)         | 108.1         | 58.1 (55.8–60.4)               | 1.17 (1.09–1.25)                               |                              | 1.17 (1.09–1.26)  |  |
| Insurance status <sup>a</sup> |        |                  |               |                                |  | <.0001                       |   | N/A  |
| Medicaid                      | 4,787  | 1,897 (40)       | 62.0          | 50.5 (48.7–52.2)               | 1.07 (0.995–1.15)                              |                              | N/A   |  |
| Medicare                      | 1,051  | 463 (44)         | 54.3          | 47.8 (44.1–51.4)               | 1.17 (1.05–1.30)                               |                              | N/A   |  |
| No insurance/<br>unknow       | 3,239  | 1,268 (39)       | 69.3          | 52.5 (50.4–54.5)               | 1.61 (1.52–1.71)                               |                              | N/A   |  |
| Private                       | 21,024 | 6,068 (29)       |               | 66.2 (65.4–66.9)               |  |                              | N/A   |  |

(continued)



**Table 3.** (continued)

| Variable       | n      | Events,<br>n (%) | Median,<br>mo | 5-yr<br>survival %<br>(95% CI) | Univariable model                              |                              | Multivariable model   |  |
|----------------|--------|------------------|---------------|--------------------------------|--|------------------------------|---|--|
|                |        |                  |               |                                | Cox<br>univariable<br>hazard ratio<br>(95% CI) | Cox<br>univariable,<br>FDR-p | Cox<br>multivariable<br>hazard ratio,<br>adjusted<br>(95% CI) | Cox<br>multivariable<br>likelihood ratio,<br>FDR-p<br>(n = 24,493) |
| Area of Living |        |                  |               |                                |  | <.0001                       |   | .0002  |
| Metro          | 25,571 | 8,118 (32)       | 155.8         | 62.3 (61.6–63.0)               |  |                              |   |  |
| Urban/rural    | 4,574  | 1,645 (36)       | 101.1         | 57.2 (55.5–58.9)               | 1.18 (1.11–1.24)                               |                              | 1.12<br>(1.05–1.19)   |  |

<sup>a</sup>Insurance status was considered as a mediator (not a confounder). Therefore, it was not included in the multivariable model to evaluate the effect of SES on survival.

Abbreviations: CI, confidence interval; NR, not reached; Poorly diff., poorly differentiated; mod., moderate differentiation; SES, socioeconomic status; Undiff., undifferentiated.

**Table 4.** Effect of insurance on survival

| Insurance status              | Effect of insurance in the<br>multivariable cox regression<br>model with insurance added <sup>a</sup> |        |
|-------------------------------|---|--------|
|                               | Hazard ratio (95% CI)   | FDR-p  |
| Insurance status 1 (2 levels) |   | <.0001 |
| Uninsured + Medicaid          | 1.28 (1.22–1.34)  |        |
| Insured                       | reference   |        |
| Insurance status 2 (2 levels) |   | <.0001 |
| Uninsured                     | 1.22 (1.14–1.31)  |        |
| Insured                       | reference   |        |
| Insurance status 3 (2 levels) |   | <.0001 |
| No private insurance          | 1.38 (1.31–1.44)  |        |
| Private insurance             | reference   |        |

<sup>a</sup>The outcome of the model was survival time. The exposure was socioeconomic status.

The confounders were age, sex, race, ethnicity, stage of diagnosis, grade, side of tumor, surgery of the primary tumor, chemotherapy, comorbidity, area of living.

Insurance was added to the model to test its mediation effect (mediation analysis results are shown in Table 5).

step in the right direction and will likely lead to early detection and diagnosis of CRC in individuals aged 45–49 years; however, this policy change will have no impact on patients younger than 45 years of age at diagnosis, which includes patients with early-onset CRC as defined in this study, for whom the increase in incidence rates is the among the highest. Hence, raising awareness of CRC symptoms through education and then providing timely access to care for younger adults is critical to early-stage diagnosis.

Patients with early-onset CRC often present with advanced-stage disease at diagnosis compared with older patients [24, 25], perhaps because CRC is least expected in younger individuals, and initial symptoms are often attributed to other etiologies, resulting in a delayed cancer diagnosis.

We demonstrated that younger patients with low SES exhibit similar tumor location and grade to patients with higher SES, yet they were more likely to present with T4 tumors, N2 disease, and stage IV disease. These findings

may be due to a lack of awareness and recognition of symptoms; inherent shame at presenting with symptoms such as diarrhea and apparent anal bleeding; limited access to health care, particularly among patients with low SES; and inability to afford necessary treatment, all of which can lead to later stage presentation and delayed diagnosis.

Previously, investigators have studied differences in outcomes of patients with CRC according to insurance status [26] and income [27]. Traditionally, young adults have the highest uninsured rate in the country [28]. The current study suggests that insurance status might account for lower survival in early-onset CRC. In this study, we show that patients with early-onset CRC with low SES were 80% less likely to have private health insurance and more likely to have no insurance than those with high SES. Furthermore, our study suggested that patients with early-onset CRC with Medicaid insurance or no insurance had a 28% increased risk of death relative to those with insurance, and patients without private insurance had a 38% increased risk of death compared with those with private insurance, even after adjusting for other factors. Private insurance status mediates 31% of the SES effect on the survival of early-onset CRC.

Insurance status impacts cancer outcomes. Several studies have demonstrated that privately insured patients with curable cancers, including CRC, have better survival than those with Medicaid insurance [26–28]. In the current study of patients with early-onset CRC, we observed an approximate 6% decline in private insurance rate over the 12-year study period, whereas the rate of stage IV disease at presentation increased by 4.7%. Of note, these trends should be further explored in population-based samples (e.g., Surveillance, Epidemiology, and End Results database). We are currently investigating the possibility of a causal relationship between lack of private insurance and advanced disease in young adults. This is an important issue in the current era of the Affordable Care Act.

In September 2010, the Dependent Coverage Expansion under the Affordable Care Act (ACA) went into effect, allowing young adults aged 26 years and younger coverage under their parents' private health insurance. Recently, novel findings highlighted the role of the ACA in improving access for patients with CRC to cancer care, including a shift to early-stage diagnosis and more timely receipt of adjuvant

**Table 5.** Mediation analysis to determine if insurance status is a significant mediator between SES and CRC survival

| SES Group | Cox multivariable regression <sup>a</sup>   |               | Cox multivariable regression with insurance <sup>b</sup> (2 levels, uninsured + Medicaid vs. insured) added |               | Cox multivariable regression with insurance <sup>b</sup> (2 levels, uninsured vs. insured) added                  |               | Cox multivariable regression with insurance <sup>b</sup> (2 levels, no private vs. private insurance) added |               |
|-----------|---|---------------|---|---------------|---|---------------|---|---------------|
|           | Hazard Ratio (95% CI)   | FDR- <i>p</i> | Hazard Ratio (95% CI)   | FDR- <i>p</i> | Hazard Ratio (95% CI)   | FDR- <i>p</i> | Hazard Ratio (95% CI)   | FDR- <i>p</i> |
|           |   | <.0001        |   | <.0001        |   | <.0001        |   | <.0001        |
| Low       | 1.35 (1.26–1.46)  |               | 1.27 (1.18–1.36)  |               | 1.33 (1.24–1.43)  |               | 1.24 (1.15–1.33)  |               |
| Mid-Low   | 1.29 (1.20–1.38)  |               | 1.24 (1.16–1.33)  |               | 1.28 (1.20–1.37)  |               | 1.22 (1.14–1.31)  |               |
| Mid-high  | 1.15 (1.07–1.23)  |               | 1.13 (1.05–1.20)  |               | 1.14 (1.07–1.22)  |               | 1.11 (1.04–1.19)  |               |
| High      | reference   |               | reference   |               | reference   |               | reference   |               |
|           | Causal mediation analysis <sup>c</sup> with insurance (2 levels, uninsured + Medicaid vs insured) as the mediator |               | Causal mediation analysis <sup>c</sup> with insurance (2 levels, uninsured vs insured) as the mediator      |               | Causal mediation analysis <sup>c</sup> with insurance (2 levels, no private vs private insurance) as the mediator |               |   |               |
|           | Percentage (95% CI) of SES effect mediated by insurance   |               | Percentage (95% CI) of SES effect mediated by insurance   |               | Percentage (95% CI) of SES effect mediated by insurance   |               |   |               |
|           |   |               | <i>p</i> value  |               | <i>p</i> value  |               |   |               |
|           | 19.45 (11.83–27.06)   |               | <.0001  |               | 5.10 (2.14–8.07)  |               | .0007   |               |
|           |   |               |   |               | 31.19 (20.95–41.44)   |               | <.0001  |               |

<sup>a</sup>The outcome of the model was survival time. The exposure was SES.

The confounders were age, sex, race, ethnicity, stage of diagnosis, grade, side of tumor, surgery of the primary tumor, chemotherapy, comorbidity, and area of living.

<sup>b</sup>Insurance was added to the model to test its mediation effect.

<sup>c</sup>Causal mediation analysis was performed to study if insurance is a significant mediator for the effect of SES on survival status (alive, death) when the confounders were adjusted.

chemotherapy [28]. However, the overall benefit for patients with early-onset CRC in terms of survival needs further study, especially given the findings that patients receiving Medicaid have worse outcomes [28].

Additionally, we show that significant racial and ethnic disparities exist among early-onset CRC patients. Hence, patients with low SES were more likely to be Black or Hispanic, more likely to have comorbidities, and less likely to undergo surgery of their primary tumors compared with patients with high SES. Furthermore, multivariate analysis showed that Black patients had a 20% increased risk of death relative to White patients, highlighting that racial and ethnic minorities with early-onset CRC have worse survival than Whites with the same disease.

Interestingly, regardless of income and race or ethnic origin, patients in metropolitan areas seemed to have a lower risk of death compared with those living in rural areas, presumably because of greater access to care, especially at centers of excellence with significant expertise (academic vs. nonacademic), and more clinical trial opportunities [29, 30].

Finally, in the current study, we show significant differences in OS of patients with early-onset CRC according to SES, where the 5-year OS rate gradually improves with increasing SES. This trend was observed at all CRC stages, including stage IV. In the univariate and multivariable Cox proportional hazard analyses for survival in the entire population, low SES was independently associated with increased risk of death after adjusting for all other covariates. This underscores the significant impact of SES and disparities on outcomes among those patients.

Most issues related to SES require community-based resolution, which might involve improved legislation with the

creation of safety nets, community-linked patient navigators, affordable health insurance, and improved social support systems; greater access to health care, including virtual care; improved health education; expansion of access to minority-specific clinical trials; and more funding for disparities research so that the true extent of the problem is known.

The scientific community is increasingly recognizing the issue of impaired survival with lower SES and is addressing possible solutions [31, 32]. The aim of the present investigation was to highlight the considerable knowledge gaps still in existence, as well as and the many details needed to facilitate optimal health planning to address SES disparities among young adults.

Because of the large sample size in this study, small effect can be found to be statistically significant. Therefore, both statistical significance (*p* value) and clinical significance (hazard ratio and its confidence interval) should be taken into consideration when interpreting the results.

To the best of our knowledge, this is the first study to investigate the relationship between all socioeconomic determinants of health and clinicopathological correlates on clinical outcomes, including survival, in the face of early-onset CRC. Nonetheless, our study has several limitations, such as the retrospective nature of the analysis, the heterogeneous nature of our patient population, and the allocation of SES solely by ZIP code. Changes in standard of care over the reported time interval have not been noted. There is also a lack of data on specific treatments received and compliance to adjuvant therapy and surveillance programs as well as details on disease recurrence, type and quality of surgical resection, and prognostic molecular characteristics

of patient tumors. Additionally, it should be noted that the NCDB is not population based, so findings do not necessarily reflect SES differences in the general population. Finally, we reported that having Medicaid or no insurance was associated with increased risk of death results, although these results should be interpreted with caution because the current analysis is limited by the lack of knowing when Medicaid insurance was obtained relative to their cancer diagnosis.

However, despite these limitations, our results clearly demonstrate the impact of SES on the OS of patients with early-onset CRC.

## CONCLUSION

We observed socioeconomic and demographic disparities in survival after a CRC diagnosis in patients with early-onset CRC across all stages of the disease. We further identified persistent disparate outcomes in young adults from low SES groups, even after adjusting for race, insurance status, cancer stage, and comorbidities. Further investigation into the clinical and geographic characteristics of early-onset CRC is warranted to eventually refine our current health care model for early detection, shift to early-stage diagnosis and timely treatment of patients with colon and rectal cancers. Only armed with all this information will we be able to address the rising incidence of early-onset CRC, a potentially curable disease. More efforts are needed to provide better education, improve access, and remove all barriers to care, thus achieving health equity.

## ACKNOWLEDGMENTS

We thank the National Cancer Database for maintaining the publicly available database that made this work possible.

## REFERENCES

- Siegel RL, Miller KD, Goding Sauer A et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin* 2020;70:145–164.
- Meyer JE, Narang T, Schnoll-Sussman FH et al. 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.
- You YN, Xing Y, Feig BW et al. Young-onset colorectal cancer: Is it time to pay attention? *Arch Intern Med* 2012;172:287–289.
- Siegel RL, Miller KD, Fedewa SA et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin* 2017;67:177–193.
- Siegel RL, Medhanie GA, Fedewa SA et al. State variation in early-onset colorectal cancer in the United States, 1995–2015. *J Natl Cancer Inst* 2019;111:1104–1106.
- Bailey CE, Hu 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.
- Siegel RL, Miller KD, Jemal A. Colorectal cancer mortality rates in adults aged 20 to 54 years in the United States, 1970–2014. *JAMA* 2017;318:572–574.
- Salem ME, Battaglin F, Goldberg RM et al. Molecular analyses of left- and right-sided tumors in adolescents and young adults with colorectal cancer. *The Oncologist* 2020;25:404–413.
- Puccini A, Lenz HJ, Marshall JL et al. Impact of patient age on molecular alterations of left-sided colorectal tumors. *The Oncologist* 2019;24:319–326.
- Vatandoust S, Price TJ, Ullah S et al. Metastatic colorectal cancer in young adults: A study from the South Australian population-based registry. *Clin Colorectal Cancer* 2016;15:32–36.
- Tricoli JV, Boardman LA, Patidar R et al. A mutational comparison of adult and adolescent and young adult (AYA) colon cancer. *Cancer* 2018;124:1070–1082.
- Rogers CR, Moore JX, Qeadan F et al. Examining factors underlying geographic disparities in early-onset colorectal cancer survival among men in the United States. *Am J Cancer Res* 2020;10:1592–1607.
- Holowatyj AN, Ruterbusch JJ, Rozek LS et al. Racial/ethnic disparities in survival among patients with young-onset colorectal cancer. *J Clin Oncol* 2016;34:2148–2156.
- American Cancer Society. *Colorectal Cancer Facts & Figures 2020–2022*. Atlanta, GA: American Cancer Society; 2020.
- Koroukian SM, Bakaki PM, Raghavan D. Survival disparities by Medicaid status: An analysis of 8 cancers. *Cancer* 2012;118:4271–4279.
- Lee DY, Teng A, Pedersen RC et al. Racial and socioeconomic treatment disparities in adolescents and young adults with stage II–III rectal cancer. *Ann Surg Oncol* 2017;24:311–318.
- Mallin K, Browner A, Palis B et al. Incident cases captured in the National Cancer Database compared with those in U.S. population based central cancer registries in 2012–2014. *Ann Surg Oncol* 2019;26:1604–1612.
- Miller BJ, Gao Y, Duchman KR. Socioeconomic measures influence survival in osteosarcoma: An analysis of the National Cancer Data Base. *Cancer Epidemiol* 2017;49:112–117.
- Du XL, Fang S, Coker AL et al. Racial disparity and socioeconomic status in association with survival in older men with local/regional stage prostate carcinoma: Findings from a large community-based cohort. *Cancer* 2006;106:1276–1285.

The Department of Solid Tumor Oncology at Levine Cancer Institute funded this work. The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request. The National Cancer Database data dictionary is publicly available: [https://www.facs.org/-/media/files/quality-programs/cancer/ncdb/puf\\_data\\_dictionary.ashx](https://www.facs.org/-/media/files/quality-programs/cancer/ncdb/puf_data_dictionary.ashx). Preliminary results of this study were presented at the Gastrointestinal Cancers Symposium in San Francisco, California, January 20–22, 2020; Abstract 13.

## AUTHOR CONTRIBUTIONS

**Conception/design:** Mohamed E. Salem, Alberto Puccini, Sally J. Trufan  
**Provision of study material or patients:** Mohamed E. Salem, Alberto Puccini, Sally J. Trufan, Wei Sha, Kunal C. Kadakia, Marion L. Hartley, Laura W. Musselwhite, James T. Symanowski, Jimmy J. Hwang, and Derek Raghavan  
**Collection and/or assembly of data:** Mohamed E. Salem, Alberto Puccini, Sally J. Trufan, Wei Sha, Kunal C. Kadakia, Marion L. Hartley, Laura W. Musselwhite, James T. Symanowski, Jimmy J. Hwang, and Derek Raghavan  
**Data analysis and interpretation:** Mohamed E. Salem, Alberto Puccini, Sally J. Trufan, Wei Sha  
**Manuscript writing:** Mohamed E. Salem, Alberto Puccini, Sally J. Trufan, Wei Sha  
**Final approval of manuscript:** Mohamed E. Salem, Alberto Puccini, Sally J. Trufan, Wei Sha, Kunal C. Kadakia, Marion L. Hartley, Laura W. Musselwhite, James T. Symanowski, Jimmy J. Hwang, and Derek Raghavan

## DISCLOSURES

**Mohamed E. Salem:** Taiho Oncology, Astrazeneca, Daiichi Sankyo, Bristol Myers Squibb, Merck, Pfizer, QED Therapeutics, Novartis, Exelixis (C/A) and speaking and relationship. **Jimmy J. Hwang:** Bayer, Bristol Myers Squibb, Deciphera, Incyte, Pfizer, QED Therapeutics (C/A), Deciphera, Incyte (H). The other authors indicated no financial relationships.

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (ET) Expert testimony; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

- 20.** VanderWeele TJ. Explanation in causal inference: Developments in mediation and interaction. *Int J Epidemiol* 2016;45:1904–1908.
- 21.** VanderWeele TJ. Mediation analysis: A practitioner's guide. *Annu Rev Public Health* 2016;37:17–32.
- 22.** Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1995;57:289–300.
- 23.** 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.
- 24.** Dozois EJ, Boardman LA, Suwanthanma W et al. Young-onset colorectal cancer in patients with no known genetic predisposition: Can we increase early recognition and improve outcome? *Medicine (Baltimore)* 2008;87:259–263.
- 25.** Willauer AN, Liu Y, Pereira AAL et al. Clinical and molecular characterization of early-onset colorectal cancer. *Cancer* 2019;125:2002–2010.
- 26.** Tawk R, Abner A, Ashford A et al. Differences in colorectal cancer outcomes by race and insurance. *Int J Environ Res Public Health* 2015;13:ijerph13010048.
- 27.** Warren Andersen S, Blot WJ, Lipworth L et al. Association of race and socioeconomic status with colorectal cancer screening, colorectal cancer risk, and mortality in southern US adults. *JAMA Netw Open* 2019;2:e1917995.
- 28.** Nogueira L, Chawla N, Han X et al. Colorectal cancer care among young adult patients after the dependent coverage expansion under the Affordable Care Act. *J Natl Cancer Inst* 2020;112:1063–1066.
- 29.** Veenstra CM, Epstein AJ, Liao K et al. The effect of care setting in the delivery of high-value colon cancer care. *Cancer* 2014;120:3237–3244.
- 30.** Cabo J, Shu X, Shu XO et al. Treatment at academic centers decreases insurance-based survival disparities in colon cancer. *J Surg Res* 2020;245:265–272.
- 31.** Groman R, Ginsburg J, American College of P. Racial and ethnic disparities in health care: A position paper of the American College of Physicians. *Ann Intern Med* 2004;141:226–232.
- 32.** Raghavan D. Disparities in cancer care: Challenges and solutions. *Oncology (Williston Park)* 2007;21:493–496; discussion 499, 503, 506.



See <http://www.TheOncologist.com> for supplemental material available online.