

Characteristics of carcinoembryonic antigen-producing colorectal cancers: A population based study

Sarah Groover¹, Rashmi Kaul¹, Sarah E. Johnston², Yan D. Zhao², Asha Karippot³, Pragatheeshwar Thirunavukarasu^{3,4}

SUMMARY

Objectives: Serum carcinoembryonic antigen is a tumor marker often found to be elevated in colorectal cancer patients. Elevated carcinoembryonic antigen has been strongly associated with poor prognosis. However, little is known about the patient and tumor characteristics between carcinoembryonic antigen-secreting and non-secreting tumors. **Methods:** We performed a retrospective analysis of all patients (N=164,187) in the Surveillance, Epidemiology and End Results database diagnosed with colorectal adenocarcinoma from 2010 to 2014. All patients were designated as having either positive/elevated (C1) or negative/normal (C0) pretreatment serum carcinoembryonic antigen level. **Results:** Of the 164,187 patients, 68,833 (57.0%) had available carcinoembryonic antigen information, and 33,412 (48.5%) had positive/elevated (C1) antigen levels. Median age was 65 years, and 36,464 (53.0%) were male. Patients with C1 cancers were more likely to be female (Odds ratio 1.06), black (Odds ratio 1.62), separated or never married (Odds ratios 1.50 and 1.49, respectively), higher grade (Odds ratios 1.35, 1.64, and 1.72 of moderately, poorly, and undifferentiated cancers, respectively), and of signet ring cell histology (Odds ratio 1.47) compared to males, whites, married participants, well differentiated grade, and adenocarcinoma histology respectively (P<0.001). Multivariate analysis showed that non-Caucasian race, female gender, unmarried status, distal to sigmoid colon location, increasing tumor invasion beyond muscular layer, increasing nodal involvement, and presence of metastases were independent factors associated with the C1 diagnosis. **Conclusions:** About half of all colorectal adenocarcinomas are associated with elevated pre-treatment serum carcinoembryonic antigen levels. Our study is the first nationwide population-based study quantifying the prevalence of serum carcinoembryonic antigen elevation in the colorectal cancer population, and identifying patient and tumor characteristics associated with elevated carcinoembryonic antigen.

KEY WORDS: Colorectal cancer; adenocarcinoma; carcinoembryonic antigen; serum; biomarkers; epidemiology; multivariate analysis; risk factors; SEER Program; United States

INTRODUCTION

In 2018, an estimated 140,250 Americans will develop colorectal cancer (CRC) and 50,630 will die from the disease (1). Carcinoembryonic antigen (CEA) is a serum tumor marker often found in colorectal cancer patients, and elevated serum CEA has been strongly associated with poor oncological prognosis (2, 3). However, not all colorectal cancer patients produce CEA, and little is known about the patient and tumor characteristics between CEA-secreting and non-CEA-secreting tumors.

The TNM staging system of the American Joint Committee on Cancer (AJCC) utilizes the size and local extent of the primary tumor (T-stage), regional lymph node involvement (N-stage), and presence or absence of metastasis (M-stage) as core elements for staging, given the importance of these components in estimating oncological prognosis (4). With increasing evidence to show that the highest pretreatment serum CEA level is associated with poor prognosis, it has also been recommended to be included in the staging system for colorectal cancer (5-7). Incorporation of CEA into the TNM staging system has shown to significantly impact survival estimates (2, 3). However, the proportion of CRC patients who have elevated CEA levels is unknown. In this study, we aim to determine the prevalence of elevated serum CEA amongst patients with colorectal adenocarcinoma and identify the factors associated with such elevation.

PARTICIPANTS AND METHODS

Data Source and Selection of Patients

The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program is a comprehensive source of US cancer data,

representing 28% of the US population, through 20 population-based cancer registries, namely the Alaska Native Tumor Registry, Arizona Indians, Cherokee Nation, Connecticut, Detroit, Atlanta, Greater Georgia, Rural Georgia, San Francisco-Oakland, San Jose-Monterey, Greater California, Hawaii, Iowa, Kentucky, Los Angeles, Louisiana, New Jersey, New Mexico, Seattle-Puget Sound, and Utah.

We extracted all patients (N = 164,187) diagnosed with adenocarcinomatous CRC between January 1, 2010, and December 31, 2014, by accessing the database named „SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2016 Sub (1973-2014 varying)” (8) (Figure 1). We used the codes „2010,” „2011,” „2012,” „2013,” and „2014” for the year of diagnosis and „Cecum,” „Ascending colon,” „Hepatic flexure of colon,” „Transverse colon,” „Splenic flexure of colon,” „Descending colon,” „Sigmoid colon,” „Overlapping lesion of colon,” „Colon, NOS,” „Rectosigmoid junction,” and „Rectum, NOS” for the site. Exclusion criteria included patients with lack of positive histological confirmation, lifetime occurrence of another primary malignancy, non-invasive or *in-situ* malignancies, or cases diagnosed at autopsy.

The interpretation of the highest CEA test result was recorded prior to treatment (i.e., preoperative CEA level) and was accessed using the variable „CS site-specific factor 1.” Results were available as positive/elevated; negative/within normal limits; borderline/undetermined whether positive or negative; test ordered (results not in chart); test not ordered; or unknown. We grouped positive/elevated (designated C1) and negative/within normal limits (C0) as those who had CEA information available for analysis. Of the 120,536 histologically proven colorectal

Arch Oncol 2019; 25(1):1-8

Published Online
December 7, 2018
<https://doi.org/10.2298/AO0181029007G>

¹Oklahoma State University Center for Health Sciences, Department of Biochemistry and Microbiology, 1111W 17th St, Tulsa, Oklahoma, USA
²University of Oklahoma Health Sciences Center, Department of Biostatistics and Epidemiology, 1100N Lindsay Ave, Oklahoma City, Oklahoma, USA
³Cancer Treatment Centers of America, 10109E 79th St, Tulsa, Oklahoma, USA
⁴Oklahoma State University Center for Health Sciences, Department of Surgery, 1111W 17th St, Tulsa, Oklahoma, USA

Correspondence to:
Pragatheeshwar Thirunavukarasu, MBBS
Oklahoma State University Center for Health Sciences, 10109E 79th Street, Suite C1155, Tulsa 74133, Oklahoma, USA
prag.t@ctca-hope.com
Office: +1-918-249-6601
Fax: +1-918-249-7514

Received 2018-10-29
Accepted 2018-11-08

ABBREVIATIONS

carcinoembryonic antigen (CEA); colorectal cancer (CRC); Surveillance, Epidemiology and End Results (SEER)



This work is licensed under a Creative Commons Attribution 4.0 license

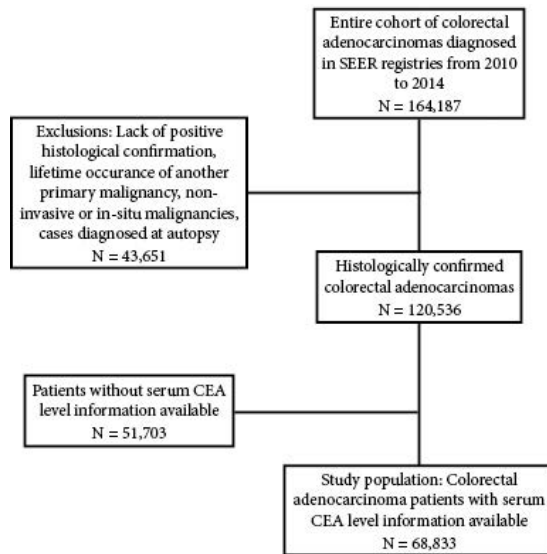


Figure 1. Flow diagram of colorectal adenocarcinoma patients selected for the study

adenocarcinomas, 68,833 patients (57.0%) had CEA information available for analysis.

Age in years, sex, and marital status when the patient was first diagnosed with cancer were obtained. Race information was available as whites; blacks; Asian or Pacific Islanders; American Indians or Alaska Natives; and unknown. Grade of primary tumor, when known, was available as

well differentiated; moderately differentiated; poorly differentiated; undifferentiated; or unknown. Tumor histology was coded in the database as per the *International Classification of Diseases for Oncology, Third Edition* (ICD-O-3) and was accessed using the variable „Site and Morphology. ICD-O-3 Hist/behav, malignant.”

Statistical Analysis

Categorical covariates were summarized with frequencies and percentages, and continuous covariates were summarized with means, medians, minimums, maximums, and standard deviations. We analyzed the data using univariate and multivariate logistic regression to model the odds of a patient having elevated serum CEA levels (i.e. being C1). Exponentiated maximum-likelihood estimates on model coefficients were reported along with p-values, based on the Wald test. From these models we obtained odds ratios (OR) and corresponding 95% confidence intervals (95% CI). As serum CEA elevation is strongly correlated with probability of metastatic disease at diagnosis, we performed further univariate and multivariate analyses on the subset of metastatic patients (defined as patients with stage IV CRC). Statistical analysis was performed using SAS software (9).

RESULTS

Patient and Tumor Characteristics

We performed a univariate analysis of CRC patients characterized by CEA status (C-stage) (Table 1). Of the 68,833 (57.0%) patients with available CEA information, 33,412 (48.5%) patients had positive/elevated (C1) antigen levels. Median age was 65 years and 36,464 (53.0%) patients were male. Compared to C0 cancers, patients with C1 cancers were significantly more likely ($P < 0.001$) to be female and non-white. On analysis of the marital status, patients who were single, divorced, widowed or separated at the time of diagnosis was associated with more than 50% chance of having elevated CEA compared to married individuals. However, individuals who were unmarried but lived with their significant other were significantly less likely to be C1 compared to married individuals. Analysis of tumor characteristics associated with CEA elevation showed that compared to C0 cancers, C1 cancers were significantly more likely ($P < 0.001$) to be of higher grade and of more advanced stage, within each TNM category, as well as overall AJCC stage (Table 1). Signet ring cell pathology was the most common morphology associated with C1 cancers, while adenocarcinoma in preexisting adenomatous polyps and medullary histology were the most common morphologies associated with C0 cancers.

Risk Factors for CEA Secretion

We performed a multivariate analysis to identify independent factors associated with diagnosis C1 tumors (Figure 2). African-American and Asian/Pacific Islander race were more likely to have elevated CEA. Female gender was also associated with slighted higher risk of C1 tumors. Unmarried marital status variables, as described in the univariate analysis, were also associated with C1 tumors. Other factors that emerged as independent factors associated with CEA elevation included left-side tumor location (sigmoid, rectosigmoid junction, and rectum); increasing tumor invasion beyond the *muscularis propria* (T stage); increasing extent of nodal disease (N stage); and presence of metastases (M stage).

| Characteristics | All Participants (N=68,833) | | OR (95% CI) | Type 3 P-value |
|--------------------------------|-----------------------------|-------------------------|------------------|----------------|
| | Normal CEA (N=35,421) | Elevated CEA (N=33,412) | | |
| Sex | | | | |
| Male | 19007 | 17457 | REF | <.001 |
| Female | 16414 | 15955 | 1.06 (1.03-1.09) | |
| Age at diagnosis | | | | |
| Mean/Standard Deviation | 64.47/13.85 | 64.86/13.93 | 1.00 (1.00-1.00) | <.001 |
| Median/Min/Max | 65/16/102 | 65/12/108 | | |
| Race | | | | |
| White | 28321 | 24547 | REF | <.001 |
| Black | 3619 | 5078 | 1.62 (1.55-1.69) | |
| American Indian/Alaska Native | 262 | 293 | 1.29 (1.09-1.53) | |
| Asian or Pacific Islander | 3062 | 3353 | 1.26 (1.20-1.33) | |
| Unknown | 157 | 141 | 1.04 (0.82-1.30) | |
| Hispanic | | | | |
| Spanish-Hispanic-Latino | 4195 | 4138 | REF | 0.029 |
| Non-Spanish-Hispanic-Latino | 31226 | 29274 | 0.95 (0.91-0.99) | |
| Marital Status | | | | |
| Married (including common law) | 19877 | 16183 | REF | <.001 |
| Single (never married) | 5369 | 6500 | 1.49 (1.43-1.55) | |
| Divorced | 3317 | 3592 | 1.33 (1.26-1.40) | |
| Widowed | 4569 | 5058 | 1.36 (1.30-1.42) | |
| Separated | 375 | 459 | 1.50 (1.31-1.73) | |
| Unmarried or Domestic Partner | 92 | 47 | 0.63 (0.44-0.89) | |
| Unknown | 1822 | 1573 | 1.06 (0.99-1.14) | |

Table 1. Univariate analyses of patient and tumor characteristics with CEA elevation

| Characteristics | All Participants (N=68,833) | | OR (95% CI) | Type 3 P-value |
|---|-----------------------------|-------------------------|------------------|----------------|
| | Normal CEA (N=35,421) | Elevated CEA (N=33,412) | | |
| SEER Site | | | | |
| Alaska Natives - 1992+ | 69 | 90 | REF | <.001 |
| Atlanta (Metropolitan) - 1975+ | 954 | 978 | 0.79 (0.57-1.09) | |
| California excluding SF/SJM/LA - 2000+ | 8260 | 7005 | 0.65 (0.47-0.89) | |
| Connecticut - 1973+ | 1302 | 1480 | 0.87 (0.63-1.20) | |
| Detroit (Metropolitan) - 1973+ | 1754 | 1680 | 0.73 (0.53-1.01) | |
| Greater Georgia - 2000+ | 2636 | 2707 | 0.79 (0.57-1.08) | |
| Hawaii - 1973+ | 728 | 772 | 0.81 (0.58-1.13) | |
| Iowa - 1973+ | 1765 | 1393 | 0.61 (0.44-0.83) | |
| Kentucky - 2000+ | 2139 | 2216 | 0.79 (0.58-1.09) | |
| Los Angeles - 1992+ | 3612 | 3407 | 0.72 (0.53-0.99) | |
| Louisiana - 2000+ | 2293 | 2471 | 0.83 (0.60-1.14) | |
| New Jersey - 2000+ | 3515 | 3704 | 0.81 (0.59-1.11) | |
| New Mexico - 1973+ | 686 | 758 | 0.85 (0.61-1.18) | |
| Rural Georgia - 1992+ | 56 | 71 | 0.97 (0.61-1.56) | |
| San Francisco-Oakland SMSA - 1973+ | 2220 | 1769 | 0.61 (0.44-0.84) | |
| San Jose-Monterey - 1992+ | 987 | 824 | 0.64 (0.46-0.89) | |
| Seattle (Puget Sound) - 1974+ | 1657 | 1408 | 0.65 (0.47-0.90) | |
| Utah - 1973+ | 788 | 679 | 0.66 (0.47-0.92) | |
| Primary Site | | | | |
| C18.0-Cecum | 5834 | 5217 | REF | <.001 |
| C18.2-Ascending colon | 5330 | 4270 | 0.90 (0.85-0.95) | |
| C18.3-Hepatic flexure of colon | 1244 | 1051 | 0.94 (0.86-1.03) | |
| C18.4-Transverse colon | 2359 | 2047 | 0.97 (0.90-1.04) | |
| C18.5-Splenic flexure of colon | 794 | 766 | 1.08 (0.97-1.20) | |
| C18.6-Descending colon | 1587 | 1420 | 1.00 (0.92-1.08) | |
| C18.7-Sigmoid colon | 7106 | 6862 | 1.08 (1.03-1.14) | |
| C18.8-Overlapping lesion of colon | 299 | 370 | 1.38 (1.18-1.62) | |
| C18.9-Colon, NOS | 192 | 754 | 4.39 (3.73-5.17) | |
| C19.9-Rectosigmoid junction | 2819 | 3015 | 1.20 (1.12-1.27) | |
| C20.9-Rectum, NOS | 7857 | 7640 | 1.09 (1.04-1.14) | |
| Histology | | | | |
| Adenocarcinoma, NOS | 25036 | 26711 | REF | <.001 |
| Adenocarcinoma in adenomatous polyp | 3254 | 1255 | 0.36 (0.34-0.39) | |
| Tubular adenocarcinoma | 22 | 11 | 0.47 (0.23-0.97) | |
| Papillary adenocarcinoma, NOS | 9 | 6 | 0.62 (0.22-1.76) | |
| Adenocarcinoma in villous adenoma | 773 | 506 | 0.61 (0.55-0.69) | |
| Villous adenocarcinoma | 19 | 11 | 0.54 (0.26-1.14) | |
| Adenocarcinoma in tubulovillous adenoma | 3578 | 1685 | 0.44 (0.42-0.47) | |
| Mucinous adenocarcinoma | 2059 | 2345 | 1.07 (1.00-1.14) | |
| Mucin-producing adenocarcinoma | 227 | 313 | 1.29 (1.09-1.53) | |
| Signet ring cell carcinoma | 327 | 514 | 1.47 (1.28-1.69) | |
| Medullary carcinoma, NOS | 94 | 34 | 0.34 (0.23-0.50) | |
| Adenosquamous carcinoma | 23 | 21 | 0.86 (0.47-1.55) | |
| Grade | | | | |
| Well differentiated; Grade I | 2898 | 1846 | REF | <.001 |
| Moderately differentiated; Grade II | 24443 | 21001 | 1.35 (1.27-1.43) | |
| Poorly differentiated; Grade III | 4988 | 5204 | 1.64 (1.53-1.76) | |
| Undifferentiated; anaplastic; Grade IV | 798 | 874 | 1.72 (1.54-1.92) | |
| Unknown | 2294 | 4487 | 3.07 (2.84-3.32) | |

Table 1. (Continued)

| Characteristics | All Participants (N=68,833) | | OR (95% CI) | Type 3 P-value |
|----------------------|-----------------------------|-------------------------|---------------------|----------------|
| | Normal CEA (N=35,421) | Elevated CEA (N=33,412) | | |
| Overall Stage | | | | |
| 0 | 674 | 171 | 0.98 (0.82-1.16) | <.001 |
| I | 9570 | 2481 | REF | |
| IIA | 9106 | 5517 | 2.34 (2.21-2.47) | |
| IIB | 646 | 571 | 3.41 (3.02-3.85) | |
| IIC | 613 | 911 | 5.73 (5.13-6.41) | |
| IINOS | 22 | 17 | 2.98 (1.58-5.62) | |
| IIIA | 1713 | 569 | 1.28 (1.15-1.42) | |
| IIIB | 7573 | 5739 | 2.92 (2.76-3.09) | |
| IIIC | 1998 | 2380 | 4.59 (4.27-4.95) | |
| IINOS | 113 | 124 | 4.23 (3.27-5.48) | |
| IVA | 1573 | 6877 | 16.86 (15.72-18.09) | |
| IVB | 1203 | 6789 | 21.76 (20.18-23.47) | |
| IVNOS | 137 | 669 | 18.83 (15.59-22.75) | |
| UNK Stage | 480 | 597 | 4.80 (4.22-5.45) | |
| T Stage | | | | |
| T1 | 5918 | 2800 | REF | <.001 |
| T2 | 5931 | 2198 | 0.78 (0.73-0.84) | |
| T3 | 17961 | 16077 | 1.89 (1.80-1.99) | |
| T4a | 2283 | 3574 | 3.31 (3.09-3.55) | |
| T4b | 1702 | 3917 | 4.86 (4.52-5.23) | |
| T4NOS | 45 | 130 | 6.11 (4.34-8.60) | |
| Tis | 674 | 171 | 0.54 (0.45-0.64) | |
| TX | 907 | 4545 | 10.59 (9.73-11.52) | |
| N Stage | | | | |
| N0 | 21728 | 14580 | REF | <.001 |
| N1a | 3803 | 3107 | 1.22 (1.16-1.28) | |
| N1b | 3589 | 3745 | 1.56 (1.48-1.64) | |
| N1c | 470 | 548 | 1.74 (1.53-1.97) | |
| N1N | 1302 | 3181 | 3.64 (3.40-3.90) | |
| N2a | 2130 | 2835 | 1.98 (1.87-2.11) | |
| N2b | 1885 | 3150 | 2.49 (2.34-2.65) | |
| N2N | 122 | 263 | 3.21 (2.59-3.99) | |
| NX | 392 | 2003 | 7.61 (6.82-8.50) | |
| M Stage | | | | |
| M0 | 32508 | 19077 | REF | <.001 |
| M1a | 1573 | 6877 | 7.45 (7.03-7.89) | |
| M1b | 1203 | 6789 | 9.61 (9.02-10.25) | |
| M1N | 137 | 669 | 8.32 (6.92-10.01) | |

Table 1. (Continued)

Increased grade was not associated with diagnosis of C1 tumors on multivariate analysis.

CEA Elevation in Metastatic Cancers

Presence of metastases was strongly associated with the highest risk of C1 tumors, but all patients with metastatic disease do not have elevated serum CEA levels. Hence, we performed an analysis of the subset of patients with metastatic disease only, to identify the factors independently associated with diagnosis of C1 cancers (Figure 3). We included site of

metastatic disease as a factor in the analysis. C1 patients with metastases were significantly more likely to have metastasis to liver (OR 2.86) and lungs (OR 1.53) compared to brain metastasis.

DISCUSSION AND CONCLUSIONS

Colorectal cancer is the third leading cause of cancer and cancer-related mortality in the United States. There have been multiple staging systems in the last several decades, in an attempt to categorize patients into different prognostic groups (i.e. 'stages') and streamline management options. In 1978, the AJCC published their first edition of their TNM system and today, it is the most widely used method for staging colorectal cancers, as well as all other types of cancer (10). The T stage describes the size and local extent of the primary tumor; the N stage describes regional lymph node involvement; and the M stages describes the presence or absence of metastasis (4). The overall stage is determined by combining the scores from these individual categories. By using serum CEA elevation to represent poor prognosis, we can describe the subset of people who do not have this marker and begin to uncover the environmental and genetic influences that play a protective role. Having a better understanding of the patient and tumor characteristics associated with this aggressive subset of colon cancers will help to develop more personalized treatment plans and effective management strategies. Using survival data analyses largely from the Surveillance, Epidemiology, and End Results (SEER) database, the AJCC regularly updates the text with changes to TNM categories, criteria, and prognostic stage groups, and in 2010 they published their seventh edition (11, 12). The limitation with the TNM system is that it is a purely anatomical method of assessing prognosis, without taking into consideration other biological markers associated with oncological natural history and pattern of aggressive behavior. The AJCC is beginning to incorporate non-anatomical prognostic factors into their site-specific staging guidelines (4, 13). Examples include age in thyroid cancer, histologic grade in esophageal cancer and sarcoma, mitotic rate in gastrointestinal stromal tumors, prostate-specific antigen levels and Gleason score in prostate cancer, and serum tumor markers in testicular cancer (11). Several prominent organizations, including the Colorectal Working Group of the AJCC, the American Society of Clinical Oncology (ASCO), the European group on tumor markers recommended the addition of serum carcinoembryonic antigen (CEA) levels to colorectal staging guidelines, (5-7) but the AJCC has yet to incorporate it into colorectal staging guidelines.

CEA is normally secreted on the apical side of colorectal epithelial cells and excreted with feces. However, in CRC, due to loss of polarity in cancer cells, CEA could get expressed on the whole tumor cell surface, allowing for secretion into blood vessels (14). High preoperative serum CEA levels have been well-correlated with increased recurrence and decreased survival in CRC patients (2, 15-25). Furthermore, elevated serum CEA has shown to be a more effective predictor of worse oncological prognosis than N stage (3, 26) and hence, an important factor that could potentially guide the course of treatment (27). CEA is a highly specific serum tumor marker for colorectal cancer, but it is not a sensitive tumor marker because not all colon cancer patients have elevated serum CEA levels (28, 29). While it is generally associated with larger tumor size and more

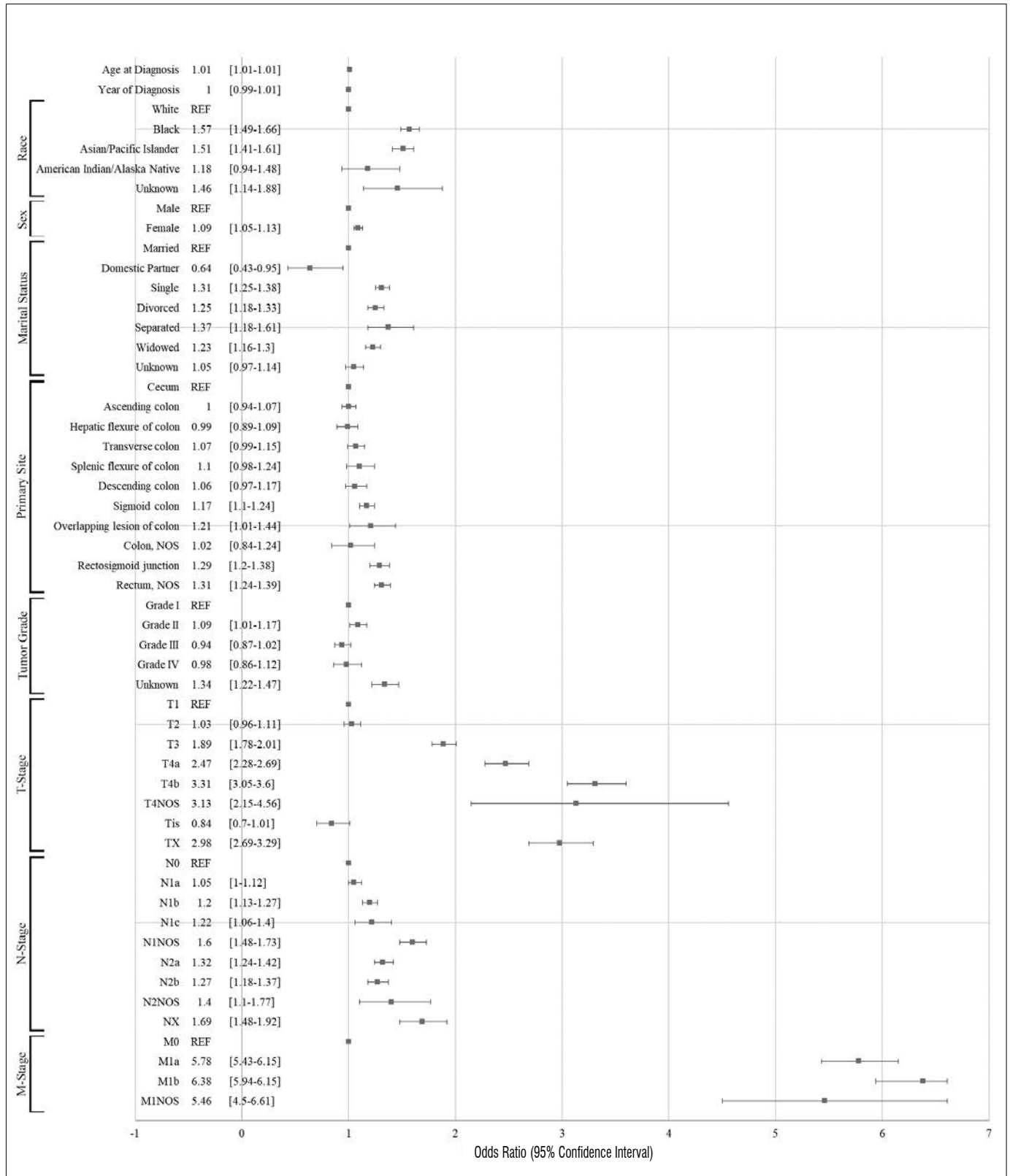


Figure 2. Flow diagram of colorectal adenocarcinoma patients selected for the study

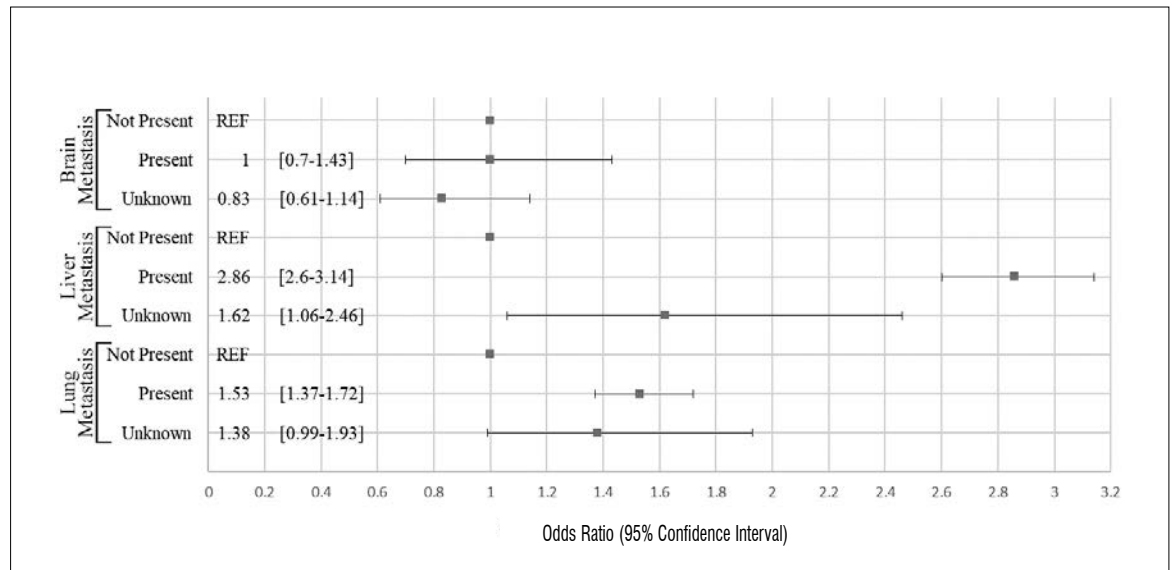


Figure 3. Multivariate analysis of site of metastasis with CEA elevation (metastatic patients only)

extensive metastasis, (30) we have shown in this study that approximately 50% of CRC cases do not have elevated CEA levels.

Another recent US study reviewing the National Cancer Database (NCDB) also reported that approximately 66% of the CRC patients had normal CEA levels (31). Their study emphasized the importance of evaluating preoperative serum CEA, as it was associated with increased hazards of mortality. However, the study did not associate these findings with any specific race or gender. In our study, African-American race was found to be an independent risk factor associated with diagnosis of colorectal cancer with elevated CEA. Supplemental data from another recent NCDB study on early onset CRC also showed that black patients were the most likely to have elevated CEA levels (32). These US population studies strongly support our findings related to elevated CEA as an important prognostic biomarker for improved disease outcome among black patients with CRC. These studies on serum CEA in the clinics can be complemented or validated by tissue CEA expression (33).

Many of the previous CRC epidemiological studies including SEER data studies have not focused on gender associations with any biomarkers in CRC. Our findings related to female gender being an independent risk factor associated with elevated CEA level are unique and need further stratification to show any association with age, race or ethnicity.

Analysis of SEER data has shown that marital status at diagnosis has been well-associated as an independent factor for poor prognosis in many types of cancers, (34-38) including colorectal cancer (39). Often, this is explained by social support mechanisms available for patients who are married at the time of diagnosis. Our study throws new light to a rather new explanation – the possibility that marital status is related to the diagnosis of biologically different cancers or other socioeconomic conditions associated with marital status may also contribute. But, such analysis or discussion is beyond the scope of this manuscript.

There are several limitations in our study, and most of them are inherent to any large population-based database retrospective analysis. The SEER database represents only 28% of the United States population via 20

different registry sites. Although such extensive coverage is appropriate for analyzing incidence and survival on a large scale, uniformity, consistency and accuracy of data entry cannot be verified by authors. Several data points have unrecorded entries, especially for race, marital status, histological grade, site of metastasis, and CEA information. However, we decided to keep all these patients and categories data-unknown patients as a variable in analysis. A drawback of our unusually large sample size is that such large sample sizes may drive even small differences to positive statistical significance, while they may not be clinically significant in daily practice.

Univariate analysis of tumor characteristics showed that C1 cancers were significantly more likely to be located in the sigmoid colon, rectosigmoid junction, or rectum and have a higher histological grade. The latter may be explained by the differences in venous drainage anatomy between colon and rectum. Colonic venous blood drains into either the inferior or superior mesenteric vein, which later drains into the portal venous system that enters the liver, where CEA is metabolized. On the other hand, rectal venous drainage enters the systemic venous system via the inferior vena cava bypassing hepatic metabolism, thereby leading to relatively higher levels of serum CEA for rectal tumors compared to colonic tumors. This is also reflected in our study results, which shows that diagnosis of C1 tumors is associated with more distal location of the primary tumor. In other words, tumors located in the ascending, transverse, and descending colon may need to produce quantitatively more CEA (i.e. to be larger in size and/or more invasive in terms of extent of the tumor) in order to demonstrate the rather equivalent levels of serum CEA as tumors located in the sigmoid colon or rectum. Further studies are needed to study this possibility in more detail.

Although CEA is a well-established marker of prognosis in CRC, about 15% of our initial patient cohort did not have the assay ordered prior to first course of treatment, and 27% were missing CEA information all together. With further accumulation of evidence to support prognostic value of CEA, we hope that the use of presence of elevation of serum CEA

on diagnosis (i.e., C-stage) would become more prevalent. This will also drive more physicians to routinely order serum CEA levels on diagnosis. In brief, serum CEA measurement on diagnosis should be considered standard of care in the initial workup of colorectal malignancies. Future studies studying the factors associated with the non-availability of serum CEA level prior to treatment might shed light on the diversity of quality of health care delivery in the USA.

This is the first and largest database-based epidemiological study quantifying the prevalence of C1 cancers in the colorectal cancer population.

ACKNOWLEDGMENTS

Statistical analyses were performed by personnel from the Oklahoma Shared Clinical & Translational Resources Institute at the University of Oklahoma Health Sciences Center. This institute is supported by the National Institutes of Health, National Institute of General Medical Sciences (U54GM104938).

Declaration of Interests

Authors declare no conflicts of interest

REFERENCES

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(1):7-30.
- Thirunavukarasu P, Sukumar S, Sathaiyah M, Mahan M, Pragatheeshwar KD, Pingpank JF, et al. C-stage in colon cancer: implications of carcinoembryonic antigen biomarker in staging, prognosis, and management. *Journal of the National Cancer Institute*. 2011;103(8):689-97.
- Thirunavukarasu P, Talati C, Munjal S, Attwood K, Edge SB, Francescotti V. Effect of Incorporation of Pretreatment Serum Carcinoembryonic Antigen Levels Into AJCC Staging for Colon Cancer on 5-Year Survival. *JAMA Surg*. 2015;150(8):747-55.
- Sobin LH. TNM: evolution and relation to other prognostic factors. *Seminars in surgical oncology*. 2003;21(1):3-7.
- Compton C, Fenoglio-Preiser CM, Pettigrew N, Fielding LP. American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. *Cancer*. 2000;88(7):1739-57.
- Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, et al. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2006;24(33):5313-27.
- Duffy MJ, Lamerz R, Haglund C, Nicolini A, Kalousova M, Holubec L, et al. Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update. *International journal of cancer*. 2014;134(11):2513-22.
- SEER*Stat Database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases - Linked To County Attributes - Total U.S., 1969-2015 Counties [Internet]. National Cancer Institute. 1973-2014. Available from: www.seer.cancer.gov.
- SAS System for Windows OS. 9.4 ed. Cary, North Carolina, United States of America: SAS Institute Inc.; 2012.
- Puppa G, Sonzogni A, Colombari R, Pelosi G. TNM staging system of colorectal carcinoma: a critical appraisal of challenging issues. *Archives of pathology & laboratory medicine*. 2010;134(6):837-52.
- Edge SB, Byrd DR, Compton C, Fritz AG, Greene FL, Trotti A. *AJCC Cancer Staging Handbook*. 7th ed. New York: Springer; 2010.
- Gunderson LL, Jessup JM, Sargent DJ, Greene FL, Stewart AK. Revised TN categorization for colon cancer based on national survival outcomes data. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2010;28(2):264-71.
- Yarbro JW, Page DL, Fielding LP, Partridge EE, Murphy GP. American Joint Committee on Cancer prognostic factors consensus conference. *Cancer*. 1999;86(11):2436-46.
- Hammarstrom S. The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Seminars in cancer biology*. 1999;9(2):67-81.
- Lou Z, Meng RG, Zhang W, Yu ED, Fu CG. Preoperative carcinoembryonic antibody is predictive of distant metastasis in pathologically T1 colorectal cancer after radical surgery. *World J Gastroenterol*. 2013;19(3):389-93.
- Chen CC, Yang SH, Lin JK, Lin TC, Chen WS, Jiang JK, et al. Is it reasonable to add preoperative serum level of CEA and CA19-9 to staging for colorectal cancer? *The Journal of surgical research*. 2005;124(2):169-74.
- Chu DZ, Erickson CA, Russell MP, Thompson C, Lang NP, Broadwater RJ, et al. Prognostic significance of carcinoembryonic antigen in colorectal carcinoma. Serum levels before and after resection and before recurrence. *Archives of surgery (Chicago, Ill : 1960)*. 1991;126(3):314-6.
- Lucha PA, Jr., Rosen L, Olenwine JA, Reed JF, 3rd, Riether RD, Stasik JJ, Jr., et al. Value of carcinoembryonic antigen monitoring in curative surgery for recurrent colorectal carcinoma. *Diseases of the colon and rectum*. 1997;40(2):145-9.
- Park IJ, Choi GS, Lim KH, Kang BM, Jun SH. Serum carcinoembryonic antigen monitoring after curative resection for colorectal cancer: clinical significance of the preoperative level. *Ann Surg Oncol*. 2009;16(11):3087-93.
- Park YJ, Park KJ, Park JG, Lee KU, Choe KJ, Kim JP. Prognostic factors in 2230 Korean colorectal cancer patients: analysis of consecutively operated cases. *World journal of surgery*. 1999;23(7):721-6.
- Steele G, Jr., Ellenberg S, Ramming K, O'Connell M, Moertel C, Lessner H, et al. CEA monitoring among patients in multi-institutional adjuvant G.I. therapy protocols. *Ann Surg*. 1982;196(2):162-9.
- Wanebo HJ, Rao B, Pinsky CM, Hoffman RG, Stearns M, Schwartz MK, et al. Preoperative carcinoembryonic antigen level as a prognostic indicator in colorectal cancer. *N Engl J Med*. 1978;299(9):448-51.
- Wang JY, Lu CY, Chu KS, Ma CJ, Wu DC, Tsai HL, et al. Prognostic significance of pre- and postoperative serum carcinoembryonic antigen levels in patients with colorectal cancer. *European surgical research Europäische chirurgische Forschung Recherches chirurgicales europeennes*. 2007;39(4):245-50.
- Wang WS, Lin JK, Chiou TJ, Liu JH, Fan FS, Yen CC, et al. Preoperative carcinoembryonic antigen level as an independent prognostic factor in colorectal cancer: Taiwan experience. *Japanese journal of clinical oncology*. 2000;30(1):12-6.
- Yeh CY, Hsieh PS, Chiang JM, Lai CC, Chen JS, Wang JY, et al. Preoperative carcinoembryonic antigen elevation in colorectal cancer. *Hepato-gastroenterology*. 2011;58(109):1171-6.
- Harrison LE, Guillem JG, Paty P, Cohen AM. Preoperative carcinoembryonic antigen predicts outcomes in node-negative colon cancer patients: a multivariate analysis of 572 patients. *J Am Coll Surg*. 1997;185(1):55-9.

- 27 Ogata Y, Murakami H, Sasatomi T, Ishibashi N, Mori S, Ushijima M, et al. Elevated preoperative serum carcinoembryonic antigen level may be an effective indicator for needing adjuvant chemotherapy after potentially curative resection of stage II colon cancer. *Journal of surgical oncology*. 2009;99(1):65-70.
- 28 Goldstein MJ, Mitchell EP. Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer. *Cancer investigation*. 2005;23(4):338-51.
- 29 Moertel CG, O'Fallon JR, Go VL, O'Connell MJ, Thynne GS. The preoperative carcinoembryonic antigen test in the diagnosis, staging, and prognosis of colorectal cancer. *Cancer*. 1986;58(3):603-10.
- 30 Amri R, Bordeianou LG, Sylla P, Berger DL. Preoperative carcinoembryonic antigen as an outcome predictor in colon cancer. *Journal of surgical oncology*. 2013;108(1):14-8.
- 31 Becerra AZ, Probst CP, Tejani MA, Aquina CT, Gonzalez MG, Hensley BJ, et al. Evaluating the Prognostic Role of Elevated Preoperative Carcinoembryonic Antigen Levels in Colon Cancer Patients: Results from the National Cancer Database. *Ann Surg Oncol*. 2016;23(5):1554-61.
- 32 Gabriel E, Attwood K, Al-Sukhni E, Erwin D, Boland P, Nurkin S. Age-related rates of colorectal cancer and the factors associated with overall survival. *J Gastrointest Oncol*. 2018;9(1):96-110.
- 33 Park JW, Chang HJ, Kim BC, Yeo HY, Kim DY. Clinical validity of tissue carcinoembryonic antigen expression as ancillary to serum carcinoembryonic antigen concentration in patients curatively resected for colorectal cancer. *Colorectal Dis*. 2013;15(9):e503-11.
- 34 Li M, Dai CY, Wang YN, Chen T, Wang L, Yang P, et al. Marital status is an independent prognostic factor for tracheal cancer patients: an analysis of the SEER database. *Oncotarget*. 2016;7(47):77152-62.
- 35 Wang X, Cao W, Zheng C, Hu W, Liu C. Marital status and survival in patients with rectal cancer: An analysis of the Surveillance, Epidemiology and End Results (SEER) database. *Cancer Epidemiol*. 2018;54:119-24.
- 36 Wang XD, Qian JJ, Bai DS, Li ZN, Jiang GQ, Yao J. Marital status independently predicts pancreatic cancer survival in patients treated with surgical resection: an analysis of the SEER database. *Oncotarget*. 2016;7(17):24880-7.
- 37 Wu C, Chen P, Qian JJ, Jin SJ, Yao J, Wang XD, et al. Effect of marital status on the survival of patients with hepatocellular carcinoma treated with surgical resection: an analysis of 13,408 patients in the surveillance, epidemiology, and end results (SEER) database. *Oncotarget*. 2016;7(48):79442-52.
- 38 Wu Y, Ai Z, Xu G. Marital status and survival in patients with non-small cell lung cancer: an analysis of 70006 patients in the SEER database. *Oncotarget*. 2017;8(61):103518-34.
- 39 Li Q, Gan L, Liang L, Li X, Cai S. The influence of marital status on stage at diagnosis and survival of patients with colorectal cancer. *Oncotarget*. 2015;6(9):7339-47.