**Name of Journal: *World Journal of Gastroenterology*
Manuscript NO: 47249
Manuscript Type: ORIGINAL ARTICLE**

***Retrospective Cohort Study***

**Prognostic value of preoperative carcinoembryonic antigen/tumor size in rectal cancer**

Cai D *et al.*CEA/tumor size in rectal cancer

Du Cai, Zeng-Hong Huang, Hui-Chuan Yu, Xiao-Lin Wang, Liang-Liang Bai, Guan-Nan Tang, Shao-Yong Peng, Ying-Jie Li, Mei-Jin Huang, Guang-Wen Cao, Jian-Ping Wang, Yan-Xin Luo

**Du Cai, Zeng-Hong Huang, Hui-Chuan Yu, Xiao-Lin Wang, Liang-Liang Bai, Guan-Nan Tang, Shao-Yong Peng, Yingjie Li, Jian-Ping Wang, Yan-Xin Luo,** Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease (Supported by National Key Clinical Discipline), The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510655, Guangdong Province, China

**Du Cai, Zeng-Hong Huang, Shao-Yong Peng, Mei-Jin Huang, Jian-Ping Wang, Yan-Xin Luo,** Department of Colorectal Surgery, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510655, Guangdong Province, China

**Zeng-Hong Huang,** Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis, Sacramento, CA 95817, United States

**Guang-Wen Cao,** Department of Epidemiology, Second Military Medical University, Shanghai 200433, China

**ORCID number:** Du Cai (0000-0002-8894-7973); Zeng-Hong Huang (0000-0003-3625-315X); Hui-Chuan Yu (0000-0001-8357-1615); Xiao-Lin Wang (0000-0002-3911-6911); Liang-Liang Bai (0000-0002-1547-7931); Guan-Nan Tang ( 0000-0001-6063-189X); Shao-Yong Peng (000-0002-5489-4869); Ying-Jie Li (0000-0003-3386-0210); Mei-Jin Huang (0000-0002-5483-4591); Guang-Wen Cao (0000-0002-8094-1278); Jian-Ping Wang (0000-0003-4654-6722), Yan-Xin Luo (0000-0002-5200-3997).

**Author contributions:** Cai D and Huang ZH contributed equally to this paper. Luo YX, Wang JP, Cao GW, and Huang MJ designed the research; Cai D and Huang ZH analyzed the data and drafted the article; Yu HC and Luo YX revised the article; Bai LL, Tang GN, Peng SY, and Li YJ collected and collated the data.

**Supported by** the National Basic Research Program of China (973 Program), No. 2015CB554001; the National Natural Science Foundation of China, No. 81472257 and No. 81502022; the Natural Science Fund for Distinguished Young Scholars of Guangdong Province, No. 2016A030306002; the Outstanding Young Scholars Supported by Special Program to Cultivate Top Experts in Guangdong Province, No. 2015TQ01R454; the Natural Science Foundation of Guangdong Province, No. 2016A030310222 and No. 2018A0303130303; the Science and Technology Program of Guangzhou, No. 2014Y2-00160; the Fundamental Research Funds for the Central Universities (Sun Yat-sen University), No. 16ykpy35; and the Program of Introducing Talents of Discipline to Universities, the National Key Clinical Discipline (2012).

**Institutional review board statement:** This study was reviewed and approved by the Ethics Committee of the Sixth Affiliated Hospital, Sun Yat-sen University.

**Informed consent statement:** Patients were not required to provide informed consent for the study because the analysis used anonymous clinical data.

**Conflict-of-interest statement:** The authors declare that they have no conflict of interest.

**Data sharing statement:** No additional data are available.

**STROBE statement:** The authors have read the STROBE Statement-checklist of items, and the manuscript was prepared and revised according to the STROBE Statement-checklist of items.

**Open-Access:** This is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

**Manuscript source:** Invited manuscript

**Corresponding author: Yan-Xin Luo, MA, MD, PhD, Associate Professor, Chief Doctor, Doctor, Surgical Oncologist,** Department of Colorectal Surgery, Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, The Sixth Affiliated Hospital, Sun Yat-sen University, 26 Yuancun Erheng Road, Guangzhou 510655, Guangdong Province, China. luoyx25@mail.sysu.edu.cn

**Telephone:** +86-13826190263

**Fax:** +86-20-38254221

**Received:** March 12, 2019

**Peer-review started:** March 12, 2019

**First decision:** March 27, 2019

**Revised:** April 4, 2019

**Accepted:** May 18, 2019 **Article in press:
Published online:**

**Abstract**

***BACKGROUND***

Carcinoembryonic antigen (CEA) is a commonly used biomarker in colorectal cancer. However, controversy exists regarding the insufficient prognostic value of preoperative serum CEA alone in rectal cancer. Here, we combined preoperative serum CEA and the maximum tumor diameter to correct the CEA level, which may better reflect the malignancy of rectal cancer.

***AIM***

To assess the prognostic impact of preoperative CEA/tumor size in rectal cancer.

***METHODS***

We retrospectively reviewed 696 stage I to III rectal cancer patients who underwent curative tumor resection from 2007 to 2012. These patients were randomly divided into two cohorts for cross-validation: training cohort and validation cohort. The training cohort was used to generate an optimal cutoff point and the validation cohort was used to further validate the model. Maximally selected rank statistics were used to identify the optimum cutoff for CEA/tumor size. The Kaplan-Meier method and log-rank test were used to plot the survival curve and to compare the survival data. Univariate and multivariate Cox regression analyses were used to determine the prognostic value of CEA/tumor size. The primary and secondary outcomes were overall survival (OS) and disease-free survival (DFS), respectively.

***RESULTS***

In all, 556 patients who satisfied both the inclusion and exclusion criteria were included and randomly divided into the training cohort (2/3 of 556, *n* = 371) and the validation cohort (1/3 of 556, *n* = 185). The cutoff was 2.429 ng/mL per cm. Comparison of the baseline data showed that high CEA/tumor size was correlated with older age, high TNM stage, the presence of perineural invasion, high CEA, and high carbohydrate antigen 19-9 (CA 19-9). Kaplan-Meier curves showed a manifest reduction in 5-year OS (training cohort: 56.7% *vs* 81.1%, *P* < 0.001; validation cohort: 58.8% *vs* 85.6%, *P* < 0.001) and DFS (training cohort: 52.5% *vs* 71.9%, *P* = 0.02; validation cohort: 50.3% *vs* 79.3%, *P* = 0.002) in the high CEA/tumor size group compared with the low CEA/tumor size group. Univariate and multivariate analyses identified CEA/tumor size as an independent prognostic factor for OS (training cohort: hazard ratio (HR) = 2.18, 95% confidence interval (CI): 1.28-3.73, *P* = 0.004; validation cohort: HR = 4.83, 95%CI: 2.21-10.52, *P* < 0.001) as well as DFS (training cohort: HR = 1.47, 95%CI: 0.93-2.33, *P* = 0.096; validation cohort: HR = 2.61, 95%CI: 1.38-4.95, *P* = 0.003).

***CONCLUSION***

Preoperative CEA/tumor size is an independent prognostic factor for patients with stage I-III rectal cancer. Higher CEA/tumor size is associated with worse OS and DFS.

**Key words:** Carcinoembryonic antigen; Carcinoembryonic antigen/tumor size; Rectal cancer; Prognosis; Survival analysis

**© The Author(s) 2019.** Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** This is a retrospective study that sought to evaluate the prognostic value of carcinoembryonic antigen (CEA)/tumor size in rectal cancer, which may better reflect the tumor malignancy. Maximally selected rank statistics identified an optimal cutoff point of 2.429 ng/mL per cm for CEA/tumor size. Kaplan-Meier curves showed a significant reduction in the 5-year overall survival and disease-free survival in the high CEA/tumor size group. Univariate and multivariate analyses identified CEA/tumor size as an independent prognostic factor for stage I to III rectal cancer.

Cai D, Huang ZH, Yu HC, Wang XL, Bai LL, Tang GN, Peng SY, Li YJ, Huang MJ, Cao GW, Wang JP, Luo YX. Prognostic value of preoperative carcinoembryonic antigen/tumor size in rectal cancer. *World J Gastroenterol* 2019; In press

**INTRODUCTION**

Colorectal cancer (CRC) is the third most frequently diagnosed malignancy and one of the leading causes of cancer-related mortality worldwide[1]. Although Western developed countries show a steady or slightly declining trend, the morbidity and mortality of CRC in developing countries like China are still on the rise[2]. Unlike Western countries, the incidence of rectal cancer is higher than that of colon cancer in China and the prognosis of rectal cancer still needs to be improved[3]. Therapy options for CRC have been developed rapidly in the past decade, but selecting optimal treatments for individuals remains a great challenge for clinicians due to the lack of effective markers[4]. In recent years, biomarkers have played an increasingly vital role in the detection and management of CRC[5]. Among the biomarkers, carcinoembryonic antigen (CEA) is one of the most common and most convenient preoperative detecting indexes in patients with colorectal cancer[6].

CEA, a large glycoprotein, has been recommended by the American Society of Clinical Oncology (ASCO) and the European Group on Tumor Markers (EGTM) as a prognostic biomarker that can be used to determine the prognosis and stage of CRC[5,7]. However, controversy still exists regarding the prognostic value of the absolute preoperative serum CEA level in colorectal cancer. Recent studies have noted that CEA is insufficiently sensitive to be used alone, and some researchers have sought new ways to improve its prognostic value by the addition of another factor, such as CD44v6, carbohydrate antigen (CA) 19-9, neutrophil-to-lymphocyte ratio (NLR), or peritoneal carcinomatosis index ratio (PCI)[6,8-11]. Intriguingly, a recent study indicated that postoperative tissue CEA (t-CEA) rather than serum CEA (s-CEA) is an independent prognostic factor in stage I to III CRC[12]. This indicated that we should pay more attention to the local CEA produced by tumor cells rather than the overall serum CEA level. Considering that detecting the CEA produced and secreted by all tumor cells is not realistic, using the ratio of CEA to tumor size may somehow reflect the ability of tumor cells to secrete CEA. Another research group demonstrated that CEA density is a prognostic factor for percutaneous ablation of pulmonary colorectal metastases[13]. Using tumor size to adjust and improve the prognostic value of tumor marker is not uncommon, such as prostate specific antigen density and tumor-infiltrating CD8+ T-cell density[14,15]. Maximum tumor diameter is also a prognostic indicator for some solid tumors including prostate cancer and colorectal liver metastases[16,17]. While the volume-adjusted prostate-specific antigen has been widely studied as a useful marker in prostate cancer[18,19], whether the combination of CEA level and tumor size serves as a novel prognostic factor for rectal cancer remains unresolved.

In this study, we considered both the preoperative serum CEA level and the rectal tumor size and devised the CEA/tumor size, which represents the CEA level adjusted by tumor size, to better reflect the malignancy of rectal cancer. We also refined the insufficient prognostic value of serum CEA. We aimed to apply this new approach to investigate the prognostic impact of the preoperative CEA/tumor size in patients with rectal cancer.

**MATERIALS AND METHODS**

***Patients***

Patients who were diagnosed with stage I to III rectal cancer and underwent a radical excision at the Sixth Afﬁliated Hospital of Sun Yat-Sen University from 2007 to 2012 were studied. This study was approved by the Medical Ethics Committee of the Sixth Affiliated Hospital of Sun Yat-sen University and did not cause any harm to the patients. All retrospective data were obtained from a database maintained by the Sixth Affiliated Hospital of Sun Yat-sen University. The inclusion criteria were as follows: (1) histologically confirmed adenocarcinoma; (2) stage I to III according to the 8th edition of the American Joint Committee on Cancer (AJCC); and (3) radical resection. The following patients were excluded: (1) those with nonprimary cancers; (2) patients who received neoadjuvant chemotherapy and/or radiotherapy; and (3) patients with missing data on preoperative CEA or tumor size. Patients who satisfied both the inclusion and exclusion criteria were randomly divided into two cohorts for cross-validation: training cohort and validation cohort. The training cohort was used to generate an optimal cutoff point and the validation cohort was used to test the applicability of this cutoff point and the model.

***Data collection***

The following data were collected using the Electronic Medical Record System: age, sex, histological features, TNM stage (AJCC), differentiation degree, presence of lymphovascular invasion, presence of perineural invasion, preoperative serum CA 19-9 and CEA levels, maximum tumor diameter, recurrence, and survival time. Follow-up was conducted every three months during the first year after resection, every six months during the next two years, and once a year thereafter. Routine physical examination, serum CEA test, and radiographic examinations including chest radiography, abdominopelvic computed tomographic scanning, or ultrasonography, whole-body bone scanning, double-contrast barium enema, and colonoscopy were performed and recorded six months after resection and yearly thereafter. The follow-up time ended in June 2016, and the follow-up interval varied from three to ten years.

***Statistical analysis***

In our study, we used the maximum diameter in the maximum cross section to represent the tumor size, which was measured by radiologists and pathologists (pathological data are preferred). We defined the CEA/tumor size as the ratio of preoperative CEA level to the maximum tumor diameter. The primary outcome was overall survival (OS), which was defined as the time in months from surgery to death. The secondary endpoint was disease-free survival (DFS), which was defined as the time in months from surgery to disease recurrence, whether radiological or histological. Maximally selected rank statistics were used to identify the optimal discriminator value for the CEA/tumor size, which was conducted in the training cohort. For every potential cutoff point, the absolute value of the standardized log-rank statistic was computed. The cutoff that provided the best separation of the survival outcome into two groups, where the standardized statistics reached their maximum, was selected as the cutoff point. Based on this cutoff, we divided the validation cohort into two groups: high CEA/tumor size group and low CEA/tumor size group. The intergroup comparisons of the clinicopathological variables were performed using the two independent samples *t*-test or Mann-Whitney *U* test for continuous variables, and the chi-square test or two-tailed Fisher’s exact test for discrete variables. The Kaplan-Meier method and log-rank test were used to plot the survival curve and to compare the survival data. Univariate analysis of potential risk factors for each variable was performed using the Cox proportional hazards regression model. Variables with a *P*-value < 0.10 in the univariate analysis were selected to fit the multivariate Cox model. Multivariate analysis using the Cox proportional hazards regression model was used to identify independent risk factors. Variable selection methods, including forward, backward, and stepwise algorithms, as determined by the Akaike information criterion (AIC), were used to construct the appropriate model. The proportional hazards assumption of the Cox regression models was tested by Schoenfeld residuals. All tests were bilateral, and *P*-values < 0.05 were considered statistically significant. All analyses were performed using the R Language for Statistical Computing (version 3.5.1).

**RESULTS**

***Baseline characteristics***

Of the 696 patients diagnosed with rectal cancer who underwent surgical resection from 2007 to 2012, 11 were not histologically confirmed to have adenocarcinoma, 70 received neoadjuvant chemotherapy and/or radiotherapy, and 59 had missing data. Excluding these patients left 566 patients who satisfied both the inclusion and exclusion criteria (Figure 1). These patients were randomly divided into two cohorts: the training cohort (*n* = 371, 2/3 of 566) and the validation cohort (*n* = 185, 1/3 of 566).

Maximally selected rank statistics were performed to determine the optimal value with maximal standardized log-rank statistics. For all 371 rectal cancer patients in the training cohort, the CEA/tumor size of 2.429 ng/mL per cm (*P* = 0.016) provided the best separation of the survival outcomes of the two groups (Figure 2). Based on this cutoff value, 371 patients from the training cohort and 185 patients from the validation cohort were divided into the high CEA/tumor size group and the low CEA/ tumor size group, respectively. As shown in Table 1, high CEA/tumor size was correlated with older age, high TNM stage, the presence of perineural invasion, and high CEA and CA 19-9 levels in the training cohort. Somewhat differently, in the validation cohort, patients with a higher CEA/tumor size only tended to have higher preoperative CEA and CA 19-9 levels. Tumor size, sex, differentiation, and lymphovascular invasion did not differ significantly between the two groups in both cohorts.

***Kaplan-Meier curves***

Kaplan-Meier curves showed a manifest reduction in the 5-year OS (56.7% *vs* 81.1%, *P* < 0.001) and DFS (52.5% *vs* 71.9%, *P* = 0.02) in the high CEA/tumor size group compared with the low CEA/tumor size group in the training cohort (Figures 3A and 4A). The worse outcome of those with high CEA/tumor size was confirmed in the validation cohort, as those patients exhibited a lower 5-year OS (58.8% *vs* 85.6%, *P* < 0.001) and DFS (50.3% *vs* 79.3%, *P* = 0.002) (Figures 3B and 4B).

***Univariate and multivariate analyses***

According to the univariate analysis, age, TNM stage, differentiation, lymphovascular invasion, preoperative CEA and CA 19-9 levels, and CEA/tumor size were selected for the multivariate analysis for OS in both cohorts. As for DFS, the univariate analysis indicated that advanced TNM stage, the presence of lymphovascular invasion, high CEA level, and high CEA/tumor size might be associated with a poor outcome in both cohorts. However, the presence of perineural invasion only showed a significant association with DFS in the training cohort, while poor differentiation and high CA 19-9 level were associated with poor DFS only in the validation cohort (Tables 2 and 3).

To adjust for the influence of potential confounders, the prognostic impact of CEA/tumor size on OS and DFS was further explored by constructing a multivariate Cox proportional hazards model. Forward, backward, and stepwise algorithms determined by the AIC were used to construct the optimum model. All of the above methods generated identical models, and the results were similar in both cohorts. According to the multivariate analysis, older age, poor differentiation, advanced TNM stage, and higher CEA/tumor size were all significantly correlated with a worse OS. With respect to DFS, the significance of TNM stage, lymphovascular invasion, and CEA/tumor size was retained in the final model in both cohorts (Table 4). As a result, CEA/tumor size was significantly associated with OS in both the training cohort (hazard ratio (HR) = 2.18, 95%CI: 1.28-3.73) and in the validation cohort (HR = 4.83, 95%CI: 2.21-10.51). However, CEA/tumor size showed a critical association with DFS in the training cohort (HR = 1.47, 95%CI: 0.93-2.33) and a significant association in the validation cohort (HR = 2.61, 95%CI: 1.38-4.95). Plotting the Schoenfeld residuals against time showed that all the covariates in the Cox proportional hazards model for OS and DFS met the proportional hazard assumption (*P* > 0.05, Figures 5 and 6).

**DISCUSSION**

CEA is reliable for the detection of rectal cancer recurrence and is recommended by the ASCO and EGTM as a prognostic biomarker during routine follow-up for CRC after surgical resection[5,7]. Despite many published studies that have demonstrated the prognostic impact of CEA among CRC patients, no agreement concerning the cutoff values has been established[20-24]. Moreover, Tong *et al*[12] found that postoperative tissue CEA is significantly associated with the prognosis of CRC, and Huo *et al*[13] illustrated that serum CEA density was an independent prognostic factor in patients with colorectal pulmonary metastasis. CEA, as a classic tumor marker, is used to evaluate the biological activity of malignancies, but biological activity will also be affected by tumor quantity. When tumors grow, no matter how clumsily or aggressively, serum CEA level will increase as the expression of CEA increases in proliferating adenocarcinoma cells. Therefore, tumor size is a confounding factor that should be minimized. A new prognostic factor that better reflects the intra-tumor CEA concentration without omission of the tumor volume will be much more accurate than a classic serum CEA test. A comprehensive study stated that tumor size, especially the maximum horizontal tumor diameter, represented a valuable prognosticator in gastric cancer[25]. Another study found a direct relationship between tumor volume in rectal cancer and overall survival[26]. Therefore, we decided to use CEA/tumor size, which is a simple parameter that could reduce the confounding effect of tumor size. Taken together, these results indicate that the ratio of serum CEA to the maximum tumor diameter might be a better marker to assess the tumor’s biological activity and to refine the insufficient prognostic value of serum CEA for rectal cancer.

This is the first study to evaluate the prognostic value of CEA/tumor size for stage I to III rectal cancer. We found that patients with a high CEA/tumor size (over 2.429 ng/mL per cm) had a significantly worse 5-year OS and DFS. Therefore, a correlation exists between the preoperative CEA/tumor size and the prognosis of rectal cancer patients after resection. Patients with high CEA/tumor size tended to have a worse outcome. In our study, no correlation was found between tumor size and survival outcome. Univariate and multivariate analyses showed that CEA/tumor size was independently associated with OS and DFS, while absolute serum CEA was not. This implied that adjusting the confounding effect of tumor size may improve the prognostic value of CEA. Thus, preoperative CEA/tumor size can be used as an independent prognostic factor for patients with stage I-III rectal cancer.

Notably, this study highlights the important relationship between serum CEA and tumor volume, which is in agreement with previous studies. With respect to the prevalence of serum CEA in clinical applications, additional improvement in the accuracy of estimating 5-year outcomes will benefit more patients. In addition, a growing tumor with little change in biological activity will exhibit an increased CEA level and a relatively unchangeable CEA/tumor size. Therefore, CEA/tumor size is not only more accurate but more stable than serum CEA. In patients with identical serum CEA levels, it is necessary to make a decision regarding clinical intervention for patients with smaller maximum tumor diameter. In contrast, a low CEA/tumor size may indicate less aggressive and malignant tumors.

However, we admit that our study has some inherent limitations. First, maximum tumor diameter as an indication of tumor volume is not so precise. Huo *et al*[13] used the spherical formula (4 × π × radius3)/3 to represent the tumor volume since they assumed that pulmonary tumors were spherical. Nevertheless, unlike pulmonary metastases, rectal tumors are not a fixed geometric shape, which means this method is unreliable[26]. Alternatively, the careful delineation of the tumor boundary combined with specific software may provide a more accurate estimation of tumor size. However, maximum tumor diameter represents a quick and convenient method that can be used to roughly estimate tumor volume, and as a result, has more prospects for clinical application. Second, CEA/tumor size cannot be used as part of a routine follow-up index to dynamically monitor the recurrence and metastasis of rectal cancer after surgery. Surgical resection will remove the local tumor, and therefore CEA/tumor size will be unable to be continually calculated. For patients with new-found relapse and metastasis, the value of CEA/tumor size requires further investigation. Beyond that, we also noticed a newly published research study suggesting that postoperative CEA is a better prognostic marker for survival than preoperative CEA in colon cancer[27]. However, postoperative CEA indicates complete resection of the tumor, while CEA/tumor size is focused on tumor malignancy. Third, we did not include patients with neoadjuvant chemotherapy and/or radiotherapy because both of them can influence preoperative CEA and tumor size and may bias our result. Finally, in both cohorts, CEA/tumor size was included in the final Cox model for DFS, which means that CEA/tumor size is an essential factor for DFS. But the *P*-value was 0.003 in the validation cohort and 0.096 in the training cohort, which may result from the insufficient sample size or discrepancy between the two cohorts. Whether CEA/tumor size is really associated with DFS still needs further study.

Preoperative CEA/tumor size is a new method that can be used to predict the outcomes of patients with stage I-III rectal cancer, which may influence the decision-making process for a specific treatment regimen and patient counselling. Since both CEA level and tumor size are routinely measured before surgery, the data of CEA/tumor size can be obtained by simple calculation. This will facilitate the application of CEA/tumor size in clinical practice. Compared with CEA, a great advantage of CEA/tumor size is the ability to figure out those patients with higher CEA but relatively small tumor size. The result of our study suggests that these easily neglected tumors may represent higher malignancy and worse outcome. With the optimization of risk stratification, clinicians can choose individualized treatment options and the outcome of rectal cancer patients can be improved accordingly.

Of course, some limitations of our study design still need to be discussed. As a retrospective study, we were not able to obtain high-level clinical evidence. We also found that some patients did not reach an enough follow-up time, which may influence the accuracy of our result. Since the estimated cutoff point was relatively high, the high-risk group and low-risk group accounted for 20% and 80%, respectively. Although the number of events per variable > 10 in our Cox model, a larger sample size would be better to obtain more reliable results[28]. Therefore, a large-scale prospective study and longer follow-up time are needed and we will try our best to validate our conclusion in future studies. It is also worthwhile for other researchers to further validate our study with new evidence, as we are looking forward to a more accurate prognostic factor for rectal cancer.

In summary, patients with a high preoperative CEA/tumor size have a worse outcome than those with a low CEA/tumor size. Preoperative CEA/tumor size may play an important role in prognosis and treatment decisions of rectal cancer patients after surgery.

**Article Highlights**

***Research background***

Colorectal cancer (CRC) is the third most frequently diagnosed malignancy and one of the leading causes of cancer-related mortality worldwide. Therapy options for CRC have been developed rapidly in the past decade, but selecting optimal treatments for individuals remains a great challenge for clinicians due to the lack of effective markers.

***Research motivation***

Controversy exists regarding the insufficient prognostic value of preoperative serum CEA alone, which is a widely used biomarker in rectal cancer. Recent studies have found that local CEA may play a more important role in the prognosis of CRC than overall serum CEA. Some studies have tried to add another factor like tumor size to improve the prognostic value of biomarker, such as prostate specific antigen density and tumor-infiltrating CD8+ T-cell density. Here, we combined preoperative serum CEA and the maximum tumor diameter to correct the CEA level, which may better reflect the malignancy of rectal cancer and improve the risk stratification system.

***Research objectives***

We aimed to investigate the prognostic impact of the preoperative CEA/tumor size in patients with rectal cancer, which may influence the decision-making process for a specific treatment regimen and patient counselling.

***Research methods***

We retrospectively reviewed 696 stage I to III rectal cancer patients who underwent curative tumor resection from 2007 to 2012. These patients were randomly divided into two cohorts for cross-validation: training cohort and validation cohort. The training cohort was used to generate an optimal cutoff point and the validation cohort was used to further validate the model. Maximally selected rank statistics were used to identify the optimum cutoff for CEA/tumor size. The Kaplan-Meier method and log-rank test were used to plot the survival curve and to compare the survival data. Univariate and multivariate Cox regression analyses were used to determine the prognostic value of CEA/tumor size. The primary and secondary outcomes were overall survival (OS) and disease-free survival (DFS), respectively.

***Research results***

In all, 556 patients who satisfied both the inclusion and exclusion criteria were included and randomly divided into a training cohort (2/3 of 556, *n* = 371) and a validation cohort (1/3 of 556, *n* = 185). The cutoff was 2.429 ng/mL per cm. Comparison of the baseline data showed that high CEA/tumor size was correlated with older age, high TNM stage, presence of perineural invasion, high CEA, and high carbohydrate antigen 19-9 (CA 19-9). Kaplan-Meier curves showed a manifest reduction in 5-year OS (training cohort: 56.7% *vs* 81.1%, *P* < 0.001; validation cohort: 58.8% *vs* 85.6%, *P* <0.001) and DFS (training cohort: 52.5% *vs* 71.9%, *P* = 0.02; validation cohort: 50.3% *vs* 79.3%, *P* = 0.002) in the high CEA/tumor size group compared with the low CEA/tumor size group. Univariate and multivariate analyses identified CEA/tumor size as an independent prognostic factor for OS (training cohort: hazard ratio (HR) = 2.18 95% confidence interval (CI): 1.28-3.73, *P* = 0.004; validation cohort: HR = 4.83, 95%CI: 2.21-10.52, *P* < 0.001) as well as DFS (training cohort: HR = 1.47, 95% CI: 0.93-2.33, *P* = 0.096; validation cohort: HR: 2.61, 95%CI = 1.38-4.95, *P* = 0.003).

***Research conclusions***

This is the first study to evaluate the prognostic value of CEA/tumor size for stage I to III rectal cancer. We found that patients with high CEA/tumor size tended to have a worse outcome. Adjusting the confounding effect of tumor size can improve the prognostic value of CEA. Compared with CEA, another great advantage of CEA/tumor size is the ability to figure out those patients with higher CEA but relatively small tumor size. The results of our study suggest that these easily neglected tumors may represent higher malignancy and worse outcome, which may challenge the conventional risk stratification system. Since both CEA level and tumor size are routinely measured before surgery, the data of CEA/tumor size can be obtained by simple calculation. Therefore, CEA/tumor size can be easily applied in clinical practice.

***Research perspectives***

As a retrospective study, we were not able to obtain high-level clinical evidence, but the current retrospective study will provide an important basis for us to carry out a prospective study. A large-scale prospective study and longer follow-up time are needed in future study.

**REFERENCES**

1 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]

2 **Arnold M**, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 2017; **66**: 683-691 [PMID: 26818619 DOI: 10.1136/gutjnl-2015-310912]

3 **Deng Y**. Rectal Cancer in Asian vs. Western Countries: Why the Variation in Incidence? *Curr Treat Options Oncol* 2017; **18**: 64 [PMID: 28948490 DOI: 10.1007/s11864-017-0500-2]

4 **Dienstmann R**, Salazar R, Tabernero J. Personalizing colon cancer adjuvant therapy: selecting optimal treatments for individual patients. *J Clin Oncol* 2015; **33**: 1787-1796 [PMID: 25918287 DOI: 10.1200/JCO.2014.60.0213]

5 **Duffy MJ**, Lamerz R, Haglund C, Nicolini A, Kalousová M, Holubec L, Sturgeon C. Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update. *Int J Cancer* 2014; **134**: 2513-2522 [PMID: 23852704 DOI: 10.1002/ijc.28384]

6 **Stiksma J**, Grootendorst DC, van der Linden PW. CA 19-9 as a marker in addition to CEA to monitor colorectal cancer. *Clin Colorectal Cancer* 2014; **13**: 239-244 [PMID: 25442815 DOI: 10.1016/j.clcc.2014.09.004]

7 **Locker GY**, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, Somerfield MR, Hayes DF, Bast RC Jr; ASCO. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol* 2006; **24**: 5313-5327 [PMID: 17060676 DOI: 10.1200/JCO.2006.08.2644]

8 **Nicholson BD**, Shinkins B, Pathiraja I, Roberts NW, James TJ, Mallett S, Perera R, Primrose JN, Mant D. Blood CEA levels for detecting recurrent colorectal cancer. *Cochrane Database Syst Rev* 2015; : CD011134 [PMID: 26661580 DOI: 10.1002/14651858.CD011134.pub2]

9 **Chen L**, Jiang B, Wang Z, Liu M, Yang H, Xing J, Zhang C, Yao Z, Zhang N, Cui M, Su X. Combined preoperative CEA and CD44v6 improves prognostic value in patients with stage I and stage II colorectal cancer. *Clin Transl Oncol* 2014; **16**: 285-292 [PMID: 23860725 DOI: 10.1007/s12094-013-1069-2]

10 **Zhan X**, Sun X, Hong Y, Wang Y, Ding K. Combined Detection of Preoperative Neutrophil-to-Lymphocyte Ratio and CEA as an Independent Prognostic Factor in Nonmetastatic Patients Undergoing Colorectal Cancer Resection Is Superior to NLR or CEA Alone. *Biomed Res Int* 2017; **2017**: 3809464 [PMID: 28685148 DOI: 10.1155/2017/3809464]

11 **Kozman MA**, Fisher OM, Rebolledo BJ, Parikh R, Valle SJ, Arrowaili A, Alzahrani N, Liauw W, Morris DL. CEA to peritoneal carcinomatosis index (PCI) ratio is prognostic in patients with colorectal cancer peritoneal carcinomatosis undergoing cytoreduction surgery and intraperitoneal chemotherapy: A retrospective cohort study. *J Surg Oncol* 2018; **117**: 725-736 [PMID: 29266235 DOI: 10.1002/jso.24911]

12 **Tong G**, Xu W, Zhang G, Liu J, Zheng Z, Chen Y, Niu P, Xu X. The role of tissue and serum carcinoembryonic antigen in stages I to III of colorectal cancer-A retrospective cohort study. *Cancer Med* 2018; **7**: 5327-5338 [PMID: 30302946 DOI: 10.1002/cam4.1814]

13 **Huo YR**, Glenn D, Liauw W, Power M, Zhao J, Morris DL. Evaluation of carcinoembryonic antigen (CEA) density as a prognostic factor for percutaneous ablation of pulmonary colorectal metastases. *Eur Radiol* 2017; **27**: 128-137 [PMID: 27165139 DOI: 10.1007/s00330-016-4352-0]

14 **Aminsharifi A**, Howard L, Wu Y, De Hoedt A, Bailey C, Freedland SJ, Polascik TJ. Prostate Specific Antigen Density as a Predictor of Clinically Significant Prostate Cancer When the Prostate Specific Antigen is in the Diagnostic Gray Zone: Defining the Optimum Cutoff Point Stratified by Race and Body Mass Index. *J Urol* 2018; **200**: 758-766 [PMID: 29758219 DOI: 10.1016/j.juro.2018.05.016]

15 **Shimizu S**, Hiratsuka H, Koike K, Tsuchihashi K, Sonoda T, Ogi K, Miyakawa A, Kobayashi J, Kaneko T, Igarashi T, Hasegawa T, Miyazaki A. Tumor-infiltrating CD8<sup>+</sup> T-cell density is an independent prognostic marker for oral squamous cell carcinoma. *Cancer Med*2019; **8**: 80-93 [PMID: 30600646 DOI: 10.1002/cam4.1889]

16 **Eichelberger LE**, Koch MO, Eble JN, Ulbright TM, Juliar BE, Cheng L. Maximum tumor diameter is an independent predictor of prostate-specific antigen recurrence in prostate cancer. *Mod Pathol* 2005; **18**: 886-890 [PMID: 15803186 DOI: 10.1038/modpathol.3800405]

17 **Yoshimoto T**, Morine Y, Imura S, Ikemoto T, Iwahashi S, Saito YU, Yamada S, Ishikawa D, Teraoku H, Yoshikawa M, Higashijima J, Takasu C, Shimada M. Maximum Diameter and Number of Tumors as a New Prognostic Indicator of Colorectal Liver Metastases. *In Vivo*2017; **31**: 419-423 [PMID: 28438872 DOI: 10.21873/invivo.11076]

18 **Tanaka N**, Fujimoto K, Chihara Y, Torimoto M, Hirao Y, Konishi N, Saito I. Prostatic volume and volume-adjusted prostate-specific antigen as predictive parameters for prostate cancer patients with intermediate PSA levels. *Prostate Cancer Prostatic Dis* 2007; **10**: 274-278 [PMID: 17339878 DOI: 10.1038/sj.pcan.4500957]

19 **Peng Y**, Shen D, Liao S, Turkbey B, Rais-Bahrami S, Wood B, Karademir I, Antic T, Yousef A, Jiang Y, Pinto PA, Choyke PL, Oto A. MRI-based prostate volume-adjusted prostate-specific antigen in the diagnosis of prostate cancer. *J Magn Reson Imaging* 2015; **42**: 1733-1739 [PMID: 25946664 DOI: 10.1002/jmri.24944]

20 **Becerra AZ**, Probst CP, Tejani MA, Aquina CT, González MG, Hensley BJ, Noyes K, Monson JR, Fleming FJ. 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**: 1554-1561 [PMID: 26759308 DOI: 10.1245/s10434-015-5014-1]

21 **Huh JW**, Oh BR, Kim HR, Kim YJ. Preoperative carcinoembryonic antigen level as an independent prognostic factor in potentially curative colon cancer. *J Surg Oncol* 2010; **101**: 396-400 [PMID: 20119979 DOI: 10.1002/jso.21495]

22 **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**: 3087-3093 [PMID: 19629600 DOI: 10.1245/s10434-009-0625-z]

23 **Peng Y**, Wang L, Gu J. Elevated preoperative carcinoembryonic antigen (CEA) and Ki67 is predictor of decreased survival in IIA stage colon cancer. *World J Surg* 2013; **37**: 208-213 [PMID: 23052808 DOI: 10.1007/s00268-012-1814-7]

24 **Thirunavukarasu P**, Sukumar S, Sathaiah M, Mahan M, Pragatheeshwar KD, Pingpank JF, Zeh H 3rd, Bartels CJ, Lee KK, Bartlett DL. C-stage in colon cancer: implications of carcinoembryonic antigen biomarker in staging, prognosis, and management. *J Natl Cancer Inst* 2011; **103**: 689-697 [PMID: 21421861 DOI: 10.1093/jnci/djr078]

25 **Jun KH**, Jung H, Baek JM, Chin HM, Park WB. Does tumor size have an impact on gastric cancer? A single institute experience. *Langenbecks Arch Surg* 2009; **394**: 631-635 [PMID: 18791731 DOI: 10.1007/s00423-008-0417-0]

26 **Tayyab M**, Razack A, Sharma A, Gunn J, Hartley JE. Correlation of rectal tumor volumes with oncological outcomes for low rectal cancers: does tumor size matter? *Surg Today* 2015; **45**: 826-833 [PMID: 25377268 DOI: 10.1007/s00595-014-1068-0]

27 **Konishi T**, Shimada Y, Hsu M, Tufts L, Jimenez-Rodriguez R, Cercek A, Yaeger R, Saltz L, Smith JJ, Nash GM, Guillem JG, Paty PB, Garcia-Aguilar J, Gonen M, Weiser MR. Association of Preoperative and Postoperative Serum Carcinoembryonic Antigen and Colon Cancer Outcome. *JAMA Oncol* 2018; **4**: 309-315 [PMID: 29270608 DOI: 10.1001/jamaoncol.2017.4420]

28 **Austin PC**, Allignol A, Fine JP. The number of primary events per variable affects estimation of the subdistribution hazard competing risks model. *J Clin Epidemiol* 2017; **83**: 75-84 [PMID: 28088594 DOI: 10.1016/j.jclinepi.2016.11.017]

**P-Reviewer:** Alkan A, Ziogas DE

**S-Editor:** Ma RY **L-Editor:** Wang TQ **E-Editor:**

**Specialty type:** Gastroenterology and hepatology

**Country of origin:** China

**Peer-review report classification**

Grade A (Excellent): 0

Grade B (Very good): 0

Grade C (Good): C, C

Grade D (Fair): 0

Grade E (Poor): 0

**Table 1 Association of carcinoembryonic antigen/tumor size with baseline characteristics of rectal cancer patients *n* (%)**

|  |  |  |
| --- | --- | --- |
|  | **Training cohort (*n* = 371)** | **Validation cohort (*n* = 185)** |
|  | **Cases** | **Low** | **High** | ***P-*value** | **Cases** | **Low** | **High** | ***P-*value** |
| Age | 371 | 58 (21-89) | 65 (32-86) | <0.001a | 185 | 61 (25-87) | 57 (35-79) | 0.149 |
| Tumor size | 371 | 4.3 (0.8-13) | 4.3 (0.8-13.5) | 0.773 | 185 | 4.5 (1-13) | 4.3 (0.8-10) | 0.472 |
| Sex |  |  |  | 0.419 |  |  |  | 0.199 |
|  Male | 218 | 177 (58) | 41 (64) |  | 103 | 82 (53) | 21 (68) |  |
|  Female | 153 | 130 (42) | 23 (36) |  | 82 | 72 (47) | 10 (32) |  |
| TNM stage |  |  |  | 0.008a |  |  |  | 0.350 |
|  I | 104 | 96 (31) | 8 (12) |  | 48 | 43 (28) | 5 (16) |  |
|  II | 127 | 99 (32) | 28 (44) |  | 74 | 61 (40) | 13 (42) |  |
|  III | 140 | 112 (36) | 28 (44) |  | 63 | 50 (32) | 13 (42) |  |
| Differentiation  |  |  |  | 0.395 |  |  |  | 0.826 |
|  Poor | 60 | 51 (17) | 9 (14) |  | 24 | 19 (12) | 5 (16) |  |
|  Moderate | 209 | 176 (57) | 33 (52) |  | 102 | 85 (55) | 17 (55) |  |
|  High | 102 | 80 (26) | 22 (34) |  | 59 | 50 (32) | 9 (29) |  |
| Lymphovascular invasion |  |  |  | 0.697 |  |  |  | 0.683 |
|  Negative | 338 | 281 (92) | 57 (89) |  | 173 | 143 (93) | 30 (97) |  |
|  Positive | 33 | 26 (8) | 7 (11) |  | 12 | 11 (7) | 1 (3) |  |
| Perineural invasion |  |  |  | 0.039a |  |  |  | 0.073 |
|  Negative  | 340 | 286 (93) | 54 (84) |  | 172 | 146 (95) | 26 (84) |  |
|  Positive | 31 | 21 (7) | 10 (16) |  | 13 | 8 (5) | 5 (16) |  |
| CEA |  |  |  | <0.001a |  |  |  | <0.001a |
|  0-5 ng/mL | 263 | 262 (85) | 1 (2) |  | 127 | 126 (82) | 1 (3) |  |
|  >5 ng/mL | 108 | 45 (15) | 63 (98) |  | 58 | 28 (18) | 30 (97) |  |
| CA 19-9 |  |  |  | 0.006a |  |  |  | 0.027a |
|  0-37 ng/mL | 325 | 276 (90) | 49 (77) |  | 158 | 136 (88) | 22 (71) |  |
|  >37 ng/mL | 46 | 31 (10) | 15 (23) |  | 27 | 18 (12) | 9 (29) |  |

a*P* < 0.05; CEA: Carcinoembryonic antigen; CA 19-9: Carbohydrate antigen 19-9.

**Table 2 Univariate analysis of prognostic factors for overall survival**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Training cohort (*n* = 371)** | **Validation cohort (*n* = 185)** |
|  | **Hazard ratio** | **95%CI** | ***P*-value** | **Hazard ratio** | **95%CI** | ***P*-value** |
| Age | 1.02 | 1.00-1.04 | 0.024a | 1.03 | 1.00-1.06 | 0.070 |
| Tumor size | 1.04 | 0.92-1.18 | 0.506 | 1.14 | 0.95-1.37 | 0.164 |
| Sex (ref = male) | 1.43 | 0.88-2.31 | 0.145 | 0.85 | 0.41-1.77 | 0.665 |
| TNM1 (ref = stage I) | 1.74 | 1.26-2.41 | 0.001a | 1.93 | 1.15-3.22 | 0.012a |
| Differentiation1 (ref = poor) | 0.58 | 0.40-0.84 | 0.004a | 0.53 | 0.30-0.94 | 0.030a |
| Lymphovascular invasion (ref = negative) | 1.88 | 0.96-3.68 | 0.066 | 3.11 | 1.19-8.13 | 0.021a |
| Perineural invasion (ref = negative) | 1.03 | 0.41-2.56 | 0.954 | 1.29 | 0.31-5.46 | 0.729 |
| CEA (ref = CEA < 5) | 1.81 | 1.11-2.94 | 0.017a | 2.72 | 1.33-5.59 | 0.006a |
| CA 19-9 (ref = CA 19-9 < 37) | 1.88 | 1.04-3.39 | 0.036a | 2.14 | 0.92-4.99 | 0.078 |
| CEA/tumor size (ref = low) | 2.45 | 1.46-4.11 | 0.001a | 3.57 | 1.70-7.52 | 0.001a |

1These variables were treated as ordinal categorical data; a*P* < 0.05. CEA: Carcinoembryonic antigen; CA 19-9: Carbohydrate antigen 19-9; CI: Confidence interval; ref: Reference.

**Table 3 Univariate analysis of prognostic factors for disease-free survival**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Training cohort (*n* = 371)** | **Validation cohort (*n* = 185)** |
|  | **Hazard ratio** | **95%CI** | ***P*-value** | **Hazard ratio** | **95%CI** | ***P*-value** |
| Age | 1 | 0.99-1.02 | 0.572 | 1.02 | 0.99-1.04 | 0.173 |
| Tumor size | 1.01 | 0.91-1.12 | 0.828 | 1.12 | 0.97-1.30 | 0.128 |
| Sex (ref = male) | 1.26 | 0.85-1.87 | 0.247 | 0.75 | 0.41-1.37 | 0.353 |
| TNM1 (ref = stage I) | 1.9 | 1.45-2.50 | <0.001a | 1.6 | 1.07-2.39 | 0.023a |
| Differentiation1 (ref = poor) | 0.78 | 0.58-1.06 | 0.113 | 0.6 | 0.38-0.95 | 0.031a |
| Lymphovascular invasion(ref = negative) | 2.44 | 1.45-4.12 | 0.001a | 2.63 | 1.11-6.22 | 0.028a |
| Perineural invasion(ref = negative) | 2.17 | 1.23-3.82 | 0.008a | 1.98 | 0.78-5.03 | 0.151 |
| CEA (ref = CEA < 5) | 1.55 | 1.03-2.32 | 0.034a | 1.9 | 1.05-3.41 | 0.033a |
| CA 19-9 (ref = CA 19-9 < 37) | 1.43 | 0.85-2.42 | 0.177 | 1.96 | 0.97-3.96 | 0.061 |
| CEA/tumor size (ref = low) | 1.72 | 1.10-2.71 | 0.018a | 2.58 | 1.37-4.85 | 0.003a |

1These variables were treated as ordinal categorical data; a*P* < 0.05. CEA: Carcinoembryonic antigen; CA 19-9: Carbohydrate antigen 19-9; CI: Confidence interval; ref: Reference.

**Table 4 Multivariate analysis of prognostic factors for overall survival and disease-free survival**

|  |  |  |
| --- | --- | --- |
|  | **Training cohort (*n* = 371)** | **Validation cohort (*n* = 185)** |
| **OS** | **Hazard radio** | **95%CI** | ***P*-value** | **Hazard radio** | **95%CI** | ***P*-value** |
| Age  | 1.02 | 1.00-1.04 | 0.023a | 1.05 | 1.02-1.09 | 0.003a |
| TNM1 (ref = stage I) | 1.47 | 1.04-2.07 | 0.031a | 1.84 | 1.04-3.24 | 0.035a |
| Differentiation1 (ref = poor) | 0.57 | 0.39-0.85 | 0.006a | 0.50 | 0.28-0.90 | 0.021a |
| CEA/tumor size (ref = low) | 2.18 | 1.28-3.73 | 0.004a | 4.83 | 2.21-10.52 | <0.001a |
| DFS |  |  |  |  |  |  |
| TNM1 (ref = stage I) | 1.75 | 1.32-2.32 | <0.001a | 1.43 | 0.94-2.17 | 0.091 |
| Lymphovascular invasion(ref = negative) | 1.85 | 1.08-3.16 | 0.024a | 2.45 | 1.00-6.03 | 0.05 |
| CEA/tumor size (ref = low) | 1.47 | 0.93-2.33 | 0.096 | 2.61 | 1.38-4.95 | 0.003a |

1These variables were treated as ordinal categorical data; a*P* < 0.05. CEA: Carcinoembryonic antigen; CI: Confidence interval; ref: Reference; OS: Overall survival; DFS: Disease-free survival.



**Figure 1 Flowchart of patient selection.**



**Figure 2 Maximally selected rank statistics for carcinoembryonic antigen/tumor size.** Maximally selected rank statistics were used to identify the optimal discriminator value for the carcinoembryonic antigen/tumor size, which was conducted in the training cohort. For every potential cutoff point, the absolute value of the standardized log-rank statistic was computed. The cutoff point that provided the best separation of the survival outcome into two groups, where the standardized statistics reached their maximum, was selected as the cutoff point. CEA: Carcinoembryonic antigen.



**Figure 3 Kaplan-Meier survival curves and risk tables for overall survival.** A: Kaplan-Meier survival curves and risk table for overall survival in the training cohort. The 5-year overall survival (OS) of the high and low carcinoembryonic antigen (CEA)/tumor size groups were 56.7% and 81.1% (*P* < 0.001), respectively. B: Kaplan-Meier survival curves and risk table for overall survival in the validation cohort. The 5-year OS of the high and low CEA/tumor size groups were 58.8% and 85.6% (*P* < 0.001), respectively. The log-rank test was used to calculate the *P*-value. OS: Overall survival; CEA: Carcinoembryonic antigen.



**Figure 4** **Kaplan-Meier survival curves and risk tables for disease-free survival.** A: Kaplan-Meier survival curves and risk table for disease-free survival (DFS) in the training cohort. The 5-year DFS of the high and low CEA/tumor size groups were 52.5% and 71.9% (*P* = 0.02), respectively. B: Kaplan-Meier survival curves and risk table for DFS in the validation cohort. The 5-year DFS of the high and low CEA/tumor size groups were 50.3% *vs* 79.3% (*P* = 0.002), respectively. The log-rank test was used to calculate the *P*-value. DFS: Disease-free survival; CEA: Carcinoembryonic antigen.



**Figure 5 Proportional hazards assumption test for overall survival by plotting the Schoenfeld residuals against time in the training cohort (A, C, E, and G) and the validation cohort (B, D, F, and H).** The X-axis represents the survival time, while the Beta values referring to age, TNM stage, differentiation, and carcinoembryonic antigen/tumor size are shown on the Y-axis. The constant mean of residuals across time confirms that the proportional hazard assumption holds for these covariate with all of the *P*-values > 0.05. CEA: Carcinoembryonic antigen.



**Figure 6 Proportional hazards assumption test for disease-free survival by plotting the Schoenfeld residuals against time in the training cohort (A, C, and E) and the validation cohort (B, D, and F).** The X-axis represents the survival time, while the Beta values referring to TNM stage, lymphovascular invasion, and carcinoembryonic antigen/tumor size are shown on the Y-axis. The constant mean of residuals across time confirms that the proportional hazard assumption holds for these covariate with all of the *P*-values > 0.05. CEA: Carcinoembryonic antigen.