Focus on patients with early esophageal cancer— a prognostic nomogram

Zhiyuan Cheng1, Zifan Zhang1, Han Lin2, Qianqian Meng2, Lei Xin2, Tianjiao Wang2, Wei Wang2, Luowei Wang2

1 College of Basic Medical Sciences, Naval Medical University, Shanghai, China; 2 Department of Gastroenterology, Changhai Hospital, Naval Medical University, Shanghai, China

Contributions: (I) Conception and design: L Wang, Z Cheng; (II) Administrative support: L Wang; (III) Provision of study materials or patients: H Lin, L Xin; (IV) Collection and assembly of data: T Wang, W Wang; (V) Data analysis and interpretation: Z Zhang, Q Meng, L Xin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

# These authors contributed equally to this work.

Correspondence to: Luowei Wang, MD, PhD. Professor of Gastroenterology, Department of Gastroenterology, Changhai Hospital, Naval Medical University, 168 Changhai Road, Yangpu District, Shanghai 200433, China. Email: wangluoweimd@126.com.

Background: Esophageal cancer is a common cancer of the digestive system, with high morbidity and poor prognosis. However, while the prognosis of early esophageal cancer is relatively good, there is no effective model to accurately predict the prognosis of early esophageal cancer. The Aims of this study are to explore risk factors for the prognosis of early esophageal cancer and to establish a prediction nomogram for patients.

Methods: Surveillance, Epidemiology and End Results (SEER) Stat 8.3.5 was used to collect 2,351 cases of early esophageal cancer from 2004 to 2015 in the SEER database. Early esophageal cancer is defined as a lesion that is confined to the lamina propria and the muscularis mucosa. Prognostic factors were analyzed with the log-rank method and a Cox proportional hazard model by SPSS (v25.0). Independent prognostic factors were used to construct a nomogram with a Cox proportional hazard model. The C-index was used to evaluate the prediction effect of the nomogram. The internal validity of the nomogram was tested by discrimination and calibration using a bootstrap method with 1,000 resamplings.

Results: The median survival time was 30 months, and the 1-, 3-, and 5-year survival rates were 65.2%, 46.8%, and 41.6%, respectively. The male to female ratio was 3:1, and 85.33% of all patients were white. Univariate analysis showed that risk factors affecting patient prognosis included age ($\chi^2=430.631$, $P<0.001$), sex ($\chi^2=48.1$, $P<0.001$), marital status ($\chi^2=107.597$, $P<0.001$), race ($\chi^2=58.928$, $P<0.001$), primary site ($\chi^2=98.675$, $P<0.001$), tumor grade ($\chi^2=116.421$, $P<0.001$), surgery ($\chi^2=1259.33$, $P<0.001$) and histologic type ($\chi^2=231.062$, $P<0.001$). Using multivariate analysis, we found that age (HR=1.787, 95% CI: 1.58–2.03), marital status (HR=0.774, 95% CI: 0.69–0.87), tumor grade (HR=1.241, 95% CI: 1.14–1.35), and surgery (HR=0.356, 95% CI: 0.33–0.39) were independent prognostic factors for patients with early esophageal cancer. We constructed the nomogram with the above independent factors, and the C-index value was 0.788.

Conclusions: This study obtained the latest epidemiological information on early esophageal cancer and determined that age, marital status, tumor grade and surgery were independent prognostic factors for early esophageal cancer. The nomogram developed with these factors could provide good prognosis prediction.

Keywords: Esophageal neoplasms; nomograms; prognosis analysis; Surveillance, Epidemiology and End Results (SEER) program

doi: 10.21037/tcr-19-1645
View this article at: http://dx.doi.org/10.21037/tcr-19-1645

ORCID: Zhiyuan Cheng, 0000-0003-1150-5816; Zifan Zhang, 0000-0002-9253-5111; Luowei Wang, 0000-0002-6647-786X.
**Introduction**

Early esophageal cancer is defined as a lesion that is confined to the mucosa (the lamina propria and the muscularis mucosa), rather than a lesion that invades the muscularis propria through the submucosa (1). The current research focuses on the treatment of early esophageal cancer (2,3), especially endoscopic treatment (4-6); thus far, no systematic analysis of prognostic factors for patients with early esophageal cancer has been conducted. Compared with other esophageal cancer patients, patients with early esophageal cancer have a better prognosis, but clinically, the prediction of these early clinical staged cancers is often too optimistic (7). To further explore early esophageal cancer and make systematic and accurate predictions, this study extracted 2,351 cases of early esophageal cancer from the Surveillance, Epidemiology and End Results (SEER) database to analyze the prognostic factors of these patients and provide guidance for clinical treatment. In addition, this study excluded confounding interference, developed a predictive model based on various factors to predict the survival of patients with early esophageal cancer, and visually showed the intensity of each independent factor and level of intervention for prognosis.

**Methods**

**Source**

Patients diagnosed with early esophageal cancer from 2004 to 2015 were identified in the SEER database using SEER Stat software (National Cancer Institute, version 8.3.5). Since any information in the SEER database does not require the patient’s explicit consent, it is not subject to the ethical approval requirements of the institutional review board.

**Patient screening**

Inclusion criteria: (I) Patients aged 18 years or older diagnosed with early esophageal cancer (II) patients diagnosed between 2004–2015; (III) patients diagnosed with pathological results; (IV) patients with complete follow-up information including age, sex, marital status, race, primary site, tumor grade, therapy and histologic type; (V) Patients died of early esophageal cancer rather than other causes.

Exclusion criteria: (I) patients with incomplete basic information on age, sex, race; (II) cases with unknown survival time; (III) patients for whom esophageal cancer is not the first primary cancer. (V) Patients received chemotherapy or radiotherapy at any time.

A total of 2,351 patients were screened and collected. The seventh edition of the tumor-node-metastasis (TNM) staging system from the American Joint Committee on Cancer (AJCC) was used to staging patients.

**Statistical analyses**

SEER Stat 8.3.5 software was used to collect 2,351 cases of early esophageal cancer from 2004 to 2015 in The SEER database. For continuous variables, we used X-tile software (Yale University, Version 3.6.1) to determine the best truncation value of the age, and it was 80. Prognostic factors of early esophageal cancer were analyzed by SPSS, version 25.0 (SPSS Inc., Chicago, IL) and the log-rank method. Introducing meaningful variables of single factor analysis into the Cox proportional hazard model and multivariate analysis, the independent influencing factors of prognosis were obtained, and $P<0.05$ was considered statistically significant. Independent prognostic factors were included in the accelerated failure-time model to construct a nomogram. The C-index was used to evaluate the prediction effect of the nomograms separately, which indicates prediction accuracy of the nomogram. The internal validity of the nomogram was tested by discrimination and calibration. Bootstrap analyses with 1,000 resamplings were applied.

**Results**

A total of 2,351 cases of early esophageal cancer were included in the study. The median survival time was 30 months, the 1-year survival rate was 65.2%, the 3-year survival rate was 46.8%, and the 5-year survival rate was 41.6%. The male to female ratio was 3:1; 85.33% of all patients were white, and their survival prognosis was significantly better than other races ($P<0.001$). Overall, 67.63% of the lesions were located in the lower third of the esophagus; 47.04% of the tumors had moderate malignancy (grade II), but 39.52% of the tumors had higher malignancy (grade III or IV). Squamous cell carcinomas (SCCs) accounted for 30.41%, and adenocarcinomas accounted for 63.5% (Table 1).

A total of 37.81% of patients underwent surgery, and 12.55% chose endoscopic treatment. Surprisingly, 49.64% of patients did not undergo surgery or endoscopic treatment, and the prognosis of this subgroup of early esophageal cancer patients was extremely poor compared with the treated patients ($P<0.001$).
<table>
<thead>
<tr>
<th>Clinicopathologic parameters</th>
<th>Number of cases</th>
<th>Average survival (months)</th>
<th>95% CI (months)</th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–79</td>
<td>1,895</td>
<td>80.60</td>
<td>75.59</td>
<td>72.39–78.79</td>
<td></td>
</tr>
<tr>
<td>≥80</td>
<td>456</td>
<td>19.40</td>
<td>20.63</td>
<td>16.70–24.56</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2,351</td>
<td>100.00</td>
<td>65.04</td>
<td>62.19–67.88</td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1,789</td>
<td>76.10</td>
<td>69.90</td>
<td>66.61–73.18</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>562</td>
<td>23.90</td>
<td>48.77</td>
<td>43.38–54.16</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2,351</td>
<td>100.00</td>
<td>65.04</td>
<td>62.19–67.88</td>
<td></td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmarried</td>
<td>1,008</td>
<td>42.88</td>
<td>50.21</td>
<td>46.11–54.31</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>1,343</td>
<td>57.12</td>
<td>75.88</td>
<td>72.09–79.66</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2,351</td>
<td>100.00</td>
<td>65.04</td>
<td>62.19–67.88</td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>2,006</td>
<td>85.33</td>
<td>68.59</td>
<td>65.49–71.70</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>233</td>
<td>9.91</td>
<td>36.07</td>
<td>29.12–43.02</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>112</td>
<td>4.76</td>
<td>59.27</td>
<td>46.62–71.92</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2,351</td>
<td>100.00</td>
<td>65.04</td>
<td>62.19–67.88</td>
<td></td>
</tr>
<tr>
<td><strong>Primary site</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cervical esophagus</td>
<td>40</td>
<td>1.70</td>
<td>45.12</td>
<td>28.67–61.58</td>
<td></td>
</tr>
<tr>
<td>Thoracic esophagus</td>
<td>102</td>
<td>4.34</td>
<td>46.17</td>
<td>37.00–55.34</td>
<td></td>
</tr>
<tr>
<td>Abdominal esophagus</td>
<td>28</td>
<td>1.19</td>
<td>42.87</td>
<td>26.38–59.36</td>
<td></td>
</tr>
<tr>
<td>Upper third of esophagus</td>
<td>142</td>
<td>6.04</td>
<td>41.05</td>
<td>31.30–50.79</td>
<td></td>
</tr>
<tr>
<td>Middle third of esophagus</td>
<td>449</td>
<td>19.10</td>
<td>45.09</td>
<td>39.26–50.92</td>
<td></td>
</tr>
<tr>
<td>Lower third of esophagus</td>
<td>1,590</td>
<td>67.63</td>
<td>74.12</td>
<td>70.61–77.64</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2,351</td>
<td>100.00</td>
<td>65.04</td>
<td>62.19–67.88</td>
<td></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>316</td>
<td>13.44</td>
<td>92.06</td>
<td>84.01–100.12</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>1,106</td>
<td>47.04</td>
<td>70.94</td>
<td>66.75–75.13</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>887</td>
<td>37.73</td>
<td>49.00</td>
<td>44.76–53.25</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>42</td>
<td>1.79</td>
<td>56.42</td>
<td>38.85–74.00</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2,351</td>
<td>100.00</td>
<td>65.04</td>
<td>62.19–67.88</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 (continued)
Univariate analysis of prognostic factors

Data analysis showed that there were eight factors included in the study had a significant impact on the prognosis of patients, including age ($\chi^2=430.631, P<0.001$), sex ($\chi^2=48.1, P<0.001$), marital status ($\chi^2=107.597, P<0.001$), race ($\chi^2=58.928, P<0.001$), primary site ($\chi^2=98.675, P<0.001$), tumor grade ($\chi^2=116.421, P<0.001$), therapy ($\chi^2=1,259.33, P<0.001$) and histologic type ($\chi^2=231.062, P<0.001$).

It should be noted that the survival time for those with lesions in the lower third of the esophagus was significantly higher than for those with lesions of the middle and upper part of the esophagus (mean survival time: 74.12 vs. 45.09 vs. 41.05 months, respectively); the survival time of patients with early esophageal cancer in the abdominal segment was lower than for those with lesions in the cervical or thoracic segment (average survival time: 42.87 vs. 45.12 vs. 46.17 months, respectively). In addition, patients who underwent endoscopic treatment had a slightly longer survival than those who underwent surgery (107.85–123.65 vs. 103.54–111.54 months), but this difference was not significant (P=0.113) (Figure 1). However, the survival times of none surgical groups was obviously shorter (19.23–23.78 months), showing the importance of early surgical intervention.

Multivariate analysis of prognostic factors and nomogram

Age, sex, marital status, race, primary site, tumor grade, therapy and histologic type were included in the Cox proportional hazard model. The results showed that only age (HR=1.787, 95% CI: 1.58–2.03), marital status (HR=0.774, 95% CI: 0.69–0.87), tumor grade (HR=1.241, 95% CI: 1.14–1.35) and surgery (HR=0.356, 95% CI: 0.33–0.39) were independent prognostic factors for patients with early esophageal cancer (Table 2).

The nomogram was constructed with the independent prognostic factors and is shown in Figure 2. Supposing an early esophageal cancer patient was younger than 80 years old (points =0), unmarried (points =0), and the degree of cancer differentiation was grade II (points =12); if he or she underwent surgical treatment (points =90), the total points was 102. Comparing the bottom three lines of Figure 2, it is intuitive to see that the probability of this patient surviving for 1-, 3-, and 5-year is 85%, 62%, and 50%, respectively (Figure 2).

Next, we evaluated the nomogram and verified it within the group. The results show that the C-index of the model was 0.788 (95% CI: 0.776–0.800), indicating that the accuracy of the prediction model was 78.8%. The calibration curve in Figure 3 shows a good fit of the model (Figure 3).

Discussion

General condition of early esophageal cancer

It was estimated that there were 17,290 new cases of esophageal cancer and 15,850 deaths from the disease in the United States in 2018 (8), with a 5-year survival rate of 19.2%. This study found that the 5-year survival rate...
Figure 1 KM curves of age (A), sex (B), marital status (C), tumor grade (D), race (E), primary site (F), surgery (G), and histologic type (H), respectively.
of early esophageal cancer was 41.2%, which was superior to other esophageal cancers, but it was still not optimistic. In addition, the male incidence rate was higher than that of females (3:1), which was consistent with general clinical observations and similar current research findings (9).

Nearly half of the patients did not undergo surgery or endoscopic treatment

Some studies have noted that endoscopic treatment could be used as an alternative to early esophageal cancer esophagectomy (5,10). The complication rate was lower; the recovery time was shorter; and the survival rate was similar (11). However, 73.9% of patients chose surgical resection instead of endoscopic treatment. It is noteworthy that 49.6% of patients with early esophageal cancer did not undergo surgery or endoscopic treatment, and their prognosis was extremely poor (survival time: 19.23–23.78 months); thus, it might be more meaningful to treat patients than to discuss the advantages and disadvantages of treatment options. This phenomenon is quite common in clinical settings. Endoscopic Ultrasound (EUS) and Computed Tomography - Positron Emission Tomography (CT-PET) technology assist in early detection of esophageal cancer and have improved patient survival (12). Patients who underwent EUS were more likely to undergo esophagectomy (P=0.01) and adjuvant therapy (P=0.008); however, only 10.7% of patients in this analysis underwent EUS. The high costs of imaging and surgery, as well as low acceptance, were the main causes of this phenomenon (13).

SCC and adenocarcinoma

SCC is the major histological type of esophageal cancer worldwide, accounting for 90% of cases. However, we found that in the early stage of esophageal cancer in the United
States from 2004 to 2015, SCC accounted for only 30.41% of cases, while adenocarcinoma accounted for 63.5%. Some epidemiological surveys showed that the incidence of esophageal cancer decreased (14-16) in North America and Europe, while esophageal adenocarcinoma increased in Western countries (17). A study (18) based on the SEER database also showed that, between 1973 and 2009, adenocarcinoma (53.9%) was significantly more prevalent than histological SCC (33.0%). This trend might explain the current phenomenon of adenocarcinoma. In terms of clinical experience, esophageal SCC was more malignant than adenocarcinoma. Our study also showed that patients with early esophageal adenocarcinoma had a better prognosis than patients with SCC. A study (19) by the American Cancer Genome Atlas Research Program (TCGA) showed that, from a molecular point of view, esophageal SCC and esophageal adenocarcinoma are completely different diseases.

**Prognostic risk factor**

After comparing the studies of Kim *et al.* (9) and Zhang *et al.* (20), we found that, in addition to treatment, the

---

**Figure 3** Calibration curves of the nomogram. (A) 1-year calibration curves; (B) 3-year calibration curves; (C) 5-year calibration curves. The x-axis shows the nomogram predicted probability, and the y-axis gives the actual survival as estimated by the Kaplan-Meier method.
prognostic factors for early esophageal cancer were the same as those for esophageal cancer overall. Zhang et al. compared the difference between radiotherapy and surgery, and Kim et al. did not include the treatment of esophageal cancer; however, neither integrated the prognostic factors to construct a predictive model.

Most of the early esophageal cancer cases originated in the lower third of the esophagus, and the survival time of these patients was significantly higher than for other sites (P<0.001). This was mainly because the SCCs, with higher malignancy, occurred more often in the upper and middle third of esophagus (18); surgery is difficult at this site and it is easy to cause multiple important organ injuries, which greatly reduced the survival time of patients. Different prognoses might be related to the difference between the cancers. Upper esophageal cancer was closer to cancer of the head and neck, while the tumors in the lower part of the esophagus were virtually indistinguishable from a subtype of gastric cancer (19).

Marital status was found to be an independent prognostic factor for patients with early esophageal cancer. The prognosis of married patients was better than that of unmarried patients. Several studies have also shown that the survival rate of gastrointestinal tumors in unmarried patients was lower than that in married patients (21-23). In general, the impact of marital status on patients is achieved through changes of sex hormones (24-26), which has often been shown in endocrine diseases, such as thyroid cancer (27), breast cancer, and melanoma (28). A review mentioned that estrogen level was a contributing factor in the pathogenesis of esophageal SCC, and hormone replacement therapy might help reduce the risk of esophageal SCC (29). A prospective cohort study (30) showed that moderate expression of estrogen receptor β (ERβ) was associated with increased survival in patients with gastroesophageal junction cancer. No evidence of estrogen receptor α (ERα) or androgen receptor (AR) expression was found in esophageal adenocarcinoma.

**Nomogram**

There was no nomogram for predicting the prognosis of patients with early esophageal cancer. Therefore, this study considered the independent factors and used the results of the Cox proportional hazard model to establish a nomogram, which visually showed the effect of each independent factor on prognosis. The nomogram was first invented by French engineer Philbert Maurice d’Ocagne in 1884. It was used to provide engineers with fast graphical calculations of complex formulas. Due to its simple and intuitive nature, the nomogram has gradually been introduced in various fields. In medicine, it was commonly used to predict the occurrence of disease (31) and the prognosis of patients (32). Nomograms are widely used in esophageal cancer-related applications (33-35), mostly for patients with certain types of esophageal cancer, such as trigeminy therapeutic regimen for esophageal adenocarcinoma (36), neoadjuvant chemotherapy (33), and recurrence after surgery (35). In our study, the evaluation results showed that the accuracy of the prediction model reached 78.8%, which indicated a surprisingly good predictive capacity compared with similar studies.

Stratified treatment and individualized treatment are gradually becoming the frontier medical concept. A multicenter trial classified breast cancer patients into low-, medium-, and high-risk groups and found no significant difference in the efficacy of postoperative endocrine therapy and endocrine therapy combined with chemotherapy in the middle-risk group, suggesting that patients in the moderate-risk group could be exempted from chemotherapy after surgery (37). Goense used a nomogram to stratify preoperative risk in patients with esophageal adenocarcinoma to assess the overall survival (OS) benefit of esophagectomy after radiotherapy and chemotherapy (CRT) (36). This suggests that we can stratify patients with early esophageal cancer in a future in-depth study and develop a targeted treatment plan to reduce ineffective treatment while improving the patient’s prognosis.

**Insufficient research**

Although this study constructed a reliable nomogram for prediction, there were several deficiencies that should be considered when interpreting our results. First, we excluded some patients because of the lack of data on important variables, including age, primary site, tumor grade, treatment and marital status. This could lead to deviations in the nomogram. Second, the SEER database did not document genetic factors and some interventions, including family history, genomic status, weight and smoking, which might influence the predictive power of the nomogram. Third, we did not have access to esophageal cancer data from other medical centers, so the predictive model could not be further evaluated. In addition, the nomogram is subject to the limitations of retrospective data collection and must be validated by prospective cohort studies before it can be used in clinical practice.
Acknowledgments

Funding: This study is funded by “Research on Intelligent Aided Diagnosis System of Esophageal Tumor Based on Big Data of Upper Digestive Endoscopy Image” (grant number: 18411952600).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr-19-1645). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This article does not contain any study of human participants or animals performed by the author. As this study is based on a publicly available database without identifying patient information, informed consent was not needed.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References
