**Name of Journal:** *World Journal of Gastrointestinal Surgery*

**Manuscript NO:** 89360

**Manuscript Type:** ORIGINAL ARTICLE

***Retrospective Cohort Study***

**Risk stratification in gastric cancer lung metastasis: Utilizing an overall survival nomogram and comparing it with previous staging**

Chen ZR *et al*. Prognostic model innovation in gastric cancer

Zhi-Ren Chen, Mei-Fang Yang, Zhi-Yuan Xie, Pei-An Wang, Liang Zhang, Ze-Hua Huang, Yao Luo

**Zhi-Ren Chen,** Department of Science and Education, Xuzhou Medical University, Xuzhou 221000, Jiangsu Province, China

**Zhi-Yuan Xie,** Department of Neurology, Clinical Laboratory, Gastrointestinal Surgery, Central Hospital of Xuzhou, Central Hospital of Xuzhou, Xuzhou 221000, Jiangsu Province, China

**Mei-Fang Yang,** Department of Neurology, Xuzhou Medical University, Xuzhou 221000, Jiangsu Province, China

**Pei-An Wang,** Department of Public Health, Xuzhou Central Hospital, Xuzhou 221000, Jiangsu Province, China

**Liang Zhang,** Department of Gastroenterology, Xuzhou Centre Hospital, Xuzhou 221000, Jiangsu Province, China

**Ze-Hua Huang, Yao Luo,** Department of Public Health, Xuzhou Medical University, Xuzhou 221000, Jiangsu Province, China

**Co-first authors:** Zhi-Ren Chen and Mei-Fang Yang.

**Author contributions:** Chen RZ and Yang MF contributed equally to this work; Chen RZ wrote a manuscript, Xie ZY conceptualized and designed the study, and Yang MF provided the study materials; Wang PA provided administrative support; Zhang L collected and assembled the data; Luo Y performed data analysis and interpretation; All authors participated in manuscript writing and approved the final manuscript.

**Supported by** Peng-Cheng Talent-Medical Young Reserve Talent Training Program, No. XWRCHT20220002; Xuzhou City Health and Health Commission Technology Project Contract, No. XWKYHT20230081; and Key Research and Development Plan Project of Xuzhou City, No. KC22179.

**Corresponding author: Pei-An Wang, MD, PhD, Professor, Researcher,** Department of Public Health, Xuzhou Central Hospital, No. 15 Building Community, Xuzhou 221000, Jiangsu Province, China. 302303121267@stu.xzhmu.edu.cn

**Received:** November 3, 2023

**Revised:** December 16, 2023

**Accepted:** January 19, 2024

**Published online:**

**Abstract**

BACKGROUND

Gastric cancer (GC) is prevalent and aggressive, especially when patients have distant lung metastases, which often places patients into advanced stages. By identifying prognostic variables for lung metastasis in GC patients, it may be possible to construct a good prediction model for both overall survival (OS) and the cumulative incidence prediction (CIP) plot of the tumour.

AIM

To investigate the predictors of GC with lung metastasis (GCLM) to produce nomograms for OS and generate CIP by using cancer-specific survival (CSS) data.

METHODS

Data from January 2000 to December 2020 involving 1652 patients with GCLM were obtained from the Surveillance, epidemiology, and end results program database. The major observational endpoint was OS; hence, patients were separated into training and validation groups. Correlation analysis determined various connections. Univariate and multivariate Cox analyses validated the independent predictive factors. Nomogram distinction and calibration were performed with the time-dependent area under the curve (AUC) and calibration curves. To evaluate the accuracy and clinical usefulness of the nomograms, decision curve analysis (DCA) was performed. The clinical utility of the novel prognostic model was compared to that of the 7th edition of the American Joint Committee on Cancer (AJCC) staging system by utilizing Net Reclassification Improvement (NRI) and Integrated Discrimination Improvement (IDI). Finally, the OS prognostic model and Cox-AJCC risk stratification model modified for the AJCC system were compared.

RESULTS

For the purpose of creating the OS nomogram, a CIP plot based on CSS was generated. Cox multivariate regression analysis identified eleven significant prognostic factors (*P* < 0.05) related to liver metastasis, bone metastasis, primary site, surgery, regional surgery, treatment sequence, chemotherapy, radiotherapy, positive lymph node count, N staging, and time from diagnosis to treatment. It was clear from the DCA (net benefit > 0), time-dependent ROC curve (training/validation set AUC > 0.7), and calibration curve (reliability slope closer to 45 degrees) results that the OS nomogram demonstrated a high level of predictive efficiency. The OS prediction model (New Model AUC = 0.83) also performed much better than the old Cox-AJCC model (AUC difference between the new model and the old model greater than 0) in terms of risk stratification (*P* < 0.0001) and verification using the IDI and NRI.

CONCLUSION

The OS nomogram for GCLM successfully predicts 1- and 3-year OS. Moreover, this approach can help to appropriately classify patients into high-risk and low-risk groups, thereby guiding treatment.

**Key Words:** Gastric cancer; Lung metastasis; Nomograms; Surveillance; Epidemiology; Surveillance epidemiology and end results program database; Overall survival; Prognosis

Chen RZ, Xie ZY, Yang MF, Wang PA, Zhang L, Huang ZH, Luo Y. Risk stratification in gastric cancer lung metastasis: Utilizing an overall survival nomogram and comparing it with previous staging. *World J Gastrointest Surg* 2024; In press

**Core Tip:** From the viewpoints of overall survival and cancer-specific survival, this study investigated the survival probability based on independent prognostic indicators in gastric cancer with lung metastasis (GCLM) patients and generated a nomogram; the cumulative incidence of disease initiation was predicted. A risk score is assigned to each patient, and vital assistance is offered for individualized treatment plans in GCLM. Moreover, this groundbreaking study provides a model for the prognosis and prevention of various malignancies.

**INTRODUCTION**

As the third most common cause of cancer-related deaths worldwide, gastric cancer (GC) continues to pose a serious threat to public health[1]. Clinical care for stomach cancer is complicated by its metastatic course. Although the liver has been the subject of most studies on GC metastases[2], lung metastases are also gradually gaining increased attention. The clinical circumstances involving the intersection of GC and lung metastasis (GCLM) are multifaceted and require accurate methods for diagnosis, treatment, and prognosis. The incidence and number of cases of lung metastasis from stomach cancer have increased over the past several decades. Clinical studies have shown that regional differences in the incidence rates of lung metastasis from stomach cancer have an effect on overall survival (OS)[3].

Early detection is crucial for improving the prognoses of GCLM patients. Imaging technology advancements, including computed tomography (CT), magnetic resonance imaging, and positron emission tomography, have improved the detection of metastatic lesions[4]. Furthermore, the introduction of liquid biopsy technologies and the discovery of novel biomarkers have ushered in a new age in diagnosis, thus allowing for noninvasive surveillance and early diagnosis of metastasis[5].

Treatment strategies for GCLM have developed over time and include a variety of surgical, chemotherapeutic, and targeted therapeutic methods. The incorporation of immunotherapy and diverse treatment techniques heralds the start of a new era in customized medicine[6]. Even with these improvements, many GC patients still experience late recurrence, and it is difficult to predict OS and disease cumulative incidence[7]. However, the prognosis of GCLM patients remains unknown, thus emphasizing the importance of continued research into novel treatment options and prognostic models. Multiple factors influence the prognosis of GCLM, including the severity of the disease, the occurrence of metastases elsewhere in the body, and the patient's overall health. An understanding of clinical prediction markers and the use of prognostic models are critical for directing treatment decisions and giving patients a realistic picture of their illness trajectory[8].

Due to their simplicity and accessibility of testing, nomograms have become standard tools for prognostic prediction in patients with various malignancies. These findings can assist clinicians in making more informed treatment decisions and provide personalized survival predictions for patients[9,10]. This study created a large GCLM dataset based on the Surveillance, epidemiology, and end results program (SEER) database, with the goal of developing an accurate predictive nomogram for GCLM patients.

**MATERIALS AND METHODS**

***Data sources and inclusion criteria***

This study used SEER-stat software version 8.4.2 to obtain the data. The SEER database, which is funded by the National Cancer Institute, is a decentralized registry spanning multiple centres and populations. It operates independently of medical ethics review processes and does not mandate the acquisition of informed consent[11]. This study used data from the 2010–2020 SEER Research Plus database, which met strict inclusion and exclusion criteria. The inclusion criteria for the study were multifaceted. First, GC patients with International Classification of Diseases-O-3/the World health Organization 2008 C160-C169 site records were included. Second, data from SEER-combined Met DX-lung patients with lung metastases were added. Age, race, sex, diagnosis-to-treatment time, and marital status were also evaluated. Additionally, the study needed comprehensive survival and follow-up data. Finally, histological information, GC tumour location and size, tumor-node-metastasis (TNM) stage, number of positive lymph nodes, and liver, brain, and bone metastases were considered. The demographic information also included surgery date, location, adjuvant therapy, chemotherapy, and radiation therapy order. We removed missing data such as "Blank" or "unknown" and excluded individuals without a stomach cancer pathology diagnosis, thus improving the quality of the data. Additional information appears in Figure 1. In addition, it is worth noting that this study successfully adhered to all of the Strengthening the Reporting Of Cohort Studies in Surgery (STROCSS) criteria, as outlined in reference[12].

***Clinical pathogenic variables and queue definition***

Patient data were divided into training and validation sets by using the R "caret" package's CreateDataPartition function. The random seed was 2345, therefore, the training set (1156 samples) was 7:3, and the validation set included 496 samples. Model development was performed with the training set, whereas parameter optimization and internal validation were performed with the validation set. The current study examined clinical and pathological factors based on information from the SEER database and the literature. The variables included age, sex, race, primary site, histologic type, T stage, N stage, bone metastases, brain metastases, and liver metastases. Radiation, chemotherapy, surgery, marital status, number of positive lymph nodes, time between diagnosis and treatment initiation, regional lymph node surgery, other regional or distant surgery, and treatment sequence were also variables. This analysis examined 19 criteria. Cancer-specific survival (CSS) was the secondary endpoint of OS.

***Statistical analysis***

All of the statistical analyses were performed in the RStudio environment by using R software, version 4.1.3. Categorical data are presented as frequencies and percentages for the creation of a three-line table and descriptive statistics of clinicopathological parameters, whereas continuous variables are presented as medians and interquartile ranges. Categorical variables were initially factorized for further analysis. To screen for variables, univariate Cox regression analyses were performed on the 19 related factors, including those with a *P* value less than 0.1 according to multivariate Cox regression analyses, to identify the final independent risk factors (*P* < 0.05). Cox regression and the training set were subsequently used to construct a nomogram for OS, which indicates the anticipated survival rates at 1 year and 3 years. A multivariate Cox regression model was used to predict OS, and another Cox model [Cox-American Joint Committee on Cancer (AJCC)] was created by using the 7th edition of the AJCC Cancer Staging Manual. To determine the best reference model for stomach cancer patients with lung metastases, the Net Reclassification Improvement (NRI) and Integrated Discrimination Improvement (IDI) were used to compare these two models. A total of 1533 patients who died entirely from cancer were identified after deaths from other causes were removed. The survfit function in R was used to construct a cumulative incidence curve for stomach cancer patients with lung metastases based on CSS data. The area under the curve (AUC) was used to assess the overall discriminative capacity of the nomogram[13]. Finally, calibration curves, receiver operating characteristic (ROC) curves, and decision curves were used to evaluate the nomogram's performance in the training and validation sets.

**RESULTS**

***Clinical characteristics and general conditions***

In this investigation describing the OS of patients with GC, 1652 patients were included. There were 529 females (32.23%) and 1123 males (67.98%) among the participants, the overwhelming majority of whom were white (76.45%). The participants were randomly assigned to the training (*n* = 1156) or validation (*n* = 499) sets at a 7:3 ratio. The majority of patients with GC (62.89%) had malignant gastric adenocarcinomas, and 69.43% of these patients were older than 60 years. Nearly half of the primary tumours (43.16%) originated in the cardia, whereas a small percentage also originated in the larger curvature of the stomach (2.91%). Significantly, 95.58% of patients did not undergo surgery at the primary site, 94.9% did not undergo surgery on regional lymph nodes, and only 4.66% underwent other noncancer-related operations. The vast majority (97.82%) of patients in this study did not receive direct radiation or cancer-related surgery, as indicated by radiation therapy sequence data. Only 18.64% of patients received radiation therapy, whereas 55.21% received chemotherapy. A substantial percentage of 96.37% of the patients had at least four positive lymph nodes. Seventy-nine percent of all of the patients did not have bone metastases; additionally, 95.7% did not have brain metastases, and 52.72% had liver metastases. At present, 59.93% of patients are solitary. The percentages for tumour staging were as follows: T0 (0.61%), T1 (19.79%), T2 (3.75%), T3 (10.9%), T4 (17.25%), and TX (47.7%). The percentages of patients with different lymph node statuses were as follows: N0 (35.59%), N1 (37.89%), N2 (3.57%), N3 (3.81%), and NX (19.13%). The median interval between diagnosis and treatment was one month. Moreover, a total of 70.16% of all tumours were between 5 and 10 cm in size, whereas 4.36% were larger than 10 cm. Table 1 provides detailed descriptions of the clinical statistics.

According to the analysis of CSS data, the competitive risk model demonstrated the following trends in cumulative incidence rates. As the follow-up period increased, there was a marked increase in the cumulative incidence among patients with GC lung metastases, particularly in individuals older than 60 years (Figure 2A), those with stage N2 tumours (Figure 2B), those with African American individuals (Figure 2C), and stage T3 tumours (Figure 2D). In addition, patients who received radiotherapy (Figure 2E), chemotherapy (Figure 2F) had a higher cumulative incidence associated with treatment in the early stage. In longitudinal analyses, we observe that cohorts devoid of hepatic (Figure 2G), osseous (Figure 2H), or cerebral metastatic involvement (Figure 2I) manifest an augmented cumulative risk profile for metastatic propagation when contrasted with counterparts who have previously undergone such metastatic transitions. From a sociological perspective, unmarried patients initially exhibit a higher cumulative incidence of disease compared to their married counterparts (Figure 2J), possibly reflecting the absence of spousal care and support. However, this discrepancy diminishes over time, suggesting that marital status does not significantly influence the long-term outcome of GCLM. Patients with a greater number of positive lymph nodes (Figure 2K) face a higher risk of lung metastasis early on, while those with fewer positive lymph nodes have a relatively lower risk. This disparity may persist over time, indicating that the number of positive lymph nodes can be a significant early indicator of the risk of lung metastasis. Clinically, this necessitates more aggressive monitoring and intervention for patients with a higher count of positive lymph nodes. Moreover, compared with patients who did undergo surgery for the primary tumour, patients who did not undergo surgery for the primary tumour exhibited a substantial increase in cumulative incidence year-over-year (Figure 2L). The primary site of the tumor (Figure 2M), the treatment sequence (Figure 2N) and the histological type of the tumor (Figure 2O) are crucial in predicting the cumulative incidence of GCLM. Notably, the cumulative incidence of signet ring cell carcinoma among gastric cancer histological types is expected to rise continually. The sequence of treatment modalities is a critical determinant of patient prognostication and risk of tumor recurrence. Observations from treatment sequences have indicated that patients receiving radiotherapy both pre- and post-operatively exhibit the most rapid early increase in cumulative incidence of disease and may have a higher probability of tumor recurrence. By comparing the different colored curves, one can observe the variations in lung metastasis risk among patients with gastric cancer originating from different sites.

***Correlation analysis of study variables***

Before conducting the Cox regression analysis, to ensure that there was no collinearity among the variables, we utilized Spearman correlation analysis. The results of the correlation analysis can be found in Figure 3.

***Variable selection for the nomogram***

The results of the univariate Cox regression analysis indicated that a total of 11 variables, including months from diagnosis to treatment, primary site, surgery, treatment sequence, surgery regarding other regional distance, chemotherapy, radiation, node-positive number, mets at bone, mets at liver, and N stage, exhibited a statistically significant association with OS. The results of the multivariate Cox regression analysis demonstrated that several factors were found to be independent prognostic factors affecting the OS of patients with GC accompanied by lung metastasis. The factors included the duration from diagnosis to treatment [*P* < 0.001, hazard ratio (HR) = 0.86, 95%CI = 0.80-0.93], body (*P* = 0.006, HR = 1.51, 95% = 1.12-2.04), greater curvature (*P* < 0.001, HR = 2.05, 95%CI = 1.39-3.02), surgery (*P* < 0.001, HR = 2.05, 95%CI = 1.39-3.02), radiation prior to surgery (*P* = 0.045, HR = 2.43, 95%CI = 1.02-5.78), nonchemotherapy (*P* < 0.001, HR = 3.63, 95%CI = 3.18-4.14), radiation (*P* = 0.046, HR = 0.84, 95%CI = 0.71-0.99), metastasis at bone (*P* < 0.001, HR = 1.32, 95% = 1.14-1.53), and metastasis at the liver (*P* < 0.001, HR = 1.30, 95%CI = 1.15-1.47), and N3 (*P* = 0.018, HR = 1.45, 95%CI = 1.07-1.98). Table 2 contains comprehensive information about the subject matter.

***Nomogram construction and validation***

Univariate Cox regression analysis was also conducted for all of the patients, with a significance level of *P* < 0.1. Covariates that were found to be significant were subsequently included in the multivariate Cox regression analysis, with a significance level of *P* < 0.05. Figure 4 shows the independent predictive variables for individuals diagnosed with GCLM. The study examined a total of ten variables, encompassing primary tumour location, surgical intervention, other regional surgeries, treatment sequence, radiation therapy, chemotherapy, the number of positive lymph nodes, the presence of bone metastases, liver metastases, and tumour N staging. These variables were utilized to predict the 1-year and 3-year survival rates for patients diagnosed with GCLM. The patients were assigned scores for each risk factor by mapping them upwards on a scale based on the classification of each risk factor. There was a negative correlation between higher scores and survival rates at both the 1-year and 3-year time points. By combining these scores, it is possible to approximate the survival rates at 1-year and 3-year time intervals. The nomogram that was created was validated by using the bootstrap technique. The self-sampling times were set at B = 1000, as shown in Figure 5. The calibration curves of the nomogram in both the training and validation sets demonstrated a high level of agreement between the predicted and actual survival results. The accuracy of the 1-year and 3-year OS prediction nomograms for GCLM patients is demonstrated in Figure 6, respectively, which utilized the ROC curve from the Cox regression model. In the training cohort, the AUC for the 1-year time point was 0.814, as depicted in Figure 6A. Similarly, for the 3-year time point, the AUC was 0.772, as shown in Figure 6B. Furthermore, decision curve analysis (DCA) yielded valuable insights into the validation of the nomogram, as shown in Figure 7. The results of the DCA demonstrated that the performance of the model on the training set was above the baseline and outperformed that of the other models. This finding suggested that the nomogram successfully achieved a favourable trade-off between true positives and false-positives, thus resulting in a greater net benefit at the specified probability threshold. These findings suggest that the nomogram has superior performance compared to other models when evaluated at certain decision thresholds. The Figure 8 depicts survival curves that were generated by using the independent risk variables that were included in the analysis. The present study employed a model to assess the potential risk of mortality due to stomach cancer resulting from several causes. The identification of the total score for each individual variable helped to achieve this effect.

***Risk stratification for GCLM***

In this study, to address various independent prognostic factors, we utilized the 'survival' package in the R language, in addition to the ggsurvplot function[14], to illustrate the Kaplan‒Meier (KM) survival curves and to perform risk stratification for both the OS model and the AJCC model. By utilizing the OS nomogram, we generated a thorough survival score for the patients using the OS nomogram. Patients were divided into two main cohorts (the high-risk group and the low-risk group) based on the median risk score (Figure 9). According to the KM survival analysis, the OS of patients in the low-risk subgroup was greater than that of patients in the high-risk subgroup (Figure 9A). Although the AJCC staging system showed low discriminative power in risk stratification, the OS risk stratification showed substantial discriminative ability (Figure 9B).

***Comparison between the new model and the previous AJCC model***

Individualized prediction has become the cornerstone of such research due to the notable prognostic variations in GC patients with lung metastases. Previously, the prognosis was based on the 7th edition of the AJCC staging method. However, the sole use of the TNM staging system within the AJCC system is not enough to ensure complete and accurate prognostic evaluations. Therefore, it is crucial to combine variables with additional clinical factors. To assess the accuracy and capacity for improvement of the recently constructed model, the nomogram that was developed in this study was contrasted with the previous AJCC staging method. The Cox-AJCC older model architecture, which was based on the Cox nomogram model, considered only age, race, sex, T stage, and N stage. By considering improvements at particular cut-off points, the NRI is used to compare the prediction skills of the old and new models. IDI is used to observe the model's capacity for overall improvement[15]. The ROC curves of the participants at 1 year (Figure 10A) and 3 years (Figure 10B) IDI of the new model indicated favourable performance (AUC > 0.8), and Figure 10C shows a bar graph demonstrating the difference in the AUC between the two models. The bar graph clearly shows that the new model has a higher AUC than the previous model for both 1-year and 3-year forecasts. This shows that the new model outperforms the previous model in terms of prediction. Furthermore, the NRI and IDI values for the first and third years were greater than 0. As shown in Figure 10D, the new model outperforms the AJCC prediction model in terms of accuracy and total improvement.

**DISCUSSION**

As a common digestive tract cancer, GC frequently faces difficulties such as poor identification, limited surgical resection options, and a high likelihood of recurrence[16]. However, stomach cancer is actually a preventable and treatable type of cancer if it is found early and combined with proactive therapeutic methods. Such tactics can significantly reduce the incidence and fatality rates of this disease[17]. Our work, which prognostically examined the 1-year and 3-year survival times of patients with GC pulmonary metastasis using the extensive sample database SEER, provided insights that surpassed the conventional thoroughness and accuracy of the AJCC staging system[18]. The established nomogram model exhibited good discriminatory capacity and calibration. Additionally, risk stratification significantly divided GCLM patients into high- and low-risk groups. The OS rates of these patients at both the 1-year and 3-year intervals showed a distinct decreasing trend as the follow-up period progressed.

The most frequent liver metastases in individuals with recurrences and metastases from GC were those that were accompanied by distant lymph node metastases. The majority of research on GC has primarily focused on liver metastasis, although frequently, the possibility of distant pulmonary metastasis has been disregarded[19]. In this study, a nomogram model was constructed by incorporating additional predictive variables (adjuvant therapy and treatment sequence).

Oh *et al*[20] reported that the development of GC and its doubling time significantly decreased as the stage progressed after confirmation of GC without the need for any therapeutic interventions and by merely observing tumour progression *via* CT and endoscopy; moreover, the survival rate also decreased during follow-up. For T1 stomach cancer, the doubling time was 11.8 months; however, for T4, it was 6.2 months. It took an average of 34 months for early-stage stomach cancer to develop into an advanced stage[20]. The urgency from diagnosis to treatment introduction is further highlighted by this rate of advancement. The risk to the patient was increased by more than 14% in comparison to timely treatment for every month of treatment delay following diagnosis according to multivariate Cox proportional hazards regression analysis (HR = 0.86, CI = 0.80-0.93; *P* = 0.001). These findings demonstrate the critical importance of early diagnosis and therapy for individuals with stomach cancer and other malignancies in terms of their prognosis. Neoadjuvant therapy for GC should be administered no later than 4 wk after surgery, according to studies on this form of treatment by Ahn *et al*[21]. If this delay is delayed past this point, the patient's chance of survival may suffer. Patients with locally progressed and late-stage GC who underwent surgery within 3-5 wk of receiving neoadjuvant therapy experienced the greatest improvements in survival according to studies by Wang *et al*[22]. However, our research showed that different treatment modalities and approaches can lead to differences in patient prognosis. In particular, patients who receive radiotherapy prior to surgery at the primary site have a 2.43-fold increased risk of poor survival (*P* = 0.045, CI = 1.02–5.78) compared to patients who did not receive radiation or surgery. This result deviates from prior research[23].

Patients with original tumours on the greater curvature of the stomach had considerably greater OS rates than did those with tumours in the antrum, cardia, or lesser curvature according to Korivi *et al*'s retrospective analysis[24]. In contrast to original tumours in the gastric antrum, predictive risk variables for lung metastasis in individuals with GC included those originating from the cardia, pylorus, stomach body, stomach fundus, greater curvature, and lesser curvature. After accounting for other confounding factors, a larger curvature (*P* = 0.001), additional stomach regions (*P* = 0.036), and the stomach body (*P* = 0.006) were particularly found to be independent predictive factors.

Through Cox regression analysis, Dong *et al*[15] reported that surgery was an independent prognostic factor impacting patient survival. When regarding the surgical management of patients with far-reaching metastases, debate has persisted in recent years[15]. In addition, the single-factor and multifactor analyses of this study's data regarding surgical therapy at the primary site showed that patients who underwent surgery at the primary lesion had a 57% lower risk than did those who did not (*P* = 0.001, HR = 0.43, CI = 0.31–0.61). Additionally, surgical intervention in these metastatic areas may be able to stop the spread of the disease. Lung metastasis from GC may also occur, as may lymph node and distant organ metastases in other regions. This study provided concrete evidence that these surgical procedures improved patients' chances of survival. Some Japanese specialists have recommended against performing palliative surgery to remove the main lesion for patients with incurable distant metastases. However, resection of the metastatic location may be an option for cancers that have limited metastatic spread[25].

For advanced stomach cancer, radiation and chemotherapy are considered essential supplementary treatment strategies. Both radiation and chemotherapy are essential components of the overall treatment plan for GC lung metastases. The combination of these two treatments can increase the success rate of treatment, increase patient survival time, and improve patient quality of life. This study emphasizes this point by identifying radiotherapy and chemotherapy as being helpful elements for lengthening patient survival. Additionally, new molecularly targeted medications and immunotherapies are in development and exhibit promise as potent adjuvant therapies. An understanding of targeted biomarkers and medication features is crucial for choosing the right tailored therapy as the number of pharmaceuticals available increases[26].

In this study, patients with liver and bone metastases had a 1.30- and 1.32-fold greater chance of dying from GC than did those without liver and bone metastases, respectively, at a very significant level (*P* < 0.001). Liang *et al*[27] conducted a retrospective investigation by using Cox survival analysis and public databases. Early bone metastasis in GC is rare, and synchronous bone metastasis is even rarer. Those individuals with GC who had surgery at the main site and who had bone metastases had a longer median survival time (9.0 months) than did those who did not have surgery (3.0 months). Furthermore, the median survival time for patients with GCLM was 7 months among those who did not have any skeletal-related disorders at the time of bone metastasis diagnosis. Additionally, treatment may improve a patient's relative survival in patients with stomach cancer that metastasizes to the bone[27,28]. CT and retrospective analysis allowed Hori *et al*[29] to categorize liver metastasis patients into H1, H2, and H3 grades. According to their study, the size and number of tumours that had spread to the liver were positively correlated with the prognosis of liver metastasis in patients with GC[29].

Younger patient populations were found to be independent risk variables for underestimating lymph node metastases and clinical N staging in Kim *et al's* study[30]. Younger patients with GC had an increased risk of lymph node metastases, as well as an inclination to underestimate staging[30]. In this study, the risk of a poor prognosis increased with increasing stage; this difference was most noticeable at the N3 stage (*P* = 0.018, HR = 1.45, CI = 1.07-1.98), when the cumulative incidence was highest.

This study used the SEER database to evaluate the long-term prognostic outcomes of patients with GC pulmonary metastasis and presented a unique model. The study utilized extensive sample data from GC patients and clinicians worldwide. The accuracy of prognostic predictions for GCLM improved when using this new approach, which differs significantly from the traditional AJCC model in that it incorporates variables such as clinical auxiliary treatment elements and demographic variables that were not present in earlier models. Precision medicine is becoming a reality with rapid progress in medical technology. Rapid advancements in fields such as radiomics, metabolomics, and genomics have led to the wider use of multiomics analytic techniques. Doctors carefully collect all of the relevant information during patient visits and perform in-depth evaluations. Thorough case studies have shown that the prognosis of patients with GC is significantly influenced by the extent of tumour resection, the depth of tumour invasion, and the extent of lymph node metastasis[31-33]. Researchers are also paying attention to the relationship between prognosis and the following three important parameters: The number of negative lymph nodes, the rate of lymph node metastases, and the presence of free tumour cells in the peritoneal cavity[34]. These discoveries have the potential to improve the survival rates of stomach cancer patients by enabling the development of more customized treatment plans. In the future, prognostic prediction models that make use of large amounts of medical data and sophisticated algorithms are expected to be fundamental to customized care. These developments are expected to improve treatment outcomes for patients with stomach cancer, thereby prolonging their survival.

There are undoubtedly certain aspects of this study that need to be further refined. In the present study, we analysed data from 1652 patients with GCLM, of which 67.98% were male, and 32.02% were female. Although this distribution reflects the sample available in the SEER database, we acknowledge that the unequal sex ratio may introduce certain biases in our findings. This disproportionate sex distribution could affect the generalizability of our results. Future studies should aim to include a more balanced male-to-female patient ratio to increase the applicability of the findings across genders. Additionally, the investigation of any sex-specific differences in the prognosis of GCLM could offer more nuanced insights into disease progression and management. Other limitations include the retrospective design of the study and the lack of additional validation through randomized controlled trials. Crucially, the study focused mainly on the ability of lung metastasis to predict the survival of patients with GC, and the clinical research characteristics in the database were limited to patients with metastases to the liver, bone, or brain. Other metastatic sites (such as the abdominal cavity and ovaries) and basic patient admission data (weight, blood pressure, and smoking history), past medical history, personal history, and important biochemical indicators (such as blood biochemistry, liver function, kidney function, and electrocardiogram) that were absent from the SEER database were not included in the study, nor were comprehensive treatment plans available for chemotherapy or radiation therapy. These missing facts may have an impact on a thorough comprehension of the prognosis. Future investigations should examine the connection between these clinicopathological factors and additional putative biomarkers, as well as how to incorporate this information into more intricate prediction models. Furthermore, external validation and use of the model in multicentre clinical trials are significant avenues that warrant consideration.

**CONCLUSION**

A risk nomogram for the OS of patients with lung metastases from stomach cancer was effectively established in this study. The suggested nomogram uses CSS to determine the cumulative incidence of patient prognosis and efficiently separates prognostic groupings. This nomogram showed constant reliability and clinical application after validation. One of the most important innovations in our study is the use of extremely extensive and precise clinicopathological variables. These characteristics are expected to substantially improve the predictive power for OS and CSS. This approach will help surgeons in creating more individualized therapeutic and prognostic strategies for these patients. Subsequent studies will require additional external verification to prospectively evaluate the model.

**ARTICLE HIGHLIGHTS**

***Research background***

Gastric cancer (GC) is a globally prevalent malignancy known for its aggressive behaviour and poor survival outcomes, especially when metastasis occurs. Recent research has focused on identifying more precise prognostic factors to tailor individual treatment strategies. By developing a nomogram using e Surveillance, epidemiology, and end results program (SEER) database, this study addresses a critical gap in understanding GC lung metastasis (GCLM). This approach goes beyond traditional American Joint Committee on Cancer staging, thus offering a more accurate predictive model for overall survival (OS) and risk categorization in GCLM patients. This contribution is significant because it can inform better clinical decision-making and potentially improve outcomes in this patient population.

***Research motivation***

This study was motivated by the need to improve prognostic predictions for GCLM, which is a condition associated with notably poor survival outcomes. The aim was to address key problems in current prognostic models, such as their limited ability to accurately predict OS and cumulative incidence prediction (CIP) in GCLM patients. The significance of solving these problems lies in providing clinicians with a more effective tool for risk stratification, which can guide personalized treatment plans and potentially improve patient outcomes. By developing a more accurate and comprehensive nomogram using data from the SEER database, this research contributes to the advancement of precision medicine in GC care, particularly for those with lung metastases.

***Research objectives***

The primary objective of this study was to develop an accurate prognostic nomogram for patients with GCLM by using data from the SEER database. This nomogram aims to predict OS and CIP more effectively than existing models. The study successfully identified significant prognostic factors related to GCLM, thus integrating these factors into a model that offers more precise survival predictions and risk stratification. These objectives are significant for future research, as they will enhance the understanding of GCLM and aid in the advancement of personalized treatment strategies, thus potentially improving patient outcomes in this challenging cancer subtype.

***Research methods***

This research utilized a retrospective analysis of data from the SEER database comprising patients with GCLM from January 2000 to December 2020. The methods included univariate and multivariate Cox regression analyses to identify independent prognostic factors, and a nomogram was developed for predicting OS. This nomogram was validated by using the time-dependent area under the curve and calibration curves. Additionally, decision curve analysis was used to assess the clinical usefulness of the model. The novelty of this research lies in the comprehensive approach combining various clinical and demographic variables that have not been previously integrated in traditional models, thereby enhancing the prognostic accuracy for GCLM patients.

***Research results***

This research established a novel prognostic nomogram for predicting OS in patients with GCLM that included factors such as age, sex, race, tumour size, and treatment modality. This model demonstrated superior predictive accuracy compared to traditional staging systems, thereby significantly contributing to personalized treatment planning and risk assessment in GCLM patients. However, challenges remain in validating the nomogram across diverse patient populations and integrating emerging biomarkers and genetic data for further refinement of the predictive model.

***Research conclusions***

The research concluded with the successful development of a prognostic nomogram for predicting the OS of patients with lung metastases from GC based on an extensive and precise collection of clinicopathological variables. Moreover, this approach helps to appropriately classify patients into high-risk and low-risk groups, thereby guiding treatment. This model, which was validated for its reliability and clinical application, represents a significant innovation in personalized treatment and prognosis strategies for GC.

***Research perspectives***

Future efforts will focus on additional external validation and prospective evaluations to further establish the model's efficacy and applicability in clinical settings.

**ACKNOWLEDGEMENTS**

I would like to express my sincere gratitude to the Surveillance, epidemiology, and end results program database for providing invaluable data resources for this study. I also wish to extend my deepest appreciation to my research team, whose expertise, dedication, and support made this research possible. A special thanks to Wei-Wei Chen for her invaluable insights and feedback; her contributions were pivotal in enhancing the quality of this research.

**REFERENCES**

1 **Ilic M**, Ilic I. Epidemiology of stomach cancer. *World J Gastroenterol* 2022; **28**: 1187-1203 [PMID: 35431510 DOI: 10.3748/wjg.v28.i12.1187]

2 **Luo Z**, Rong Z, Huang C. Surgery Strategies for Gastric Cancer With Liver Metastasis. *Front Oncol* 2019; **9**: 1353 [PMID: 31921626 DOI: 10.3389/fonc.2019.01353]

3 **Gu J**, Chu X, Huo Y, Liu C, Chen Q, Hu S, Pei Y, Ding P, Pang S, Wang M. Gastric cancer-derived exosomes facilitate pulmonary metastasis by activating ERK-mediated immunosuppressive macrophage polarization. *J Cell Biochem* 2023; **124**: 557-572 [PMID: 36842167 DOI: 10.1002/jcb.30390]

4 **Yashima K**, Onoyama T, Kurumi H, Takeda Y, Yoshida A, Kawaguchi K, Yamaguchi N, Isomoto H. Current status and future perspective of linked color imaging for gastric cancer screening: a literature review. *J Gastroenterol* 2023; **58**: 1-13 [PMID: 36287268 DOI: 10.1007/s00535-022-01934-z]

5 **Röcken C**. Predictive biomarkers in gastric cancer. *J Cancer Res Clin Oncol* 2023; **149**: 467-481 [PMID: 36260159 DOI: 10.1007/s00432-022-04408-0]

6 **Rihawi K**, Ricci AD, Rizzo A, Brocchi S, Marasco G, Pastore LV, Llimpe FLR, Golfieri R, Renzulli M. Tumor-Associated Macrophages and Inflammatory Microenvironment in Gastric Cancer: Novel Translational Implications. *Int J Mol Sci* 2021; **22** [PMID: 33916915 DOI: 10.3390/ijms22083805]

7 **Yago A**, Haruta S, Ueno M, Hamada Y, Ogawa Y, Ohkura Y, Urabe M, Udagawa H. Adequate period of surveillance in each stage for curatively resected gastric cancer: analyzing the time and rates of recurrence. *Gastric Cancer* 2021; **24**: 752-761 [PMID: 33400037 DOI: 10.1007/s10120-020-01147-4]

8 **Liu S**, Chen L, Zeng J, Chen Y. A prognostic model based on the COL1A1-network in gastric cancer. *Am J Transl Res* 2023; **15**: 1640-1653 [PMID: 37056863]

9 **Kong J**, Zheng J, Cai J, Wu S, Diao X, Xie W, Chen X, Liao C, Yu H, Fan X, Huang C, Liu Z, Chen W, Lv Q, Qin H, Huang J, Lin T. A nomogram for individualized estimation of survival among adult patients with adrenocortical carcinoma after surgery: a retrospective analysis and multicenter validation study. *Cancer Commun (Lond)* 2019; **39**: 80 [PMID: 31775884 DOI: 10.1186/s40880-019-0426-0]

10 **Balachandran VP**, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* 2015; **16**: e173-e180 [PMID: 25846097 DOI: 10.1016/S1470-2045(14)71116-7]

11 **Sampson JH**, Lad SP, Herndon JE 2nd, Starke RM, Kondziolka D. SEER insights. *J Neurosurg* 2014; **120**: 297-298 [PMID: 24286150 DOI: 10.3171/2013.6.JNS13993]

12 **Agha R**, Abdall-Razak A, Crossley E, Dowlut N, Iosifidis C, Mathew G; STROCSS Group. STROCSS 2019 Guideline: Strengthening the reporting of cohort studies in surgery. *Int J Surg* 2019; **72**: 156-165 [PMID: 31704426 DOI: 10.1016/j.ijsu.2019.11.002]

13 **Avalos-Pacheco A**, Ventz S, Arfè A, Alexander BM, Rahman R, Wen PY, Trippa L. Validation of Predictive Analyses for Interim Decisions in Clinical Trials. *JCO Precis Oncol* 2023; **7**: e2200606 [PMID: 36848613 DOI: 10.1200/PO.22.00606]

14 **Sonabend R**, Király FJ, Bender A, Bischl B, Lang M. mlr3proba: an R package for machine learning in survival analysis. *Bioinformatics* 2021; **37**: 2789-2791 [PMID: 33523131 DOI: 10.1093/bioinformatics/btab039]

15 **Dong Z**, Zhang Y, Geng H, Ni B, Xia X, Zhu C, Liu J, Zhang Z. Development and validation of two nomograms for predicting overall survival and cancer-specific survival in gastric cancer patients with liver metastases: A retrospective cohort study from SEER database. *Transl Oncol* 2022; **24**: 101480 [PMID: 35868142 DOI: 10.1016/j.tranon.2022.101480]

16 **Zhou L**, Li SH, Wu Y, Xin L. Establishment of a prognostic model of four genes in gastric cancer based on multiple data sets. *Cancer Med* 2021; **10**: 3309-3322 [PMID: 33934516 DOI: 10.1002/cam4.3654]

17 **Norwood DA**, Montalvan-Sanchez E, Dominguez RL, Morgan DR. Gastric Cancer: Emerging Trends in Prevention, Diagnosis, and Treatment. *Gastroenterol Clin North Am* 2022; **51**: 501-518 [PMID: 36153107 DOI: 10.1016/j.gtc.2022.05.001]

18 **RuiYang W**, ZhiMing Y, Jiao F, Liang Z, Gang Z. Evaluation and Recommendation of the 8th Edition of American Joint Committee on Cancer (AJCC) Staging System for Intrahepatic Cholangiocarcinoma (ICC) in 820 Patients from the Surveillance, Epidemiology, and End Results (SEER) Database. *J Gastrointest Surg* 2021; **25**: 145-154 [PMID: 32193847 DOI: 10.1007/s11605-020-04557-y]

19 **Muroi T**, Okui M, Sasaki A. [Surgical Resection of Pulmonary Metastases from Gastric Cancer;Report of a Case]. *Kyobu Geka* 2020; **73**: 1045-1047 [PMID: 33268760]

20 **Oh SY**, Lee JH, Lee HJ, Kim TH, Huh YJ, Ahn HS, Suh YS, Kong SH, Kim GH, Ahn SJ, Kim SH, Choi Y, Yang HK. Natural History of Gastric Cancer: Observational Study of Gastric Cancer Patients Not Treated During Follow-Up. *Ann Surg Oncol* 2019; **26**: 2905-2911 [PMID: 31190210 DOI: 10.1245/s10434-019-07455-z]

21 **Ahn GT**, Baek SK, Han JJ, Kim HJ, Jeong SJ, Maeng CH. Optimal time interval from surgery to adjuvant chemotherapy in gastric cancer. *Oncol Lett* 2020; **20**: 32 [PMID: 32774505 DOI: 10.3892/ol.2020.11893]

22 **Wang Y**, Liu Z, Shan F, Ying X, Zhang Y, Li S, Jia Y, Li Z, Ji J. Optimal Timing to Surgery After Neoadjuvant Chemotherapy for Locally Advanced Gastric Cancer. *Front Oncol* 2020; **10**: 613988 [PMID: 33392098 DOI: 10.3389/fonc.2020.613988]

23 **Chen J**, Guo Y, Fang M, Yuan Y, Zhu Y, Xin Y, Zhang L. Neoadjuvant chemoradiotherapy for resectable gastric cancer: A meta-analysis. *Front Oncol* 2022; **12**: 927119 [PMID: 35992846 DOI: 10.3389/fonc.2022.927119]

24 **Korivi BR**, Faria S, Aly A, Sun J, Patnana M, Jensen CT, Wagner-Bartak N, Bhosale PR. Intestinal and diffuse gastric cancer: a retrospective study comparing primary sites. *Clin Imaging* 2019; **56**: 33-40 [PMID: 30870726 DOI: 10.1016/j.clinimag.2019.03.002]

25 **Japanese Gastric Cancer Association**. Japanese gastric cancer treatment guidelines 2018 (5th edition). *Gastric Cancer* 2021; **24**: 1-21 [PMID: 32060757 DOI: 10.1007/s10120-020-01042-y]

26 **Sato Y**, Okamoto K, Kida Y, Mitsui Y, Kawano Y, Sogabe M, Miyamoto H, Takayama T. Overview of Chemotherapy for Gastric Cancer. *J Clin Med* 2023; **12** [PMID: 36835872 DOI: 10.3390/jcm12041336]

27 **Liang C**, Chen H, Yang Z, Han C, Ren C. Risk factors and prognosis of bone metastases in newly diagnosed gastric cancer. *Future Oncol* 2020; **16**: 733-748 [PMID: 32285687 DOI: 10.2217/fon-2019-0728]

28 **Imura Y**, Tateiwa D, Sugimoto N, Inoue A, Wakamatsu T, Outani H, Tanaka T, Tamiya H, Yagi T, Naka N, Okawa S, Tabuchi T, Takenaka S. Prognostic factors and skeletal-related events in patients with bone metastasis from gastric cancer. *Mol Clin Oncol* 2020; **13**: 31 [PMID: 32765878 DOI: 10.3892/mco.2020.2101]

29 **Hori S**, Honda M, Kobayashi H, Kawamura H, Takiguchi K, Muto A, Yamazaki S, Teranishi Y, Shiraso S, Kono K, Kamiga T, Iwao T, Yamashita N. A grading system for predicting the prognosis of gastric cancer with liver metastasis. *Jpn J Clin Oncol* 2021; **51**: 1601-1607 [PMID: 34491361 DOI: 10.1093/jjco/hyab140]

30 **Kim W**, Youn S, Won Y, Min S, Park YS, Ahn SH, Park DJ, Kim HH. Clinicopathologic Characteristics of Young Gastric Cancer Patients: Diagnostic Staging Accuracy and Survival. *J Minim Invasive Surg* 2020; **23**: 163-171 [PMID: 35601641 DOI: 10.7602/jmis.2020.23.4.163]

31 **Li X**, Yang J. Association of tumor deposits with tumor-infiltrating lymphocytes and prognosis in gastric cancer. *World J Surg Oncol* 2022; **20**: 58 [PMID: 35220959 DOI: 10.1186/s12957-022-02507-3]

32 **Woo Y**, Behrendt CE, Yang A, Hahn M, Goel A, Li H, Yuan YC, Fong Y. Tumor Epigenetic Signature and Survival in Resected Gastric Cancer Patients. *J Am Coll Surg* 2021; **232**: 483-491.e1 [PMID: 33465468 DOI: 10.1016/j.jamcollsurg.2020.12.023]

33 **Wang X**, Chen Y, Gao Y, Zhang H, Guan Z, Dong Z, Zheng Y, Jiang J, Yang H, Wang L, Huang X, Ai L, Yu W, Li H, Dong C, Zhou Z, Liu X, Yu G. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun* 2021; **12**: 1637 [PMID: 33712598 DOI: 10.1038/s41467-021-21674-7]

34 **Chen ZD**, Zhang PF, Xi HQ, Wei B, Chen L, Tang Y. Recent Advances in the Diagnosis, Staging, Treatment, and Prognosis of Advanced Gastric Cancer: A Literature Review. *Front Med (Lausanne)* 2021; **8**: 744839 [PMID: 34765619 DOI: 10.3389/fmed.2021.744839]

**Footnotes**

**Institutional review board statement:** The SEER database is a nationwide cancer registry funded by the National Cancer Institute, which operates across multiple centers and populations. It does not undergo medical ethics review and does not necessitate informed consent. The data used in this study is from the United States public database SEER.

**Informed consent statement:** The SEER database is a multi-center and multi-population registry funded by the National Cancer Institute that is not subject to medical ethics review and does not require informed consent.

**Conflict-of-interest statement:** All the authors report no relevant conflicts of interest for this article.

**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 article 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 NonCommercial (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: https://creativecommons.org/Licenses/by-nc/4.0/

**Provenance and peer review:** Unsolicited article; Externally peer reviewed.

**Peer-review model:** Single blind

**Peer-review started:** November 3, 2023

**First decision:** December 6, 2023

**Article in press:**

**Specialty type:** Gastroenterology and hepatology

**Country/Territory of origin:** China

**Peer-review report’s scientific quality classification**

Grade A (Excellent): 0

Grade B (Very good): B

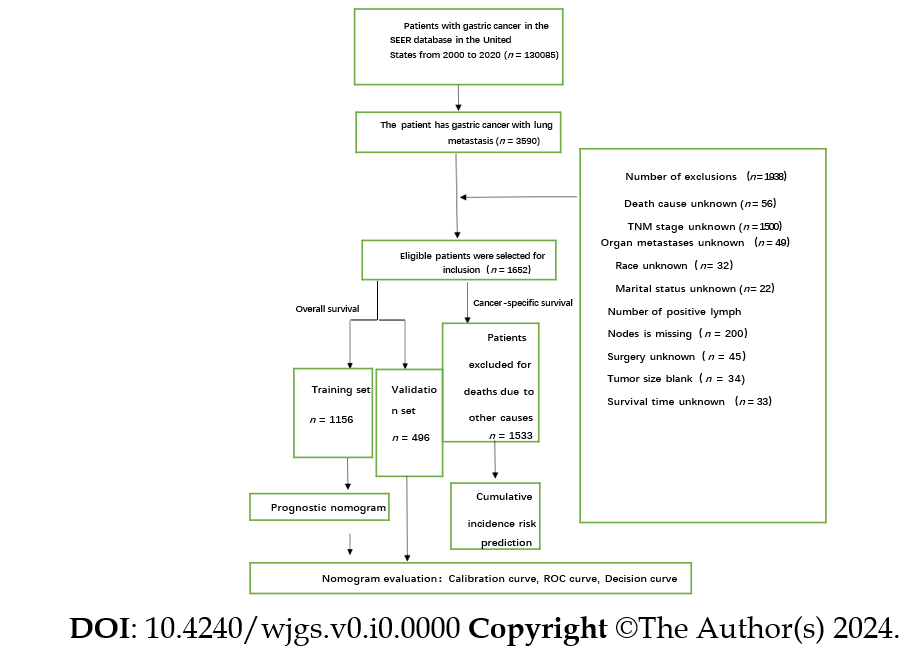
Grade C (Good): 0

Grade D (Fair): 0

Grade E (Poor): 0

**P-Reviewer:** Nittari G, Italy **S-Editor:** Li L **L-Editor:** A **P-Editor:**

**Figure Legends**



**Figure 1 Gastric cancer with lung metastasis patient screening process flowchart.** GCLM: Gastric cancer with lung metastasis.

图表

描述已自动生成 图表

描述已自动生成

图表, 散点图

描述已自动生成 图表, 折线图

描述已自动生成

图表

描述已自动生成 图表

描述已自动生成

图表

描述已自动生成 图表, 散点图

描述已自动生成

图表, 散点图

描述已自动生成 图表

描述已自动生成

图表

描述已自动生成 图表

描述已自动生成

图表, 箱线图

描述已自动生成 图表

描述已自动生成

图表

描述已自动生成



**Figure 2 Cumulative Incidence Prediction plot of cancer special survival in gastric cancer with lung metastasis.** A: Age; B: N stage; C: Race; D: T stage; E: Radiotherapy; F: Chemotherapy; G: Liver Metastasis; H: Bone Metastasis; I: Brain metastasis; J: Marital status; K: Node positive number; L: Surgery; M: Primary site; N: Treatment sequence; O: Histological type. CSS: Cancer special survival; GCLM: Gastric cancer with lung metastasis; CIF: Cumulative incidence function.

图片包含 日历

描述已自动生成



**Figure 3 All included variables' Pearson correlation analysis.** AJCC: American Joint Committee on Cancer; Surg: Surgery; T: Tumor; N: Node; LN: Lymph node; Reg: Regional.



**Figure 4 The overall survival Nomgram for gastric cancer with lung metastasis.** AJCC: American Joint Committee on Cancer; Surg: Surgery; T: Tumor; N: Node; LN: Lymph node; Reg: Regional; Surg: Surgery; Oth: Other; Dis: Disease.

图表, 折线图

描述已自动生成图表, 折线图

描述已自动生成

图表, 折线图

描述已自动生成图表, 折线图

描述已自动生成

**Figure 5 Gastric cancer with lung metastasis calibration curves.** A: 12-month likelihoods of overall survival (OS) in the training dataset; B: 36-month likelihoods of OS in the training dataset; C: 12-month likelihoods of OS in the validation dataset; D: 36-month likelihoods of OS in the validation dataset.

图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成

**Figure 6 Time-dependent area under the curve and receiver operating characteristic curves of overall survival.** A: Receiver operating characteristic (ROC) curves corresponding to 1-year in the training cohort; B: ROC curves corresponding to 3-year overall survival in the training cohort; C: ROC curves corresponding to 1-year in the validation cohort; D: ROC curves corresponding to 3-year cancer-specific survival in the validation cohort. AUC: Area under the curve.

图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成

**Figure 7** **Decision curve analysis of the nomogram in the estimation of overall survival.** A: The Decision curve analysis (DCA) curve for the 1-year overall survival of the training dataset; B: The DCA curve for the 1-year overall survival of the validation dataset; C: The DCA curve for the 3-year overall survival of the training dataset; D: The DCA curve for the 3-year overall survival of the validation dataset.

图片包含 日程表

描述已自动生成 图片包含 图形用户界面

描述已自动生成 图片包含 图形用户界面

描述已自动生成 图表, 直方图

描述已自动生成 图片包含 图示

描述已自动生成 图片包含 图表

描述已自动生成 图形用户界面, 应用程序

描述已自动生成 图片包含 图形用户界面

描述已自动生成 图片包含 图形用户界面

描述已自动生成

**Figure 8 Survival curves for different features of overall survival.** A: Primary site; B: Surgery; C: Surgery other regional distant; D: Treatment Sequence; E: Radiation; F: Chemotherapy; G: Node stage; H: Metastasis at bone; I: Metastasis at liver. HR: Hazard ratio; Surg: Surgery; Oth: Other; Reg: Regional; Dis: Disease.

直方图

低可信度描述已自动生成 图片包含 图表

描述已自动生成

**Figure 9 Comparison of Kaplan–Meier curves of gastric cancer with lung metastasis patients between new Cox model and Cox-American Joint Committee on Cancer.** A: Kaplan–Meier overall survival curves of gastric cancer with lung metastasis (GCLM) patients with different risks stratified; B: Kaplan–Meier Cox-American Joint Committee on Cancer curves of GCLM patients with different risks stratified.

图表, 折线图

描述已自动生成 图表, 折线图

描述已自动生成 图表, 条形图

描述已自动生成 图表, 直方图

描述已自动生成

**Figure 10 Comparison between the new and old models of net reclassification improvement and integrated discrimination improvement.** A: Receiver operating characteristic (ROC) curve of the participants in the 1-year integrated discrimination improvement (IDI) of the new model; B: ROC curve of the participants in the 3-year IDI of the new model; C: Area under the receiver operating characteristic curve difference between new and old model in 1-year and 3-year; D: New model net reclassification improvement and IDI for 1 Year and 3 Years. ROC: Receiver operating characteristic; AUC: Area under the receiver operating characteristic curve; NRI: Net reclassification improvement; IDI: Integrated discrimination improvement.

**Table 1 Comparison of clinical features between the training and validation sets in the overall survival group, *n* (%)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Total (*n* = 1652)** | **Train set (*n* = 1156)** | **Valid set (*n* = 496)** | **Statistic** | ***P* value** |
| Months from diagnosis to treatment, M (Q1, Q3) | 1.00 (1.00 - 1.00) | 1.00 (1.00 - 1.00) | 1.00 (1.00 - 1.00) | Z = 1.394 | 0.252 |
| Tumor size |  |  |  | *χ*2 = 1.983 | 0.371 |
| ≤ 5 | 421 (25.48) | 302 (26.12) | 119 (23.99) |  |  |
| 5-10 | 1159 (70.16) | 800 (69.20) | 359 (72.38) |  |  |
| ≥ 10 | 72 (4.36) | 54 (4.67) | 18 (3.63) |  |  |
| Age |  |  |  | *χ*2 = 0.077 | 0.782 |
| < 60 | 505 (30.57) | 351 (30.36) | 154 (31.05) |  |  |
| ≥ 60 | 1147 (69.43) | 805 (69.64) | 342 (68.95) |  |  |
| Race |  |  |  | *χ*2 = 1.121 | 0.571 |
| Black | 190 (11.5) | 127 (10.99) | 63 (12.70) |  |  |
| Other | 199 (12.05) | 138 (11.94) | 61 (12.30) |  |  |
| White | 1263 (76.45) | 891 (77.08) | 372 (75.00) |  |  |
| Sex |  |  |  | *χ*2 = 1.552 | 0.213 |
| Female | 529 (32.02) | 381 (32.96) | 148 (29.84) |  |  |
| Male | 1123 (67.98) | 775 (67.04) | 348 (70.16) |  |  |
| Primary site |  |  |  | *χ*2 = 3.797 | 0.803 |
| Antrum | 165 (9.99) | 119 (10.29) | 46 (9.27) |  |  |
| Body | 116 (7.02) | 78 (6.75) | 38 (7.66) |  |  |
| Cardia | 713 (43.16) | 494 (42.73) | 219 (44.15) |  |  |
| Fundus | 91 (5.51) | 62 (5.36) | 29 (5.85) |  |  |
| Greater curvature | 48 (2.91) | 34 (2.94) | 14 (2.82) |  |  |
| Lesser curvature | 75 (4.54) | 48 (4.15) | 27 (5.44) |  |  |
| Others | 426 (25.79) | 309 (26.73) | 117 (23.59) |  |  |
| Pylorus | 18 (1.09) | 12 (1.04) | 6 (1.21) |  |  |
| Histologic |  |  |  | *χ*2 = 6.574 | 0.037a |
| Adenocarcinoma | 1039 (62.89) | 704 (60.90) | 335 (67.54) |  |  |
| Others | 372 (22.52) | 275 (23.79) | 97 (19.56) |  |  |
| Signet ring cell carcinoma | 241 (14.59) | 177 (15.31) | 64 (12.90) |  |  |
| Surgery |  |  |  | *χ*2 = 1.650 | 0.199 |
| No | 1579 (95.58) | 1100 (95.16) | 479 (96.57) |  |  |
| Yes | 73 (4.42) | 56 (4.84) | 17 (3.43) |  |  |
| Regional lymph node surgery |  |  |  | *χ*2 = 1.063 | 0.302 |
| No | 1568 (94.92) | 1093 (94.55) | 475 (95.77) |  |  |
| Yes | 84 (5.08) | 63 (5.45) | 21 (4.23) |  |  |
| Surgery other regional/distant |  |  |  | *χ*2 = 0.001 | 0.976 |
| No | 1575 (95.34) | 1102 (95.33) | 473 (95.36) |  |  |
| Yes | 77 (4.66) | 54 (4.67) | 23 (4.64) |  |  |
| Treatment sequence |  |  |  | - | 0.057 |
| No radiation and/or cancer-directed surgery | 1616 (97.82) | 1126 (97.40) | 490 (98.79) |  |  |
| Radiation after surgery | 24 (1.45) | 22 (1.90) | 2 (0.40) |  |  |
| Radiation before and after surgery | 1 (0.06) | 1 (0.09) | 0 (0.00) |  |  |
| Radiation prior to surgery | 11 (0.67) | 7 (0.61) | 4 (0.81) |  |  |
| Radiation |  |  |  | *χ*2 = 0.380 | 0.537 |
| No | 1344 (81.36) | 936 (80.97) | 408 (82.26) |  |  |
| Yes | 308 (18.64) | 220 (19.03) | 88 (17.74) |  |  |
| Chemotherapy |  |  |  | *χ*2 = 0.601 | 0.438 |
| No | 740 (44.79) | 525 (45.42) | 215 (43.35) |  |  |
| Yes | 912 (55.21) | 631 (54.58) | 281 (56.65) |  |  |
| Nodes positive |  |  |  | *χ*2 = 5.288 | 0.021 |
| < 4 | 60 (3.63) | 50 (4.33) | 10 (2.02) |  |  |
| ≥ 4 | 1592 (96.37) | 1106 (95.67) | 486 (97.98) |  |  |
| Metastases at bone |  |  |  | *χ*2 = 2.578 | 0.108 |
| No | 1305 (79) | 901 (77.94) | 404 (81.45) |  |  |
| Yes | 347 (21) | 255 (22.06) | 92 (18.55) |  |  |
| Metastases at brain |  |  |  | *χ*2 = 3.751 | 0.053 |
| No | 1581 (95.7) | 1099 (95.07) | 482 (97.18) |  |  |
| Yes | 71 (4.3) | 57 (4.93) | 14 (2.82) |  |  |
| Metastases at liver |  |  |  | *χ*2 = 2.773 | 0.096 |
| No | 781 (47.28) | 562 (48.62) | 219 (44.15) |  |  |
| Yes | 871 (52.72) | 594 (51.38) | 277 (55.85) |  |  |
| Marital status at diagnosis |  |  |  | *χ*2 = 0.216 | 0.642 |
| 0 | 990 (59.93) | 697 (60.29) | 293 (59.07) |  |  |
| 1 | 662 (40.07) | 459 (39.71) | 203 (40.93) |  |  |
| T stage |  |  |  | *χ*2 = 3.078 | 0.688 |
| T0 | 10 (0.61) | 5 (0.43) | 5 (1.01) |  |  |
| T1 | 327 (19.79) | 227 (19.64) | 100 (20.16) |  |  |
| T2 | 62 (3.75) | 42 (3.63) | 20 (4.03) |  |  |
| T3 | 180 (10.9) | 132 (11.42) | 48 (9.68) |  |  |
| T4 | 285 (17.25) | 200 (17.30) | 85 (17.14) |  |  |
| TX | 788 (47.7) | 550 (47.58) | 238 (47.98) |  |  |
| N stage |  |  |  | *χ*2 = 0.708 | 0.950 |
| N0 | 588 (35.59) | 408 (35.29) | 180 (36.29) |  |  |
| N1 | 626 (37.89) | 440 (38.06) | 186 (37.50) |  |  |
| N2 | 59 (3.57) | 43 (3.72) | 16 (3.23) |  |  |
| N3 | 63 (3.81) | 46 (3.98) | 17 (3.43) |  |  |
| NX | 316 (19.13) | 219 (18.94) | 97 (19.56) |  |  |

a*P* < 0.05.

M: Median.

**Table 2 Univariate and multivariate analysis of gastric cancer overall survival**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variables** | **Beta** | **SE** | **Z** | ***P* value** | **HR (95%CI)** | **m\_Beta** | **m\_SE** | **m\_Z** | ***P* value** | **HR (95%CI)** |
| Months from diagnosis to treatment | -0.15 | 0.03 | -4.90 | < 0.001a | 0.86 (0.81-0.92) | -0.15 | 0.04 | -4.11 | < 0.001b | 0.86 (0.80-0.93) |
| Tumor size |  |  |  |  |  |  |  |  |  |  |
| ≤ 5 |  |  |  |  | Ref. |  |  |  |  |  |
| 5-10 | 0.13 | 0.07 | 1.96 | 0.05a | 1.14 (1.01-1.31) |  |  |  |  |  |
| ≥ 10 | 0.08 | 0.15 | 0.53 | 0.593 | 1.08 (0.81-1.45) |  |  |  |  |  |
| Age |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  | Ref. |  |  |  |  |  |
| 1 | -0.10 | 0.07 | -1.52 | 0.127 | 0.91 (0.80-1.03) |  |  |  |  |  |
| Race |  |  |  |  |  |  |  |  |  |  |
| White |  |  |  |  | Ref. |  |  |  |  |  |
| Other | 0.19 | 0.09 | 1.99 | 0.047a | 1.20 (1.01-1.45) |  |  |  |  |  |
| Black | 0.14 | 0.10 | 1.42 | 0.154 | 1.15 (0.95-1.39) |  |  |  |  |  |
| Sex |  |  |  |  |  |  |  |  |  |  |
| Female |  |  |  |  | Ref. |  |  |  |  |  |
| Male | 0.03 | 0.06 | 0.42 | 0.677 | 1.03 (0.91-1.16) |  |  |  |  |  |
| Primary site |  |  |  |  |  |  |  |  |  |  |
| Antrum |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Others | 0.19 | 0.11 | 1.72 | 0.085a | 1.21 (0.97-1.50) | 0.24 | 0.11 | 2.10 | 0.036b | 1.26 (1.02-1.58) |
| Cardia | -0.06 | 0.10 | -0.54 | 0.591 | 0.95 (0.77-1.16) | 0.13 | 0.11 | 1.22 | 0.222 | 1.14 (0.92-1.41) |
| Body | 0.14 | 0.15 | 0.97 | 0.331 | 1.16 (0.86-1.55) | 0.41 | 0.15 | 2.73 | 0.006b | 1.51 (1.12-2.04) |
| Pylorus | 0.03 | 0.30 | 0.09 | 0.928 | 1.03 (0.57-1.86) | 0.44 | 0.31 | 1.42 | 0.154 | 1.55 (0.85-2.85) |
| Fundus | 0.19 | 0.16 | 1.19 | 0.234 | 1.21 (0.88-1.65) | 0.12 | 0.16 | 0.77 | 0.442 | 1.13 (0.83-1.55) |
| Lesser curvature | 0.10 | 0.17 | 0.57 | 0.571 | 1.10 (0.79-1.55) | 0.28 | 0.18 | 1.61 | 0.107 | 1.33 (0.94-1.88) |
| Greater curvature | 0.50 | 0.20 | 2.57 | 0.010a | 1.66 (1.13-2.43) | 0.72 | 0.20 | 3.61 | < 0.001b | 2.05 (1.39-3.02) |
| Histologic |  |  |  |  |  |  |  |  |  |  |
| Signet ring cell carcinoma |  |  |  |  | Ref. |  |  |  |  |  |
| Adenocarcinoma | -0.14 | 0.09 | -1.62 | 0.106 | 0.87 (0.74-1.03) |  |  |  |  |  |
| Others | -0.11 | 0.10 | -1.15 | 0.250 | 0.89 (0.74-1.08) |  |  |  |  |  |
| Surgery |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Yes | -0.50 | 0.14 | -3.54 | < 0.001a | 0.60 (0.46-0.80) | -0.83 | 0.17 | -4.77 | < 0.001b | 0.43 (0.31-0.61) |
| Regional lymph node surgery |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  |  |
| Yes | -0.44 | 0.13 | -3.31 | < 0.001a | 0.64 (0.50-0.84) |  |  |  |  |  |
| Treatment sequence |  |  |  |  |  |  |  |  |  |  |
| No radiation and/or cancer-directed surgery |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Radiation prior to surgery | -0.44 | 0.38 | -1.17 | 0.243 | 0.64 (0.31-1.35) | 0.89 | 0.44 | 2.00 | 0.045b | 2.43 (1.02-5.78) |
| Radiation after surgery | -0.78 | 0.23 | -3.44 | < 0.001a | 0.46 (0.29-0.72) | -0.21 | 0.25 | -0.84 | 0.400 | 0.81 (0.49-1.33) |
| Radiation before and after surgery | 0.33 | 1.00 | 0.33 | 0.742 | 1.39 (0.20-9.89) | 1.60 | 1.02 | 1.58 | 0.115 | 4.98 (0.68-36.63) |
| Surgery other regional distance |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Yes | -0.53 | 0.14 | -3.74 | < 0.001a | 0.59 (0.44-0.78) | -0.32 | 0.16 | -1.95 | 0.051 | 0.73 (0.53-1.00) |
| Chemotherapy |  |  |  |  |  |  |  |  |  |  |
| Yes |  |  |  |  | Ref. |  |  |  |  | Ref. |
| No | 1.21 | 0.06 | 19.12 | < 0.001a | 3.35 (2.96-3.79) | 1.29 | 0.07 | 19.02 | < 0.001b | 3.63 (3.18-4.14) |
| Radiation |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Yes | -0.27 | 0.08 | -3.59 | < 0.001a | 0.76 (0.66-0.88) | -0.17 | 0.09 | -1.99 | 0.046b | 0.84 (0.71-0.99) |
| Nodes positive |  |  |  |  |  |  |  |  |  |  |
| ≥ 4 |  |  |  |  | Ref. |  |  |  |  | Ref. |
| < 4 | -0.38 | 0.15 | -2.56 | 0.011a | 0.68 (0.51-0.91) | 0.28 | 0.18 | 1.54 | 0.125 | 1.32 (0.93-1.89) |
| Metastases at brain |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  |  |
| Yes | -0.04 | 0.14 | -0.28 | 0.777 | 0.96 (0.73-1.26) |  |  |  |  |  |
| Metastases at bone |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Yes | 0.17 | 0.07 | 2.40 | 0.017a | 1.19 (1.03-1.37) | 0.28 | 0.07 | 3.71 | < 0.001b | 1.32 (1.14-1.53) |
| Metastases at liver |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | Ref. |  |  |  |  | Ref. |
| Yes | 0.22 | 0.06 | 3.60 | < 0.001a | 1.24 (1.10-1.39) | 0.26 | 0.06 | 4.14 | < 0.001b | 1.30 (1.15-1.47) |
| Marital status at diagnosis |  |  |  |  |  |  |  |  |  |  |
| 1 |  |  |  |  | Ref. |  |  |  |  |  |
| 0 | -0.20 | 0.06 | -3.19 | 0.001a | 0.82 (0.73-0.93) |  |  |  |  |  |
| T stage |  |  |  |  |  |  |  |  |  |  |
| TX |  |  |  |  | Ref. |  |  |  |  |  |
| T3 | -0.24 | 0.10 | -2.40 | 0.016a | 0.79 (0.65-0.96) |  |  |  |  |  |
| T4 | 0.02 | 0.08 | 0.24 | 0.808 | 1.02 (0.87-1.20) |  |  |  |  |  |
| T2 | -0.20 | 0.16 | -1.21 | 0.228 | 0.82 (0.59-1.13) |  |  |  |  |  |
| T1 | -0.11 | 0.08 | -1.35 | 0.177 | 0.90 (0.77-1.05) |  |  |  |  |  |
| T0 | -0.47 | 0.50 | -0.94 | 0.347 | 0.62 (0.23-1.67) |  |  |  |  |  |
| N stage |  |  |  |  |  |  |  |  |  |  |
| N1 |  |  |  |  | Ref. |  |  |  |  | Ref. |
| NX | 0.18 | 0.08 | 2.09 | 0.037a | 1.19 (1.01-1.40) | 0.06 | 0.09 | 0.73 | 0.463 | 1.07 (0.90-1.26) |
| N2 | 0.00 | 0.16 | 0.03 | 0.979 | 1.00 (0.73-1.37) | 0.30 | 0.17 | 1.79 | 0.073 | 1.35 (0.97-1.86) |
| N0 | 0.05 | 0.07 | 0.67 | 0.506 | 1.05 (0.91-1.20) | -0.03 | 0.07 | -0.38 | 0.703 | 0.97 (0.84-1.12) |
| N3 | 0.14 | 0.16 | 0.91 | 0.361 | 1.15 (0.85-1.56) | 0.37 | 0.16 | 2.36 | 0.018b | 1.45 (1.07-1.98) |

a*P* < 0.1: Included in the multivariate analysis.

b*P* < 0.05: There is a significant difference in the multivariate analysis.

HR: Hazard ratio; T: Tumor; N: Node; m\_Beta: Multivariate beta (the coefficient estimate from multivariate analysis); m\_SE: Multivariate standard error (the standard error of the multivariate beta coefficient); m\_Z: Multivariate Z (the Z-score or Z-statistic associated with the multivariate beta coefficient).