Mining prognostic factors of extensive-stage small-cell lung cancer patients using nomogram model

Hongxiang Gao, MD, Yazheng Dang, MD, Tao Qi, MD, Shigao Huang, PhD, Xiaozhi Zhang, MD, PhD

Abstract
This study is to establish the nomogram model and provide clinical therapy decision-making for extensive-stage small-cell lung cancer (ES-SCLC) patients with different metastatic sites using the Surveillance, Epidemiology, and End Results (SEER) Program. A total of 10,025 patients of ES-SCLC with metastasis from January 2010 to December 2016 were enrolled from the SEER database. All samples were randomly divided into a derivation cohort and a validation cohort, and the derivation cohort was divided into 6 groups by different metastatic sites: bone, liver, lung, brain, multiple organs, and other organs. Using Cox proportional hazards models to analyze candidate prognostic factors, screening out the independent prognostic factors to establish the nomogram. Compare the different models by Net reclassification improvement and integrated discrimination improvement. Concordance index (C-index) and the calibration curve were used to verify the prediction efficiency of the nomogram in the derivation cohort and validation cohort.

In the derivation cohort, the median overall survival was 7 months. The overall survival rates at 6-month, 1-year, and 2-year were 55.07%, 24.61%, and 7.56%, respectively. The median survival time was 10, 8, 7, 9, 7, and 6 months for the 6 groups of different metastatic sites: other, bone, liver, lung, brain, and multiple organs, respectively. Age, sex, race, T, N, distant metastatic site, and chemotherapy were contained in the final nomogram prognostic model. The C-index was 0.6569777 in the derivation cohort and 0.8386301 in the validation cohort.

The survival time of ES-SCLC patients with different metastatic sites was significantly different. The nomogram can effectively predict the prognosis of individuals and provide a basis for clinical decision-making.

Abbreviations: AI = American Indian/Alaska native, AIC = Akake information criterion, API = Asian or Pacific Islander, Chemo = chemotherapy, C-index = concordance index, DM = distant metastatic site, ES-SCLC = extensive-stage small-cell lung cancer, IDI = integrated discrimination improvement, NRI = net reclassification improvement.

Keywords: extensive-stage small-cell lung cancer, multiple metastatic, nomogram model, prediction, Surveillance, Epidemiology, and End Results

1. Introduction
The small-cell lung cancer (SCLC) accounts for 16% of all lung cancer, which is characterized by rapid growth, early metastatic spread, and widespread dissemination. Approximately two-thirds of SCLC patients showed significant metastasis at the time of clinical diagnosis, which was classified as ES-SCLC. Liver, brain, bone, adrenals, and lungs are the most common metastatic sites in ES-SCLC patients, which is the main factor for the failure of treatment and poor prognosis. Despite the rapid development of clinical oncology, the prognosis of ES-SCLC is still very poor at present, with the survival time of 60% to 70% of patients ranging from 7 to 11 months, and only 2% of patients reaching the 5-year survival period. To this kind of patient, except the routine chemotherapy, the treatment method is various, but the curative effect is not satisfactory. A previous study based on large cohort data revealed that the survival time of ES-SCLC patients varied with different metastatic sites. Therefore, predicting the patient's survival time based on the patient's metastatic site was essential for the choice of treatment decisions. In this study, the nomogram was established to analyze the survival time of ES-SCLC patients from the SEER database and the individualized patient survival prediction can provide accurate clinical decision-making.
2. Patients and methods

2.1. Patients enrollment

We applied the SEER database of incidence-seer 18 Regs Custom Data with additional treatment fields and the data of the November 2017 Sub (1973–2016) version, which concluded at the end of 2016. This study takes cases of ES-SCLC in adults over 18 years old as the research object. All patients were divided randomly into a derivation cohort and a validation cohort. Pathologic types include ICD-O-3 Hist 8041/3 small cell carcinoma, NOS; 8042/3 oat cell carcinoma; 8043/3 small cell carcinoma, fusiform cell; 8044/3 small cell carcinoma, intermediate cell; 8045/3 combined small cell carcinoma. Because the database did not clearly record the distant metastatic sites before 2010, only cases from 2010 to 2016 were selected in this study. The purpose of this study was to analyze the survival of different metastatic sites of ES-SCLC, so cases with no distant metastasis and cases with unclear metastatic sites were excluded. Cases in which the cause of death was unclear or died of other causes should also be excluded. Figure 1 illustrates the screening flow chart of patients in SEER database.

The data on race, sex, age, American Joint Committee on Cancer (AJCC) 7th edition TNM staging, different metastatic site, and chemotherapy were extracted for all cases. In the SEER database, only 4 sites of distant metastasis were recorded: bone, liver, lung, and brain. According to the research needs, all cases were divided into 6 groups by different metastatic sites: bone; liver; lung; brain; multiple organs; others organs, specifically refers to the exception of bone, liver, lung, brain metastases. This study was approved by the Chang An Hospital Ethics Committee (approval no: CA2019-001-023; Xi’an, China)

2.2. Survival analysis

The life-table method was used to calculate the survival time of different metastatic sites. The survival curve was drawn by the Kaplan–Meier method, and the survival difference was evaluated by the Log-rank test. Categorical variables were grouped based on clinical findings, and decisions on the groups were made before modeling.

2.3. Prognostic factors analysis

The COX proportional risk model was used for multivariate analysis of candidate prognostic factors in the derivation cohort. Independent prognostic factors were screened by the backward LR method.


2.4. Development of an individualized prediction model

Two different prediction models were established based on these independent prognostic factors, and R 3.6.1 version software was used to calculate net reclassification improvement (NRI) and integrated discrimination improvement (IDI) to compare the advantages and disadvantages of the 2 models and select the best model.

2.5. Validation of the nomogram

Internal validation: Demonstrate the reproducibility of the model and verify it in the Derivation cohort. Calibration curves were plotted to assess the calibration of the nomogram, accompanied by the Hosmer–Lemeshow test. To quantify the discrimination performance of the nomogram, Harrell C-index was measured. The nomogram was subjected to bootstrapping validation (1000 bootstraps resamples) to calculate a relatively corrected C-index. External validation: Demonstrate the transportability and generalizability of the model and verify it in the validation cohort. The expected survival rate for each case was calculated in the validation cohort using the regression model and the baseline survival rate from the derivation cohort. Then, the COX regression analysis was performed again with the expected survival rate as the only independent variable in the validation cohort. And finally, the C-index and calibration curve were derived based on the regression analysis.

2.6. Statistical analysis

Statistical analysis was performed using SPSS 24.0 software. All the data were expressed as the mean ± standard deviation. Statistical significance was determined as $P < .05$. The nomogram was established by R 3.6.1 version software on the basis of multivariable COX analysis in the derivation cohort.

3. Results

3.1. Patient characteristics

A total of 10,025 SCLC patients with metastatic were chosen from the SEER database. Among them, 8796 patients died by the end of follow-up, accounting for 87.7% of the total cases. The age ranged from 19 to 94 years old, with a median age of 66 years old. Except for Race ($P = .028$), there was no selection bias of clinical characteristics between the derivation cohort and the validation cohort, which is shown in Table 1.

3.2. Survival analysis

The median survival time of the derivation cohort was 7 months. The 6-month, 1-year, and 2-year survival rates were 55.07%, 24.61%, and 7.56%, respectively (Fig. 2A). The median survival time was 10, 8, 7, 7, 9, and 6 months in the 6 metastatic site groups of the other organs, bone, brain, liver, lung, and multiple organs, respectively. The Log-rank method showed a significant difference in survival between the groups ($\chi^2 = 218.851, P = .000$) (Fig. 3B).

3.3. Prognostic factors analysis

Multivariate analysis of Cox regression models including age, sex, race, T, N, metastatic sites, and chemotherapy. The results in Table 3 show that these 7 factors were independent prognostic factors affecting survival ($P < .05$). There were significant differences among 6 groups of metastatic sites except for brain metastasis ($P = .574$).

3.4. Development of an individualized prediction model

Mod1 was established by the independent prognostic factors including age, sex, race, T, N, metastatic site and chemotherapy,
and Mod2 was established without race. The prediction ability of Mod2 was equivalent compared with that of Mod1 (NRI 0.006, 95% confidence interval [CI] −0.139 to 0.090, $P = .786$), and the prediction ability was improved by 0.1% (IDI 0.001, 95% CI −0.008 to 0.002, $P = .856$), but there was no statistical difference.

3.5. Establish the nomogram

Taking Mod2 as the final model, 6 prognostic factors including age, sex, T, N, metastatic site and chemotherapy were selected to establish the nomogram. According to the levels of each prognostic factor, the individualized survival probability of patients with different metastatic sites was inquired at 6 months, 1 year, and 2 years (Fig. 4).

3.6. Validation of the nomogram

Nomogram was validated in terms of its discrimination and calibration. The C-index in the derivation cohort and validation cohort is 0.6569777 and 0.8386301, respectively. And the calibration curve in Figure 5 shows that the nomogram can effectively predict the prognosis.

4. Discussion

The ES-SCLC patients with multiple organ metastatic were treated with chemotherapy, radiotherapy, or supportive care because of its late onset and short survival expectations. Treatment choices were often made based on expected survival time and response to the treatment. However, there are a lot of prognostic factors, including demographics, tumor-related, and therapy-related characteristics. Analysis based on 14 trials showed that sex, age, PS (performance status), creatinine levels, and a number of metastatic sites were important prognostic factors of ES-SCLC. Furthermore, there was a significant interaction between sex and PS within ES-SCLC, suggesting that PS was highly prognostic in males, with no significant impact on females. Our previous study[^6–^8^] tried to find several strong prognostic factors to establish a prognostic scoring system to predict survival. Sculier et al[^9^] established a Recursive Partitioning Analysis (RPA) grading system by screening out 4 prognostic factors including TNM staging, PS, age, and gender through the analysis of cases from the international staging database of the International Association for the Study of Lung Cancer. In this study, our result showed that race, sex,

| Distance metastatic | 6-mo survival rate | 1-yr survival rate | 2-yr survival rate | Median survival time |
|---------------------|--------------------|--------------------|--------------------|---------------------|
| **Derivation cohort** |                    |                    |                    |                     |
| Total               | 55.07              | 24.61              | 7.56               | 7                   |
| Others              | 66.02              | 37.16              | 15.05              | 10                  |
| Bone                | 61.25              | 28.23              | 6.01               | 8                   |
| Brain               | 52.89              | 28.13              | 9.80               | 7                   |
| Liver               | 52.41              | 20.45              | 4.88               | 7                   |
| Lung                | 62.34              | 32.90              | 14.76              | 9                   |
| Multorgan           | 47.78              | 15.43              | 2.83               | 6                   |
| **Validation cohort** |                    |                    |                    |                     |
| Total               | 56.57              | 24.44              | 6.85               | 8                   |
| Others              | 64.23              | 35.20              | 13.83              | 9                   |
| Bone                | 61.98              | 25.11              | 7.77               | 9                   |
| Brain               | 53.42              | 27.46              | 8.82               | 7                   |
| Liver               | 53.01              | 21.90              | 4.11               | 7                   |
| Lung                | 63.48              | 36.30              | 13.96              | 9                   |
| Multorgan           | 52.41              | 15.08              | 2.16               | 7                   |

ES-SCLC = extensive-stage small cell lung cancer.
age, metastatic site, T stage, N stage, and initial chemotherapy were independent prognostic factors, which were consistent with the previous findings.\textsuperscript{5,8–13} In addition, male patients over 75 years old, with T3N2 stage and above, multiple organ metastases, and without chemotherapy were adverse prognostic factors. Although there was metastasis at the time of diagnosis, the survival period of patients was different due to different metastatic

Table 3
COX regression model multivariate analysis for the 2 models in derivation cohort.

| Prognostic factors | Mod1 |         |         | Mod2 |         |         |
|-------------------|------|---------|---------|------|---------|---------|
|                   | \(\beta\) | Exp(\(\beta\)) (95\% CI) | \(P\) | \(\beta\) | Exp(\(\beta\)) (95\% CI) | \(P\) |
| Race              |       |         |         |      |         |         |
| Black             | -0.118 | 0.801–0.986 | .025 | NA     | NA     | NA     |
| Others            | -0.199 | 0.699–0.961 | .014 | NA     | NA     | NA     |
| Sex               | -0.126 | 0.630–0.936 | <.001 | -1.22  | 0.834–0.940 | <.001 |
| Age \(\geq 45, < 55\) | 0.063 | 0.792–1.433 | .677 | 0.048  | 0.780–1.411 | .753 |
| Age \(\geq 55, < 65\) | 0.050 | 0.787–1.403 | .736 | 0.041  | 0.781–1.391 | .780 |
| Age \(\geq 65, < 75\) | 0.175 | 0.893–1.589 | .235 | 0.162  | 0.882–1.569 | .270 |
| Age \(\geq 75\) | 0.415 | 1.131–2.028 | .005 | 0.400  | 1.114–1.997 | .007 |
| T                 |       |         |         |      |         |         |
| T1a               | -0.088 | 0.661–1.269 | .598 | -0.079 | 0.667–1.280 | .634 |
| T1b               | 0.073  | 0.781–1.484 | .654 | 0.076  | 0.783–1.487 | .643 |
| T2a               | -0.033 | 0.716–1.308 | .832 | -0.035 | 0.714–1.305 | .820 |
| T2b               | 0.086  | 0.800–1.485 | .586 | 0.080  | 0.795–1.475 | .614 |
| T3                | 0.136  | 0.852–1.541 | .368 | 0.135  | 0.851–1.540 | .371 |
| T4                | 0.121  | 0.841–1.513 | .421 | 0.114  | 0.836–1.503 | .446 |
| N                 |       |         |         |      |         |         |
| N1                | -0.030 | 0.843–1.118 | .680 | -0.023 | 0.848–1.125 | .746 |
| N2                | 0.198  | 1.106–1.340 | <.001 | 0.203  | 1.114–1.347 | <.001 |
| N3                | 0.200  | 1.101–1.357 | <.001 | 0.201  | 1.101–1.358 | <.001 |
| DM                |       |         |         |      |         |         |
| Bone              | 0.298  | 1.202–1.510 | <.001 | 0.299  | 1.203–1.512 | <.001 |
| Liver             | 0.311  | 1.224–1.522 | <.001 | 0.305  | 1.217–1.512 | <.001 |
| Lung              | 0.459  | 1.426–1.756 | <.001 | 0.461  | 1.429–1.760 | <.001 |
| Brain             | 0.039  | 0.907–1.192 | .574 | 0.041  | 0.909–1.195 | .553 |
| Multiorgan        | 0.617  | 1.700–2.021 | <.001 | 0.617  | 1.700–2.022 | <.001 |
| Chemo             | -1.098 | 0.309–0.360 | <.001 | -1.092 | 0.311–0.362 | <.001 |

DM = distant metastatic site.
Figure 4. Nomogram for the different metastatic site in extensive-stage small-cell lung cancer patients.

Figure 5. Calibration plot for nomogram of (A) derivation cohort mod2 and (B) validation cohort.
sites and the number of metastatic sites. Therefore, it is very important to predict the survival time of patients according to the different metastatic sites and the number of metastatic sites. However, there was still a lack of studies on the correlation between different metastatic sites and prognosis. Foster et al[5] found that the number of metastatic tumors was an important prognostic factor, but he ignored the impact of metastatic sites on survival. At present, studies had indicated that the most common metastatic sites of ES-SCLC were liver, brain, bone, lung, and adrenal gland.[10] In this study, 42.8% of the patients were initially diagnosed with liver metastasis, followed by bone metastasis, brain metastasis, and lung metastasis. Survival analysis showed that patients with liver metastasis and multiple organ metastases had the shortest survival period, which was consistent with the report of Cai et al.[4] The decrease of survival expectation caused by liver metastasis may be related to the decrease of liver function caused by liver metastatic tumors, and the occurrence of severe symptoms such as jaundice and ascites, which lead to the decrease of patients’ quality of life.

The significance of the clinical prognostic model was that patients could be grouped according to their survival expectations before treatment, so as to determine whether a specific treatment scheme was worthy of implementation. Chemotherapy, as the most expected survival time, overtreatment should be avoided and treatment, to benefit. Doctors select patients with long expected survival time for active treatment. Relevant research[10,13,16] showed that patients with metastatic tumors, and the occurrence of severe symptoms such as jaundice and ascites, which lead to the decrease of patients’ quality of life.

The prognostic analysis of different metastatic patterns in extensive-stage small-cell lung cancer patients: a large population-based study. Future Oncol 2018;14:1397–407.

5. Conclusion

The survival time of ES-SCLC patients with different metastatic sites was significantly different. The nomogram model can effectively predict the survival of individuals and provide a basis for clinical decision-making.

Author contributions

Conceptualization, SH and HG; Data curation, HG, TQ, and YD; Formal analysis, SH, XH, TQ, and YD; Funding acquisition, XZ and HG; Methodology, HG and SH; Resources, HG; Supervision, XZ and SH; Writing—original draft, SH and HG; Writing—review & editing, SH and HG. All authors read and approved the final manuscript.

References

[1] Ruhimaki M, Hemminki A, Fallah M, et al. Metastatic sites and survival in lung cancer. Lung Cancer 2014;86:78–84.
[2] Jackman DM, Johnson BE. Small-cell lung cancer. Lancet 2005;366:1385–96.
[3] Sundstrom S, Brennes RM, Kaasa S, et al. Cisplatin and etoposide regimen is superior to cyclophosphamide, epirubicin, and vincristine regimen in small-cell lung cancer: results from a randomized phase III trial with 5 years’ follow-up. J Clin Oncol 2001;19:5153–5.153.
[4] Cai H, Wang H, Li Z, et al. The prognostic analysis of different metastatic patterns in extensive-stage small-cell lung cancer patients: a large population-based study. Future Oncol 2018;14:1397–407.
[5] Foster NR, Mandrekar SJ, Schild SE, et al. Prognostic factors differ by tumor stage for small cell lung cancer: a pooled analysis of North Central Cancer Treatment Group trials. Cancer 2009;115:2721–31.
[6] Gao HX, Huang SG, Du JF, et al. Comparison of prognostic indices in NSCLC patients with brain metastases after radiosurgery. Int J Biol Sci 2018;14:2065–72.
[7] Huang S, Yang J, Feng S, et al. Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. Cancer Lett 2020;471:61–71.
[8] Huang S, Yang J, Feng S, et al. Mining prognostic index of brain metastases using artificial intelligence. Cancers 2019;11:1140.
[9] Sculier JP, Changsuy K, Crowley JJ, et al. The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th Edition of the TNM Classification of Malignant Tumors and the proposals for the 7th Edition. J Thorac Oncol 2008;3:457–66.
[10] Hong X, Xu Q, Yang Z, et al. The value of prognostic factors in Chinese patients with small cell lung cancer: a retrospective study of 999 patients. Clin Respir J 2018;12:433–47.
[11] Zhao W, Fang L, Shun-Chang J. Prognostic factors for survival of patients with extensive stage small cell lung cancer—a retrospective single institution analysis. Asian Pac J Cancer Prev 2012;13:4959–62.
[12] Wang S, Yang L, Gi B, et al. Development and validation of a nomogram prognostic model for SCLC patients. J Thorac Oncol 2018;13:1338–48.
[13] Lara PN, Natale R, Crowley J, et al. Phase III trial of irinotecan/cisplatin compared with etoposide/cisplatin in extensive-stage small-cell lung cancer: clinical and pharmacogenomic results from SWOG S0124. J Clin Oncol 2009;27:2530–5.
[14] Nakazawa K, Kurushima K, Tamura T, et al. Specific organ metastases and survival in small cell lung cancer. Oncol Lett 2012;4:617–20.
[15] Spigel DR, Townley PM, Waterhouse DM, et al. Randomized phase II study of bevacizumab in combination with chemotherapy in previously untreated extensive-stage small-cell lung cancer: results from the SALUTE trial. J Clin Oncol 2011;29:2215–22.
[16] Nasser H, Bunn PA, Langer C, et al. Randomized phase III trial comparing irinotecan/cisplatin with etoposide/cisplatin in patients with previously untreated extensive-stage disease small-cell lung cancer. J Clin Oncol 2006;24:2038–43.
[17] Slotman BJ, van Tinteren H, Praag JO, et al. Use of thoracic radiotherapy for extensive stage small-cell lung cancer: a phase 3 randomised controlled trial. Lancet 2015;385:36–42.
[18] Slotman BJ, Faire-Finn C, van Tinteren H, et al. Which patients with ES-SCLC are most likely to benefit from more aggressive radiotherapy? a secondary analysis of the Phase III CREST trial. Lung Cancer 2017;108:150–3.