Prognostic Factors in Patients with Brain Metastases of Lung
Adenocarcinoma with the Surveillance Epidemiology and End Results
Database

Rongjiong Zheng¹, Dina Guo¹, Yongquan Dong¹, Mingming Wang¹, Meiling Hu²*, Yifeng
Ren¹*

¹Department of Pulmonology, Ningbo Yinzhou Second Hospital, Ningbo 315192, China

²Cixi People’s Hospital of Zhejiang Province, Zhejiang, China.

Running title: Prognostic factors in patients with brain metastases of lung
adenocarcinoma

Email addresses:

RJZ: 842429171@qq.com  MLH: humeiling826@163.com  MMW: wmmrespiratory@163.com

DNG: 156980386@qq.com  YFR: 84479482@qq.com  YQD: 12919819@qq.com

*Correspondence to: (The following authors are equally)

Meiling Hu, MD, Cixi People’s Hospital of Zhejiang Province, Zhejiang, China. E-mail address:
humeiling826@163.com; and Yifeng Ren, MD, Ningbo Yinzhou Second Hospital, Ningbo 1
Qianhe Road, Ningbo 315192, China. E-mail address: 84479482@qq.com
Abstract

Background: The study aimed to distinguish the risk factors associated with overall survival (OS) of the patients suffering from lung adenocarcinoma (LACA) with brain metastases, and built a prognostic tool (nomogram) for these patients.

Methods: LACA patients with brain metastases between 2010 and 2013 were selected from the incidents collecting database i.e., SEER (Surveillance, Epidemiology, and End Results). Kaplan-Meier method along with the Cox regression model was used to assess the prognostic effect of each variable on the rate of survival. The nomogram was developed for the prediction of a 3-, 6- and 9-month OS rate.

Results: About 2631 LACA patients with brain metastases were included in this study. A nomogram was developed according to those variables that were considerably affecting the OS, followed by validating via internal bootstrap resampling method which revealed that the nomogram exhibited an appropriate power of discrimination.

Conclusions: The nomogram was able to predict a 3-, 6- and 9-month OS rate of patients suffering from LACA with brain metastases.

Keywords: nomogram, lung adenocarcinoma, brain metastases, hazards model, SEER database
Introduction

Lung carcinoma is considered to be the most prevalent carcinoma across the globe. Because of their correlation with serious health problems, lung carcinoma has been gained great attention in recent years. In 2012, around 1.8 million people were diagnosed with lung carcinoma with 1.6 million deaths worldwide. Non-small cell carcinoma (NSCLC) is the most prevalent cancer and among lung carcinoma, approximately 85% of patients diagnosed with NSCLC, particularly the lung adenocarcinoma (LACA, a histologic subtype of NSCLC). As a common advanced phenomenon in malignancies, brain metastasis takes up about 20-40% of the cancers. Furthermore, intracranial involvement accounts for about 40-60% of the LACA patients, and approximately 10% are diagnosed at the first time. Considering the severity of the disease, an available and feasible grading system is needed to stratify the prognosis of the patients with brain metastases of LACA.

In recent years, a nomogram, which is a graphical predictive tool, has been extensively applied to predict the prognosis of some diseases. It can estimate the survival rate of individual patients more accurately by integrating some important prognostic variables. However, the imaging application of LACA patients with brain metastases based on crowd data is still rare so far. Thus, we have used data from the incidents collecting database i.e., SEER (Surveillance, Epidemiology, and End Results) to recognize the OS associated risk factors and to build a nomogram for the forecasting of the patients suffering from LACA with brain metastases.
Materials and Methods

Data Source

The data on cancer prevalence, incidence, mortality, and treatment was from the SEER database, which covered about 30% of the US population across 17 geographical regions in the United States. The information about patients with brain metastases of LACA in this study was retrieved from the SEER database via SEER*Stat software (version 8.3.5). The Ethics Committee of the Institutional Review Board of Ningbo No.2 Hospital provided approval for the existing study and consent form. The existing study was conducted based on recommended suggestions of the Declaration of Helsinki.

A total of 2631 patients with brain metastases of LACA were included from January 2010 to December 2013. The exclusion standards have been given as: (1) the diagnosis age was less than 18 years old; (2) participants diagnosed only from autopsy or a death certification; (3) participants with more than one primary carcinoma and had gotten LACA before; (4) participants with omitted or deficient data associated with the rate of survival, grade, follow-up months and the main reason of mortality.

Variables

A series of clinical and demographic variables were extracted from the SEER program in this study, including race, sex, age, year of diagnosis, grade, surgery history, radiotherapy history, chemotherapy history, cause of death, and follow-up information. The classifications of the patients were as follows: Patients were divided into two groups in accordance with the diagnostic age, including people who were younger than 60 and older than 60. Races were divided into black, white, and others. The primary endpoint was OS. The former was characterized as the time interval from
LACA diagnosis to its censor. The latter was characterized as the time interval from diagnosis of LACA to death or last follow-up with no restriction to the causes of mortality.

Statistical analysis

Categorical variables in this study were displayed in frequency and scale, followed by the comparison with a Fisher’s exact or Chi-squared test. The log-rank test was employed for the evaluation of the variations in survival among subgroups. Cox analyses were conducted to predict 95% confidence intervals (CIs) and hazard ratios (HRs). 25.0 version of SPSS (Chicago, IL, USA) and R-3.5.1 (http://www.r-project.org/) were employed for statistical analysis. Moreover, the nomogram for predicting OS was performed based on the Cox regression and the R packages rms and mstate were also applied to frame the model. The C-index for OS in the nomogram model measured the variations in prognostic power between observed and predicted data that were used to analyze the nomogram discrimination. A larger C-index showed a relatively more capability of patients with various outcomes of survival rate. The calibration plot and C-index were obtained based on the regression analysis. A calibration curve (along the 45-degree line) would show an appropriate calibration model having a high level of similarity in predicted probabilities with the actual results. All the p-values were two-sided. P-values lower than 0.05 were regarded as statistically considerable.

Results

Patient characteristics

In accordance with the standards of patient selection, a total of 2631 eligible patients diagnosed with brain metastases of LACA between 2010 and 2013 were included. Table 1 showed the demographic
and clinical features of the patients. Among these patients, 49.7% were male, and most of the patients were white (75.9%), and older than 60 years (61.0%). The increasingly common grades were found to be badly differentiated, moderately differentiated, and well-differentiated i.e., 64.5%, 29.5%, and 4.7% accordingly.

From the perspective of treatment, patients were treated with surgery (7.1%), radiotherapy (17.2%), and chemotherapy (62.0%) before. By the end of follow-up, 2303 (87.5%) people in this study had died, including 2200 (95.5%) people who died of brain metastases of LACA and 103 (4.5%) people died of other reasons.

**Survival analysis**

In this study, all 2631 patients were involved in cox regression analyses to investigate the predictors of survival. As depicted in Figure 1, survival outcome differed with age, sex, and race. Also, the grade was the risk factor affecting survival. As shown in Figure 1 e, f and g, surgery history, radiotherapy history, and chemotherapy history also remarkably affected the prognosis of patients.

Furthermore, we used cox regression analyses to extensively explore the influence of race, sex, age, grade, chemotherapy history, surgery history, and radiotherapy history (Table 2). By univariate analyses, each variable was linked with OS. The independent prognostic variables were surgery history, radiotherapy history, chemotherapy history, age, sex, race, and grade ($p < 0.001$). After adjustment of other risk factors, the multivariate cox regression model via step by step selection method determined the same variables as univariate analyses as the independent predictors ($p < 0.01$).

Figure 2 depicted the developed nomogram based on significant risk factors determined via multivariate analysis using OS for 3-, 6- and 9-month. For the measurement of the 3-, 6- and 9-
month OS rate, each factor was recognized based on total points at the top scale of the nomogram, and the total points were summed. Eventually, the 3-, 6- and 9-month OS rates were achieved in view of the point scale at the bottom of the nomogram. Models showed significant precision with C-index of 0.697 (95% CI = 0.685-0.709) for the OS model, which indicated a relatively good model discriminating ability for predicting the 3-, 6- and 9-month OS rate of the brain metastases of the LACA patients. The calibration curves based on bootstrap resampling validation were well standardized, where the points were close to a 45° (Figure 3).

Discussion

In the present study, we use 2631 cases from the incidents collecting database i.e., SEER for the establishment of a prognostic nomogram to predict OS incidence in 3, 6, and 9 months of the LACA patients with brain metastases. In our cox model, age, sex, race, grade, surgery history, radiotherapy history, and chemotherapy history are all independent prognostic factors. The internal Bootstrap resampling method is adopted to analyze the performance and discriminant features of the nomogram. And the result shows that the nomogram fits well with the actual observation for the prediction of the 3-, 6- and 9-month OS rate according to the C-index.

The brain metastases of LACA remain a huge public health burden in the world because of the poor prognosis, high incidence, and recurrence; however, there is no unified and efficient survival prediction model internationally. Hence, there is a compelling need for effective approaches to determine high-risk people with bad survival. Here, we established a nomogram that defined significant prognostic factors, which were readily accessible in everyday clinical practice. The nomogram comprised of some independent prognostic factors (from clinical practice). Sex, age, and
race were key prognostic variables on OS in numerous studies\textsuperscript{20-22}. The existing study also revealed that older and male patients had poorer survival, and married patients had a better prognosis, possibly resulting from more emotional comfort. Furthermore, grade, surgery history, radiotherapy history, and chemotherapy history were also the independent factors for survival prediction in comparable reports\textsuperscript{23-24} which were found to be persistent with our obtained results.

The nomogram provides a very basic graphical representation of complex models for the quantification of the personal risk factors, showing a comprehensive implementation in clinical practice and research\textsuperscript{25}. The prediction of survival rate has been made easy and simple by using the nomogram. Firstly, from each clinical variable, a vertical line is drawn to the “points” line in the nomogram. Secondly, total “points” are added to obtain “total points”, and from “total points” a vertical line is drawn to the “OS” lines to attained corresponding survival. For instance, consider a 61-year-old (3.1 points) male (1.8 points) with grade I (0.4 points) and chemotherapy history (0 points). Using the nomograms, the 3-, 6- and 9-month OS are 79\%, 65\%, and 54\%, respectively. Therefore, the nomogram in this study would enable clinicians to determine patients with an elevated level of risk and low survival rate before treatment and to make better clinical decisions and follow-up checks for patients.

The main advantages of this study are the detailed and rich samples and the simplicity of models, confirming the establishment of a perfect predictive model. All 2631 suitable patients were included in the existing study and statistical data in the underlined samples from population-based cancer registries (PBCR) were more general and authentic as compared to those from single-center studies. Furthermore, variables in the models can be acquired easily and the predictive hazard assessment for the patients suffering from LACA with brain metastases was more complete. Our
nomogram displays a high capability for prediction of OS, and the presentation of the nomogram was also confirmed by calibration. As in the case of earlier reported studies that are based on different databases of SEER, our study also has some shortcomings in investigation via a large population-based dataset. First, potential selection bias may be appeared due to the retrospective review from the SEER database. Second, the SEER dataset does not give certain effective clinical parameters correlated with prognosis i.e., vascular invasion, driver gene status, and KPS (Karnofsky performance status); future studies should contain the underlined significant factors. Third, the predictive value calculated from the nomograms is only for clinicians' reference as all prognostic factors are not present in the nomograms. In view of this fact, it is not all time possible to have precise prognosis in clinical practices.

Conclusions

By summing up the main points, we develop nomograms for forecasting a 3-, 6- and 9-month OS cohort study based on an enormous population, which show a good prognosis in LACA patients with brain metastases.

Consent for publication

Not applicable
Data Availability

All data are fully available without restriction.

Conflicts of Interest

The authors have nothing to disclose.

Funding

None

Acknowledgements

We appreciate all participants who took part in our study.

Authors’ contributions

RJZ and DNG carried out the study analysis and interpretation of data, and drafted the manuscript.

MMW, YFR, YQD and MLH participated in the analysis and interpretation of data, conceived the study, and helped in drafting the manuscript. All authors read and approved the final manuscript.

References

1. Locher C, Debieuvre D, Coëtmeur D, et al. Major changes in lung cancer over the last ten years
in France: the KBP-CPHG studies. Lung Cancer. 2013;81(1):32-8. (PMID: 23541463)

2. Jemal A, Miller KD, Ma J, et al. Higher Lung Cancer Incidence in Young Women Than Young Men in the United States. N Engl J Med. 2018; 24;378(21):1999-2009. (PMID: 29791813)

3. Malvezzi M, Carioli G, Bertuccio P, et al. European cancer mortality predictions for the year 2017, with focus on lung cancer. Ann Oncol. 2017;1;28(5):1117-1123. (PMID: 28327906)

4. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J. Clin. 2015;65(2):87-108. (PMID: 25651787)

5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin. 2015;65(1):5–29. (PMID:25559415)

6. Ettinger DS, Akerley W, Borghaei H, et al. National comprehensive cancer network. Non-small cell lung cancer, version 2. 2013. J Natl Compr Cancer Netw. 2013;11:645–53. (PMID: 23744864)

7. Lemjabbar-Alaoui H, Hassan OU, Yang YW, et al. Lung cancer: biology and treatment options. Biochim Biophys Acta. 2015;1856:189–210. (PMID: 26297204)

8. Nayak L, Lee EQ, Wen PY. Epidemiology of brain metastases. Curr Oncol Rep. 2012;14(1):48-54. (PMID: 22012633)

9. Schouten LJ, Rutten J, Huveneers HA, et al. Incidence of brain metastases in a cohort of patients with carcinoma of the breast, colon, kidney, and lung and melanoma. Cancer 94:2698–2705. (PMID: 12173339)

10. Ding X, Dai H, Hui Z, et al. Risk factors of brain metastases in completely resected pathological stageIII-A-N2non-smallcell lung cancer. Radiat Oncol. 2012;30;7:119. (PMID: 22846375)

11. Fang C, Wang W, Feng X, et al. Nomogram individually predicts the overall survival of patients
with gastroenteropancreatic neuroendocrine neoplasms. Br J Cancer 2017;117:1544e50. (PMID: 28949958)

12. Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008;26(8):1364–70. (PMID: 18323559)

13. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population, Med Care. 2002;40(8 Suppl):IV-3-18. (PMID: 12187163)

14. Surveillance Research Program, National Cancer Institute SEER*Stat software (www.seer.cancer.gov/seerstat) version 8.3.5.

15. Rongjiong Zheng, Xiaolong Gu, Mingming Wang, et al. Nomograms to Predict Survival in Patients with Lung Squamous Cell Cancer: A Population-Based Study, J Nippon Med Sch 2019;86(6):336–44.

16. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.2014.

17. Frank E, Harrell Jr. rms: Regression Modeling Strategies. R package version 3.4–0. http://cran.r-project/package=rms.

18. Wolbers M, Koller MT, Witteman JC, et al. Prognostic models with competing risks: methods and application to coronary risk prediction. Epidemiology. 2009;20(4):555–61. (PMID: 19367167)

19. Huitzil-Melendez FD, Capanu M, O'Reilly EM, et al. Advanced hepatocellular carcinoma: which staging systems best predict prognosis? J Clin Oncol. 2010;10;28(17):2889-95. (PMID: 20458042)
20. Wray CJ, Phatak UR, Robinson EK, et al. The effect of age on race-related breast cancer survival disparities. Ann Surg Oncol. 2013;20:2541–2547. (PMID: 23435633)

21. Oweira H, Petrausch U, Helbling D, et al. Prognostic value of site-specific metastases in pancreatic adenocarcinoma: a Surveillance Epidemiology and End Results database analysis. World J Gastroenterol WJG. 2017;23:1872–1880. (PMID: 28348494)

22. Martínez ME, Anderson K, Murphy JD, et al. Differences in marital status and mortality by race/ethnicity and nativity among California cancer patients. Cancer. 2016;122:1570–1578. (PMID: 27065455)

23. Fang C, Wang W, Feng X, et al. Nomogram individually predicts the overall survival of patients with gastroenteropancreatic neuroendocrine neoplasms. Br J Cancer. 2017;117:1544–1550. (PMID: 28949958)

24. Deng W, Wang Q, Xiao Z, et al. A prognostic nomogram for overall survival after neoadjuvant radiotherapy or chemoradiotherapy in thoracic esophageal squamous cell carcinoma: a retrospective analysis. Oncotarget. 2017;20;8(25):41102-41112. (PMID: 28456788)

25. Bianco FJ Jr. Nomograms and medicine. Eur Urol. 2006; 50:884–886. (PMID: 16973258)
Figure 1 Overall Kaplan-Meier survival curves for patients according to (a) age, (b) sex, (c) race, (d) grade, (e) radiotherapy history, (f) surgery history, (g) chemotherapy history.
Figure 2 Nomogram for predicting 3-, 6- and 9-month overall survival (OS) of LACA patients with brain metastases.
**Figure 3** Calibration plots of the nomogram for 3-, 6- and 9-month OS prediction (a, b, c). The X-axis represents the nomogram-predicted probability of survival; the Y-axis represents the actual OS probability.
Table 1. Baseline characteristics of the patients

| Variables           | All patients (n = 2631) |
|---------------------|-------------------------|
|                     | No. | %       |
| Sex                 |     |         |
| Female              | 1324 | 50.3    |
| Male                | 1307 | 49.7    |
| Age                 |     |         |
| ≤ 60                | 1027 | 39.0    |
| > 60                | 1604 | 61.0    |
| Race                |     |         |
| White               | 1997 | 75.9    |
| Black               | 335  | 12.7    |
| Other a             | 299  | 11.4    |
| Grade               |     |         |
| I                   | 123  | 4.7     |
| II                  | 776  | 29.5    |
| III                 | 1696 | 64.5    |
| IV                  | 36   | 1.4     |
| Surgery history     |     |         |
| No                  | 2444 | 92.9    |
| Yes                 | 187  | 7.1     |
| Radiotherapy history|     |         |
| No                  | 2178 | 82.8    |
| Yes                 | 453  | 17.2    |
| Chemotherapy history|     |         |
| No                  | 1001 | 38.1    |
| Yes                 | 1630 | 62.0    |

a Other includes American Indian/AK Native, Asian/Pacific Islander, and unknown.
| Variables            | Univariate |         |         | Multivariate |         |         |
|----------------------|------------|---------|---------|--------------|---------|---------|
|                      | HR (95%CI) | p value | HR (95%CI) | p value      | HR (95%CI) | p value |
| Sex                  |            |         |         |              |         |         |
| Female               | 1.00       | < 0.001 | 1.00    | = 0.002      | 1.00    |         |
| Male                 | 1.20(1.11-1.31) |         | 1.14(1.05-1.24) |         |         |
| Age                  |            |         |         |              |         |         |
| < 60                 | 1.00       | < 0.001 | 1.00    | < 0.001      | 1.00    | < 0.001 |
| ≥ 60                 | 2.16(1.65-2.83) |         | 1.27(1.16-1.38) |         |         |
| Race                 |            |         |         |              |         |         |
| White                | 1.00       |         | 1.00    | < 0.001      | 1.00    | < 0.001 |
| Black                | 0.99(0.88-1.12) |         | 0.94(0.83-1.06) |         |         |
| Other*               | 0.66(0.58-0.76) |         | 0.64(0.56-0.73) |         |         |
| Grade                |            | < 0.001 |         | < 0.001      |         |         |
| I                    | 1.00       |         | 1.00    |              | 1.00    | < 0.001 |
| II                   | 0.91(0.74-1.11) |         | 0.97(0.79-1.19) |         |         |
| III                  | 1.17(0.96-1.42) |         | 1.27(1.05-1.55) |         |         |
| IV                   | 1.27(0.86-1.86) |         | 1.33(0.90-1.95) |         |         |
| Surgery history      |            |         |         |              |         |         |
| No                   | 1.00       | < 0.001 | 1.00    | < 0.001      | 1.00    | < 0.001 |
| Yes                  | 0.40(0.33-0.48) |         | 0.45(0.37-0.55) |         |         |
| Radiotherapy history |            |         |         |              |         |         |
| No                   | 1.00       | < 0.001 | 1.00    | < 0.001      | 1.00    | < 0.001 |
| Yes                  | 0.56(0.50-0.63) |         | 0.74(0.65-0.84) |         |         |
| Chemotherapy history |            |         |         |              |         |         |
| No                   | 1.00       | < 0.001 | 1.00    | < 0.001      | 1.00    | < 0.001 |
| Yes                  | 0.46(0.42-0.50) |         | 0.46(0.42-0.50) |         |         |

* Other includes American Indian/AK Native, Asian/Pacific Islander, and unknown.