Prognostic Factors and Diseases-Specific Survival Outcome in Patients with Glioblastoma: A Population-Based Study

Lihua Wu  
Cancer Center of Shanxi Bethune Hospital, Shanxi Academy of Medical Science

Jianbo Song  
Cancer Center of Shanxi Bethune Hospital, Shanxi Academy of Medical Science

Junping Zhang  
Cancer Center of Shanxi Bethune Hospital, Shanxi Academy of Medical Science

Wenhui Yang  
Cancer Center of Shanxi Bethune Hospital, Shanxi Academy of Medical Science

Mengxian Zhang  
Cancer Center of Shanxi Bethune Hospital, Shanxi Academy of Medical Science

Jing Wei (✉ weijing3024@163.com)  
Cancer Center of Shanxi Bethune Hospital, Shanxi Academy of Medical Science  
https://orcid.org/0000-0001-5834-8403

Research article

Keywords: Glioblastoma, Disease-specific survival, Lasso regression, Restricted cubic spline cox regression, Nomogram

DOI: https://doi.org/10.21203/rs.3.rs-111364/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Objective

This study aimed to determine the prognostic factors for disease-specific survival (DSS) of glioblastoma (GBM) and establish a corresponding effective nomogram for clinical prediction.

Methods

This study was based on Surveillance, Epidemiology, and End Results database between 2004 and 2015. Kaplan-Meier survival analysis was used to evaluate the effect of various prognostic factors on DSS. Lasso regression was used to determine the independent prognostic factors of DSS and multivariate cox regression analysis was performed correspondingly. Additional restricted cubic spline cox regression was used to analyze the trend of the risk effect (hazard ratio) of continuous variables on DSS. Based on the multivariate cox regression model, a nomogram was established to predict DSS.

Results

The average age at diagnosis of all enrolled patients was 59.8±12.2 years, of which 40.5% were women and 59.5% were men. Lasso regression analysis showed that age at diagnosis, sex, marital status, race, tumor size, primary site, laterality, surgery, radiotherapy and chemotherapy, radiotherapy sequence with surgery, and year of diagnosis were independent prognostic factors for DSS. Multivariate cox regression analysis showed that elderly, males, unmarried status, larger tumors were all risk factors for DSS. Restricted cubic spline cox regression showed that the risk of death from GBM was significantly increased for the elderly, especially older than 75 years. When the tumor was smaller than 75mm, an increasing risk linearly was associated with DSS, but the risk effect remained constant after 75mm. Constructing the nomogram to predict the DSS probability of 1-, 3- and 5-year respectively, and its good predictive performance was proved by the calibration curve.

Conclusion

The advanced age was one of the significant risk factors for GBM. How the change of tumor size affected DSS needed further study and discussion. The established nomogram was robust in predicting 1-, 3-, and 5-year DSS.

Introduction

Glioblastoma (GBM) associated with particularly invasive activities and high lethality is the main histological subtype of glioma, accounting for 48.3% of malignant brain tumors in the United States, and the five-year relative survival rate for patients is merely 6.8 percent[1–3]. Female patients have a significant survival preponderance over males[4]. The predominant hallmarks of GBM are highly infiltrating and often affecting the surrounding brain parenchyma, but the tumor is usually limited to the central nervous system and does not metastasize[5, 6]. Recognized risk factor for GBM is exposure to
ionizing radiation, and it can be potentially altered\textsuperscript{[7]}. Additionally, several studies have explored the factors that affect the overall survival and prognosis of GBM, including age at diagnosis, sex, surgery, radiotherapy and chemotherapy, etc\textsuperscript{[8–11]}. Prognostic factors to predict diseases-specific survival (DSS) of patients are generally less studied. Therefore, we attempted to evaluate the predictive factors for DSS of GBM patients and develop a clinical prediction nomogram.

The morbidity of GBM is relatively high about 3.22/100000 in the United States\textsuperscript{[12, 1]}. We used the SEER-18 database, which covers approximately 28\% of the population in the United States, to analyze the significant prognostic factors related to the DSS of GBM. In the construction of the multivariable cox regression model, different from using traditional univariate cox regression, we used least absolute shrinkage and selection operator (Lasso) regression\textsuperscript{[13]}, a regularized regression calculation method, to select appropriate variables, which could optimize the bias-variance trade-off and improve the fitness of the regression model. And the regularization of coefficients in Lasso regression can be used to solve the over-fitting problem caused by multicollinearity. By regressing and penalizing all variable coefficients to make the relatively unimportant independent variable coefficients 0, main characteristic variables were selected.

In addition, continuous independent variables in cox regression model are usually classified, but the number of classifications and classification thresholds are often subjective, which can cause bias in research results. So, restricted cubic spline cox regression, one of the commonly used methods to analyses the nonlinear relationship between continuous variables and DSS, was used in our study. This provides support for in-depth understanding of prognostic factors on DSS.

**Materials And Methods**

**Data source and subjects**

Our study was based on the public access database (SEER) with corresponding information from the reports of 18 cancer registries in the United States, using SEER*Stat 8.3.8 to extract the GBM cases diagnosed in 2004 and later. The third edition of International Classification of Disease for Oncology (ICD-O-3) was used to confirm the histological type (9440/3) limited to “Brain and Other Nervous System” in SEER database. All selected cases were confirmed by microscopically pathological diagnosis.

Related diagnostic exclusion criteria were established according to the following variables for GBM in adults: (1) One primary tumor only (2) Age at diagnosis $> 18$ (3) IV grade (4) Tumor size recorded with accurate numbers (1-989 millimeters), excluding uncertain records (5) Primary sites: Frontal lobe(C71.1), Temporal lobe(C71.2), Parietal lobe(C71.3), Occipital lobe(C71.4), Overlapping lesion of brain(C71.8) (6) Laterality: Left-origin of primary, Right-origin of primary, and Not a paried site (7) According to the treatment information of radiotherapy and chemotherapy, we generated a new variable “Radiation and Chemotherapy”: Both, and Radiation only.
All variables analyzed in this study including age at diagnosis, sex, race, marital status, tumor size, primary site, laterality, surgery, radiotherapy and chemotherapy, and radiotherapy sequence with surgery, eliminating unknown or missing information. A total of 4310 cases were incorporated in the analysis (Table 1).

Statistics and Analysis

For better observing the long-term trend of age-adjusted incidence of GBM over time, the Supplementary Fig. 1 based on SEER-18(2000–2017) were made. Kaplan-Meier survival analysis were used to determine the influence of each research variable on DSS, and comparison were performed by Log-rank test. Lasso regression was used to screen the characteristic variables that affected the prognosis of GBM, and then multivariate analysis was performed based on the cox regression model. The associations between continuous variables and DSS were flexibility modeled by restricted cubic splines cox regression, adjusted by other covariates in cox regression, to explore the potential non-linear correlation. The multivariate cox regression model was represented as a nomogram to predict DSS for clinical application, and calibration curves of predicting 1-, 3-, and 5-year DSS were plotted to assess the predictive accuracy. All statistical analysis was conducted using RStudio (Version 4.0.2)

Results

Characteristics of the study subjects

Based on the SEER database, we plotted the age-adjusted incidence of GBM patients from 2000 to 2017 (Supplementary Fig. 1). The incidence rate of GBM showed ubiquitous steady high trends for total population, males only and females only. Taken together, incidence rate of GBM was significantly higher in males than in females.

Our study included 4310 patients of GBM from 2004 to 2015, and the demographic characteristics for all subjects are summarized in Table 1. The average age at diagnosis for all cases in study was 59.8 (standard deviation: 12.2), and the median age was 75 (quartile: 52, 68). After converting the continuous variable “age at diagnosis” into a categorical variable, the number of patients in the 40–65 stage was the largest, accounting for 58% cases. There were more males than females for all subjects: 59.5% male. Most of the patients were white (91.2%), and the proportions of black (5.1%) and others (3.7%) were relatively low. Details for other prognostic factors can be found in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Characteristics of the study subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on the SEER database, we plotted the age-adjusted incidence of GBM patients from 2000 to 2017 (Supplementary Fig. 1). The incidence rate of GBM showed ubiquitous steady high trends for total population, males only and females only. Taken together, incidence rate of GBM was significantly higher in males than in females.</td>
</tr>
<tr>
<td>Our study included 4310 patients of GBM from 2004 to 2015, and the demographic characteristics for all subjects are summarized in Table 1. The average age at diagnosis for all cases in study was 59.8 (standard deviation: 12.2), and the median age was 75 (quartile: 52, 68). After converting the continuous variable “age at diagnosis” into a categorical variable, the number of patients in the 40–65 stage was the largest, accounting for 58% cases. There were more males than females for all subjects: 59.5% male. Most of the patients were white (91.2%), and the proportions of black (5.1%) and others (3.7%) were relatively low. Details for other prognostic factors can be found in Table 1.</td>
</tr>
</tbody>
</table>
CI, confidence interval; SD, standard deviation; IQR, interquartile range; Statistically significant p values are bolded; Red bars indicate risk factors, green bars indicate protective factors, and blue bars indicate no statistical significance.

### Survival analysis

The cases included in this study had a total of 3490 deaths due to GBM at the end of follow-up. Supplementary Table 1 shows the complete 1-, 3-, and 5-year DSS list. DSS was strongly age-dependent and the elderly populations (> 65 years) had a poorer prognosis compared to younger populations (< 40 years, 40–65 years). Likewise, males and white were more likely to have a poor DSS, respectively. Married patients had an advantage in short-term survival than unmarried patients, but it was opposite in long-term survival. The group with tumor size less than 30 mm was significantly better in 1-year DSS (64.1%) than the groups with 30–50 mm (55.9%) and > 50 mm (52.4%). The highest 3-year DSS of the tumor located in occipital lobe compared to those located in other primary site was observed and the highest 5-
year DSS was observed in the tumor located in frontal lobe. Left-origin of primary laterality, and diagnosed after 2010, were all related to a better DSS prognosis.

In terms of treatment methods, patients who received surgery had a more obvious advantage in DSS than patients who did not; Patients who received radiotherapy alone had a worse prognosis than patients who received a combination treatment with radiotherapy and chemotherapy; In addition, the sequence of radiotherapy with surgery did not show a significant effect on DSS. The above results were showed pictorially by Fig. 1 and Supplementary Fig. 2–3.

**Variable selection**

In Fig. 2(A), the ordinate represented the coefficient value, and the upper abscissa represented the number of non-zero coefficient in the model. This figure showed that as the lambda decreasing, the parameters compressed were decreased and the absolute values of the coefficients were increased. Figure 2(B) used cross validation to select variables and then fit the model. The default ten-fold cross validation method was to divide the data into ten parts, and nine parts were used to build the model, the other part was used to validate. By establishing and verifying the model, the results obtained were more stable. Overall, Figs. 2(A) and (B) showed the changes in the calculation operation of the model with the change of lambda.

Finally, we selected the best lambda value (lambda.min) and output the corresponding coefficient values of different variables (Supplementary Table 2). If the variable coefficient was not 0, it indicated that it was meaningful for predicting DSS. Lasso regression results showed that age at diagnosis, sex, marital status, race, tumor size, primary site, laterality, surgery, radiotherapy and chemotherapy, radiotherapy sequence with surgery, and year of diagnosis all showed significance for DSS (coefficient not 0).

**Multivariate cox regression model**

The cox regression model is a semi-parametric model, and uses survival outcome and survival time as dependent variables to analyze the impact of multiple prognostic factors on survival. According to the Lasso regression results, factors that may affect the prognosis of GBM were selected into the cox regression model for multivariate analysis (Table 1). The results showed that age at diagnosis, sex, marital status, race, tumor size, surgery, radiotherapy and chemotherapy, and year of diagnosis were all significantly correlated with DSS ($P < 0.05$). Individually, taking the < 40 years old group as a reference, the group of 40–65 years ($HR = 2.16, 95\% CI = 1.84–2.54, p < 0.001$) and ≥ 65 years ($HR = 3.61, 95\% CI = 3.05–4.27, p < 0.001$) showed significant increases in risk. Compared with white, the risks of others (Asian or Pacific Islander) were reduced by 20%. The larger the tumor size was, the higher was the risk of GBM (30–50mm: $HR = 1.17, 95\% CI = 1.06–1.31, p = 0.003$; >50mm: $HR = 1.27, 95\% CI = 1.15–1.41, p < 0.001$). Patients who did not choose surgery were more likely to have an increased risk for DSS than those who chose ($HR = 1.77, 95\% CI = 1.00–3.12, p = 0.050$). Males, unmarried status, and choosing radiotherapy only were also shown to be risk factors.
In addition, tumor located in the temporal lobe, later year of diagnosis indicated certain protective effects for DSS. No other prognostic factors were found to be correlated significantly with DSS ($p > 0.05$).

**Restricted cubic spline cox regression**

In the cox regression model, the continuous variables age at diagnosis and tumor size were classified to convert into categorical variables and included in the analysis, which was impossible to actually observe the continuous risk variation trend. Artificially transforming continuous variables into categorical variables may cause bias, because there were difficulties in realizing the association between continuous variables changing and DSS. In this study, we used restricted cubic spline cox regression to better observe the effect of continuous changes with age and tumor size on DSS. Results showed that taking the minimum age of 19 years as a reference for the included population, the risk effect value (hazard ratio) gradually increased as the age increasing and had a faster growth trend in the elderly population (> 75 years), suggesting that advanced age was an significant risk factor for DSS (Fig. 3(A)). When 30 mm of tumor size was used as a reference, the value of hazard ratio rapidly increased below 75 mm, and after that it remained a rather stable risk trend (Fig. 3(B)).

The non-linear associations between continuous variables (age, tumor size) and DSS were statistically significant ($p < 0.0001$), indicating that the application of restricted cubic spline regression was necessary. Additionally, it still maintained similar trends in age and tumor size categorted by other variables (Supplement Fig. 4–7).

**Nomogram**

The nomogram with multivariable cox regression model was constructed by scoring independent prognostic factors and it could provide guidance for clinical prediction (Fig. 4 and Supplementary table 3). New samples generated by bootstrap self-sampling were used to evaluate the accuracy of the nomogram. The horizontal axis of the calibration chart was the predicted survival rate, and the vertical axis was the actual survival rate. In theory, the standard curve is a straight line that passes through the origin of the coordinate axis and its slope is 1. If the predicted calibration curve is closer to the standard curve, the better is the predictive ability of the nomogram. Supplementary Fig. 8 showed that the 1-, 3-, and 5-year predicted probabilities of the nomogram constructed based on the variables selected by Lasso regression fit better with the true probability. Moreover, we selected variables based on the traditional univariate cox regression analysis method (Supplementary table 4), and then constructed a nomogram (Supplementary Fig. 9), which showed that the prediction probability level was poor (Supplementary Fig. 10). Therefore, the choice of disease prognostic factors should not be limited to use traditional analysis methods.

**Discussion**

This study analyzed the prognostic factors related to DSS in the GBM patients from 18 cancer registries in SEER. Despite our data only covers part of the U.S. population, it was necessary to study the effects of
its prognostic factors and DSS outcomes for rapid disease progression, poor prognosis, and high incidence level of GBM.

Results from the multivariable cox regression showed the cases diagnosed in 2011–2015 had better DSS than those diagnosed in 2004–2010, which may benefit from the continuous improvement of medical treatments over time. Such as, radiotherapy plus temozolomide chemotherapy to GBM replaced radiotherapy alone from 2005[14–16]. And this protocol was supported by a randomized Phase 3 study[17]: adding temozolomide could extend the median survival time to 15 months, while radiotherapy alone could increase the median survival to 12 months ($HR = 0.63, P < 0.001$). Then, bevacizumab was used as a rescue treatment for GBM in the United States in 2009. Bevacizumab is a humanized monoclonal antibody generally being recommended for those patients in severe symptoms and can reduce brain edema and tumor shrinkage[18]. In addition, tumor treatment fields are being tested as a novel anti-mitotic therapy that interferes with cell division by delivering low-intensity alternating electric fields to the tumor for GBM patients[19]. Extensive efforts are needed to explore new treatments in areas such as immunotherapy and precision oncology. However, compared with other solid tumors, biological factors such as the blood-brain barrier, unique tumors and immune microenvironment pose major challenges in developing new therapies[20].

Advanced age was an increased significantly risk factor for DSS of GBM, particularly if patients aged more than 75 years. Studies have found that T cell dysfunction is prone to occur in GBM, and the effective immunity in the brain tumor microenvironment is suppressed, which easily leads to poor prognosis[21]. However, the increasing age will accelerate the senescence of T cells[22, 23], and the human mRNA expression factors related to immunosuppression in the brain will increase[24], which may reflect one of the possible reasons for the increased risk of the elderly. It was suggested that elderly patients, especially those over 75 years, should be considered their immune status and carefully cared to minimize the risk of death and improve the quality of life[25].

In the restricted cubic spline regression analysis, we found that for tumors smaller than 75 mm, the risk of DSS increased with the increasing of tumor size, but the risk value (HR) remained at a stable level and no longer increased from 75 mm. The size of the tumor was demonstrated that having an impact on survival prognosis[26–28]. However, there is still a lack of research on the impact of larger diameter tumors on DSS. We speculate that larger tumors may be more likely to detect and receive clinical treatment, which may place patients at a fixed risk for DSS. Further exploration about how larger tumors affect prognosis remained warranted.

Furthermore, we also found that male and white population were associated with poor prognosis. Unmarried population had a short-term protection advantage compared to married, but in long-term survival, the protective effect of married began to become obvious. Surgery and the combined treatment with radiotherapy and chemotherapy were also beneficial to the prognosis of GBM. Above-mentioned findings have good agreements with other studies[4, 11, 29–33].
Our study only analyzed GBM patients diagnosed with the only one primary tumor to avoid the influence of secondary or other tumor history on the research results. And two methodological strengths were existed in the current study: 1) Lasso regression was used to select feature variables to avoid eliminating variables based on p-value in univariate cox regression analysis, while considering the collinearity between variables to avoid over-fitting of the model. 2) Restricted cubic spline cox analysis could evaluate the nonlinear exposure response effects of continuous variables in cox regression and avoid resulting in the loss of important information. Meanwhile, the cases included in the SEER database have a large time span and involve a wide range of populations, and are still of great significance for clinical analysis. So, the nomogram was constructed, which provides useful information for clinical prediction of DSS. However, this study also had some limitations: Firstly, this study is a retrospective study based on the SEER database, which may have biases; Secondly, the variable recording in the database are limited, and information related to biomarkers and immune statuses are not recorded, which is not conducive to comprehensive analysis of GBM; Thirdly, the construction and verification of the nomogram were based on the same sample population, and it is necessary to perform verification and reliability evaluation in other populations.

Conclusion

GBM is a malignant brain tumor with high incidence and poor prognosis. Patients over the age of 75 had a significantly increased risk of DSS, so special attention should be paid. Tumors smaller than 75 mm had an approximately linear increasing risk of death from GBM as the diameter increasing, but tumors larger than 75 mm maintained a stable high-risk state. In-depth research on the relevant impact mechanisms about tumor size is required. The constructed nomogram had a robust prediction effect.

Declarations

Availability of data and materials

All data used to generate the main analyses and figures are available at https://seer.cancer.gov/data-software/.

Conflicts of interest

The authors declare no actual or potential competing financial interests.

Funding

This work was supported by Program of the National Natural Science Foundation of China, grant number 81772680.

Acknowledgements
Jing Wei conceived the study idea, analyzed the data, and critically revised the manuscript. Lihua Wu reviewed the literature, and drafted the manuscript. Jianbo Song, Junping Zhang, Wenhui Yang, and Mengxian Zhang contributed to study design. All authors read and approved the final version of the manuscript as submitted and agree to be held accountable for all aspects of the work.

References


