

## RESEARCH ARTICLE

# The role of deep learning-based survival model in improving survival prediction of patients with glioblastoma

Hajar Moradmand<sup>1</sup>  | Seyed Mahmoud Reza Aghamiri<sup>1</sup> | Reza Ghaderi<sup>2</sup> | Hamid Emami<sup>3</sup>

<sup>1</sup>Medical Radiation Engineering, Shahid Beheshti University, Tehran, Iran

<sup>2</sup>Electrical Engineering, Shahid Beheshti University, Tehran, Iran

<sup>3</sup>Department of Radiation Oncology, Isfahan University of Medical Sciences, Seyed Al-Shohada Charity Hospital, Isfahan, Iran

## Correspondence

Hajar Moradmand, Shahid Beheshti University, Shahid Shahriari Square, Daneshjou Boulevard, 19839 69411, Tehran, Iran.  
Email: h\_moradmand@sbu.ac.ir

## Funding information

This research received no funding.

## Abstract

This retrospective study has been conducted to validate the performance of deep learning-based survival models in glioblastoma (GBM) patients alongside the Cox proportional hazards model (CoxPH) and the random survival forest (RSF). Furthermore, the effect of hyperparameters optimization methods on improving the prediction accuracy of deep learning-based survival models was investigated. Of the 305 cases, 260 GBM patients were included in our analysis based on the following criteria: demographic information (i.e., age, Karnofsky performance score, gender, and race), tumor characteristic (i.e., laterality and location), details of post-surgical treatment (i.e., time to initiate concurrent chemoradiation therapy, standard treatment, and radiotherapy techniques), and last follow-up time as well as the molecular markers (i.e., O-6-methylguanine methyltransferase and isocitrate dehydrogenase 1 status). Experimental results have demonstrated that age (Elderly > 65: hazard ratio [HR] = 1.63; 95% confidence interval [CI]: 1.213–2.18;  $p$  value = 0.001) and tumors located at multiple lobes ([HR] = 1.75; 95% [CI]: 1.177–2.61;  $p$  value = 0.006) were associated with poorer prognosis. In contrast, age (young < 40: [HR] = 0.57; 95% [CI]: 0.343–0.96;  $p$  value = 0.034) and type of radiotherapy (others include stereotactic and brachytherapy: [HR] = 0.5; 95% [CI]: 0.266–0.95;  $p$  value = 0.035) were significantly related to better prognosis. Furthermore, the proposed deep learning-based survival model (concordance index [ $c$ -index] = 0.823 configured by Bayesian hyperparameter optimization), outperformed the RSF ( $c$ -index = 0.728), and the CoxPH model ( $c$ -index = 0.713) in the training dataset. Our results show the ability of deep learning in learning a complex association of risk factors. Moreover, the remarkable performance of the deep-learning-based survival model could be promising to support decision-making systems in personalized medicine for patients with GBM.

## KEYWORDS

decision support systems, deep learning, glioblastoma, hyperparameter optimization, survival analysis

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Cancer Medicine* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Glioblastoma (GBM) is the most common fatal malignant brain tumor in adults, with an incidence rate of 3.2 per 100,000 populations.<sup>1</sup> The current approved treatment of GBM is the maximum safe resection surgery of the tumor with a minimum side effect, followed by the combination of radiotherapy and chemotherapy generally with temozolomide.<sup>2</sup> The prognosis of GBM patients (median overall survival of 14 months) has remained poor in the past three decades, even with severe multi-pronged therapies. In some cases ( $\leq 10\%$ ), a 5-year survival rate has been reported with the same routine treatment procedure.

Accurate prediction of individual patients' prognosis is a crucially important task not only for patients and their families but also for physicians to support personalized treatment and to identify who benefits from aggressive or moderate treatment and avoid ineffective treatment.

Conventionally, the Cox proportional hazard (CoxPH) model<sup>3</sup>, as represented in Equation (1), evaluates the hazard function of the event occurring at time  $t$ ,  $h_i(t, z_i)$ , for a patient  $i$  based on the linear combination of the covariates ( $Z$ ), in which  $\beta$  is a regression coefficient,  $p$  is a vector of unknown variables, and  $h_0(t)$  is an indefinite baseline hazard function.

$$h_i(t, z_i) = h_0(t) \prod_{k=1}^p e^{\beta_k z_{ik}} \quad (1)$$

whereas the CoxPH model assumes that each covariate influences patient's risk factor independent of another covariate, in other words, it presumes a linear combination of covariates, it may be too naive to model the effect of nonlinear risk factors on patient's survival.

To address these drawbacks, tremendous effort and methods have been employed in survival analysis.

Hitherto, machine learning algorithms such as random forest, artificial neural networks (ANNs), and support vector machines have shown striking results in many applications. Machine learning algorithms also have been effectively adopted, either as competition (e.g., the random survival forest [RSF]<sup>4</sup>) or as a complement (e.g., Cox-net), with the standard survival analysis model such as CoxPH.<sup>5</sup>

The concept of using an ANN in survival analysis, for learning nonlinear risk functions, was first proposed by the Faraggi-Simon network.<sup>6</sup> In this approach, the amount of  $\beta z_i$  in Equation (1), was replaced with the output of a single-layer feed-forward neural network to determine the vector of the unknown parameters  $\theta$ . Though the Faraggi-Simon model did not significantly outperform the standard CoxPH model, it suggested that a similar extension can be constructed. Since then, many attempts have been performed to acquisition and handle the superb capacity of the neural network in the survival analysis.

Recently, deep learning has attracted remarkable attention for modeling the complex interactions between the covariates in the survival analysis,<sup>7,8</sup> among them the deep learning-based survival model (DeepSurv)<sup>9</sup> has provided striking results. Hitherto, some studies suggested that the DeepSurv models have learned efficiently the complex patient's risk factor obtained from multiple parameters and outperformed in estimating the failure of treatment for different cancer types such as cervical, oral, and lung cancers.<sup>10,11</sup> However, to our knowledge, no study has been performed on patients with GBM. Furthermore, the performance of the deep learning-based survival model is intimately affected by the appropriate configuration of the model hyperparameters. Albeit, it has remained a challenging and time-consuming task due to computational and process limitations.

On the other hand, given that GBM is a highly heterogeneous tumor at both molecular and histological levels, the combination of both clinical manifestations of the patients and the molecular marker of GBM may give a better survival prediction.<sup>12</sup> However, rare studies have investigated the combined influence of characteristics such as clinical data, tumor characteristics, treatment options, and molecular markers of GBM in the survival model's predictive accuracy.<sup>13</sup> Accurate patients' survival prediction remains a challenging work.

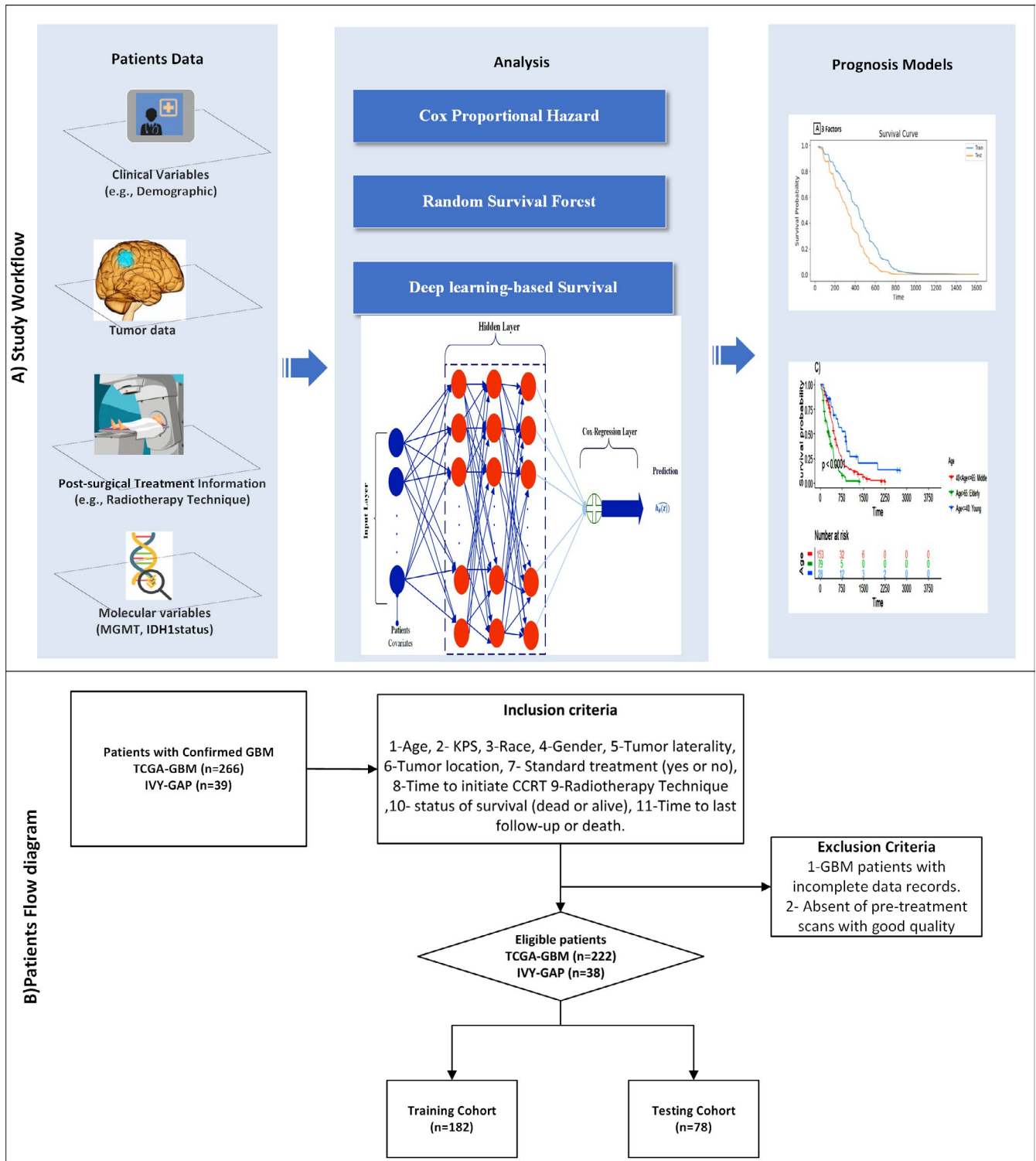
Therefore, this study was designed to investigate three issues: (i) To evaluate the effect of concurrent multivariate risk factors including patient characteristics, tumor site, post-surgery treatment specification, and molecular marker on patient's survival; (ii) To validate the deep learning-based survival model performance in comparison with two other survival reference models (i.e., CoxPH and RSF); and (iii) To assess the impact of hyperparameter optimization methods (random search and Bayesian optimization) on the performance of deep learning-based survival model. The workflow of our work is summarized in Figure 1A.

The rest of the paper is as follows. In the next section, our dataset, the survival models, particularly the deep learning-based survival model, are clarified. Furthermore, hyperparameter optimization techniques and statistic metrics used for our assessment are explained in detail. Experimental results are illustrated in Section 3. Finally, a comparison analysis with the results of previous studies and the conclusion of our work is described in Section 4.

## 2 | MATERIALS AND METHODS

### 2.1 | Data collection

Of the 305 histopathological confirmed GBM patients, 260 cases (163 males and 97 females with a mean age of 59 years) were selected. Patients were eligible who meet our inclusion



**FIGURE 1** Study workflow and patients flow diagram. Panel (A) workflow of this study. Panel (B) study flow diagram of patients' recruitment, with the exclusion criteria. CCRT, concurrent chemoradiotherapy

criteria as follows: (i) demographic characteristics encompass the gender, age at surgery, Karnofsky performance score (KPS), and race; (ii) tumor laterality and location, (iii) detailed of post-surgical treatment, that is, time to start concurrent chemoradiotherapy (CCRT) after surgery, receiving standard treatment (yes or no); (iv) radiotherapy techniques

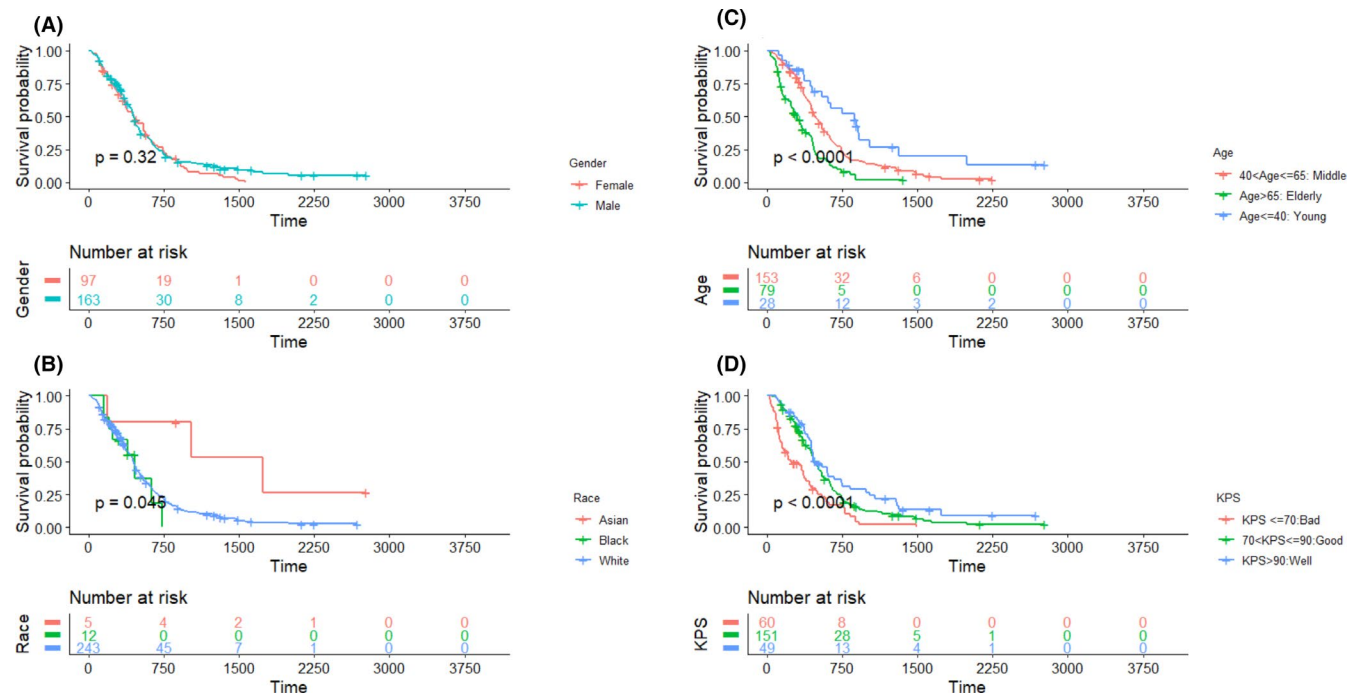
including three-dimensional conformal radiotherapy (3D-CRT), intensity-modulated radiotherapy (IMRT), other (stereotactic or brachytherapy); and (v) access to last follow-up time and death status of survival (dead or alive). The flow diagram of patients' recruitment, with the inclusion criteria in the study, is depicted in Figure 1B.

**TABLE 1** Demographic information and tumor characteristics of eligible patients (from the TCGA-GBM and IVY Gap datasets) with glioblastoma

Collection (n)	Gender (n, %)	Age (n, %)	KPS (n, %)	Race (n, %)	Tumor laterality (n, %)	Tumor location (n, %)
TCGA-GBM (221)	Female (80, 36%)	Age < 40 (25, 11%)	KPS ≤ 70 (53, 24%)	White (204, 93%)	Right (103, 47%)	Frontal (64, 29%)
	Male (141, 64%)	40 ≤ Age < 65 (125, 57%)	70 < KPS ≤ 90 (130, 58%)	Black (12, 5%)	Left (118, 53%)	Temporal (63, 28%)
		Age ≥ 65 (71, 32%)	KPS > 90 (38, 18%)	Asian (5, 2%)		Parietal (37, 17%)
						Occipital (19, 9%)
					Other <sup>a</sup> (38, 17%)	
IVY-GAP (39)	Female (17, 44%)	Age < 40 (3, 7%)	KPS ≤ 70 (7, 18%)	NA	Right (27, 69%)	Frontal (12, 31%)
	Male (22, 54%)	40 ≤ Age < 65 (28, 72%)	70 < KPS ≤ 90 (21, 53%)		Left (12, 31%)	Temporal (10, 26%)
		Age ≥ 65 (8, 21%)	KPS > 90 (11, 29%)			Parietal (13, 33%)
					Occipital (3, 8%)	
					Other (1, 2%)	

Abbreviations: GAP, Glioblastoma Atlas Project; KPS, Karnofsky performance score; NA, not available; TCGA, The Cancer Genome Atlas Glioblastoma Multiforme.

<sup>a</sup>Other is related to tumors located at other lobes or more than one lobe.



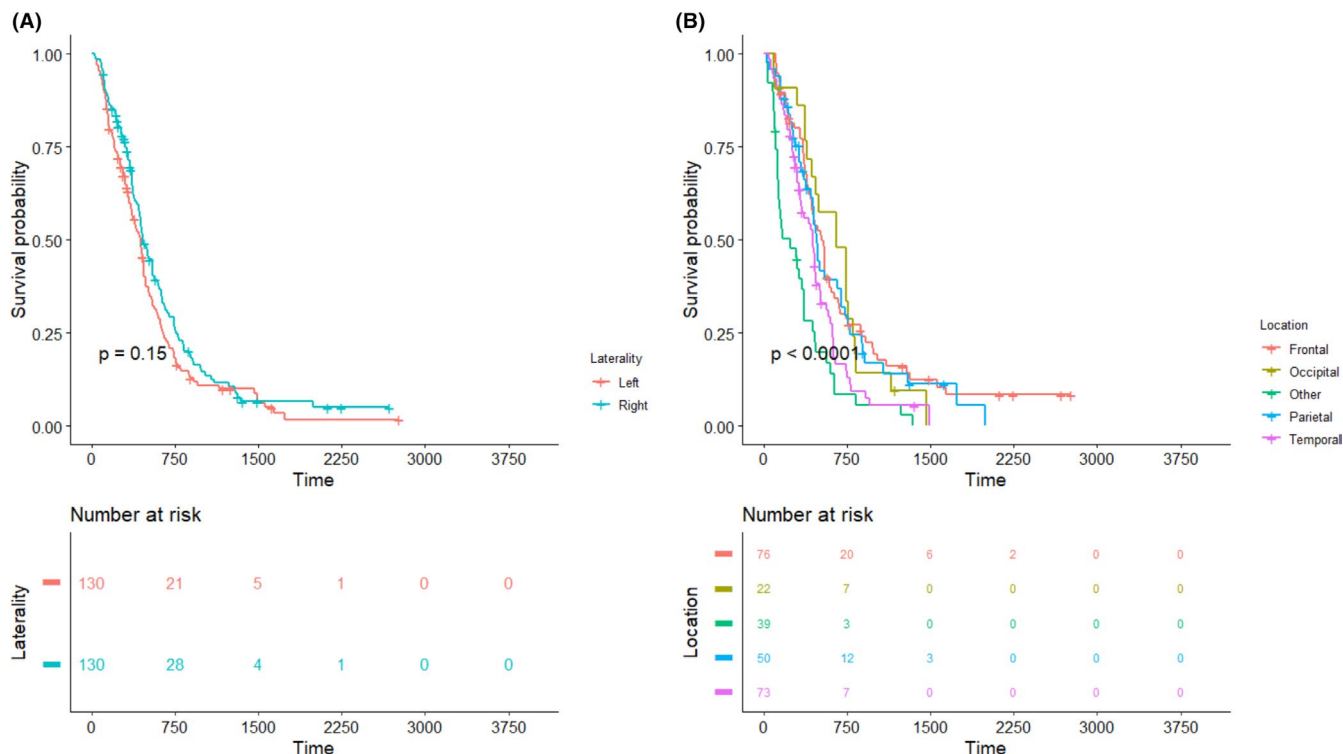
**FIGURE 2** Kaplan–Meier curves on the overall dataset (260 patients) are represented for clinical factors, with pairwise comparisons using the log-rank test and risk table. Panels (A–C), respectively, confirm that the KPS, age, and race are statistically significant prognoses. While panel (D) shows, gender is not a statistically significant prognostic factor. KPS, Karnofsky performance score

These data are available from the The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) series<sup>14</sup> and IVY Glioblastoma Atlas Project (Ivy GAP) database<sup>15</sup> and provides the pathological, genetic, clinical data, and radiological data of patients. Both datasets are without a patient identifier, so the approval of the institutional review board is not required.

The per-operatives magnetic resonance volumes were employed to determine tumor laterality and tumor location of

the selected TCGA-GBM patients by a specialist. In Table 1, the clinical and tumor information of the eligible cases for each series (TCGA-GBM and IVY GAP) is summarized.

Moreover, isocitrate dehydrogenase 1 (IDH1) and O-6-methylguanine-methyltransferase (MGMT) were integrated into our survival analysis as molecular markers. These markers have frequently been reported as favorable prognostic factors of GBM patients. The molecular characterization of the TCGA-GBM dataset is available



**FIGURE 3** Kaplan–Meier curves on the 260 patients are represented for tumor factors, with pairwise comparisons using the log-rank test and risk table. (A) Present a statistically significant differences in the log-rank test for tumor location, but the (B) shows no statistically significant difference in the tumor laterality

through the Genomic Data Commons Data Portal and was extracted by the Bioconductor (open development and free, open-source software) packages in R (version 3.6.2) languages. The clinical and genomic data of IVY GAP are accessible via <https://glioblastoma.alleninstitute.org>. The distribution of molecular markers, that is, IDH1 and MGMT, for all eligible patients is represented in Table S1.

## 2.2 | Survival models

Survival analysis (or time-to-event analysis) is an actuarial method that has tremendous applications in clinical oncology. One of the main objectives in survival analysis is to designate the probability of occurrence of the event of interest (e.g., death time) beyond any specified time ( $t$ ), that is, survival function  $S(t) = \Pr(T > t)$ . Alternatively, the survival function can be accessed by  $S(t) = \exp(-H(t))$ , where  $H(t)$  is the cumulative hazard function (CHF) and is defined as  $H(t) = \int_0^t h(x)dx$ . Heretofore, a wide range of statistical methods has been presented in three main categories (i.e., parametric, non-parametric, and semi-parametric) to estimate the survival function and hazard ratio. In this study, the performance of the state-of-the-art survival model, that is, deep learning-based survival models, is compared alongside the two reference survival models, that is, CoxPH and RSF.

### 2.2.1 | Cox proportional hazard

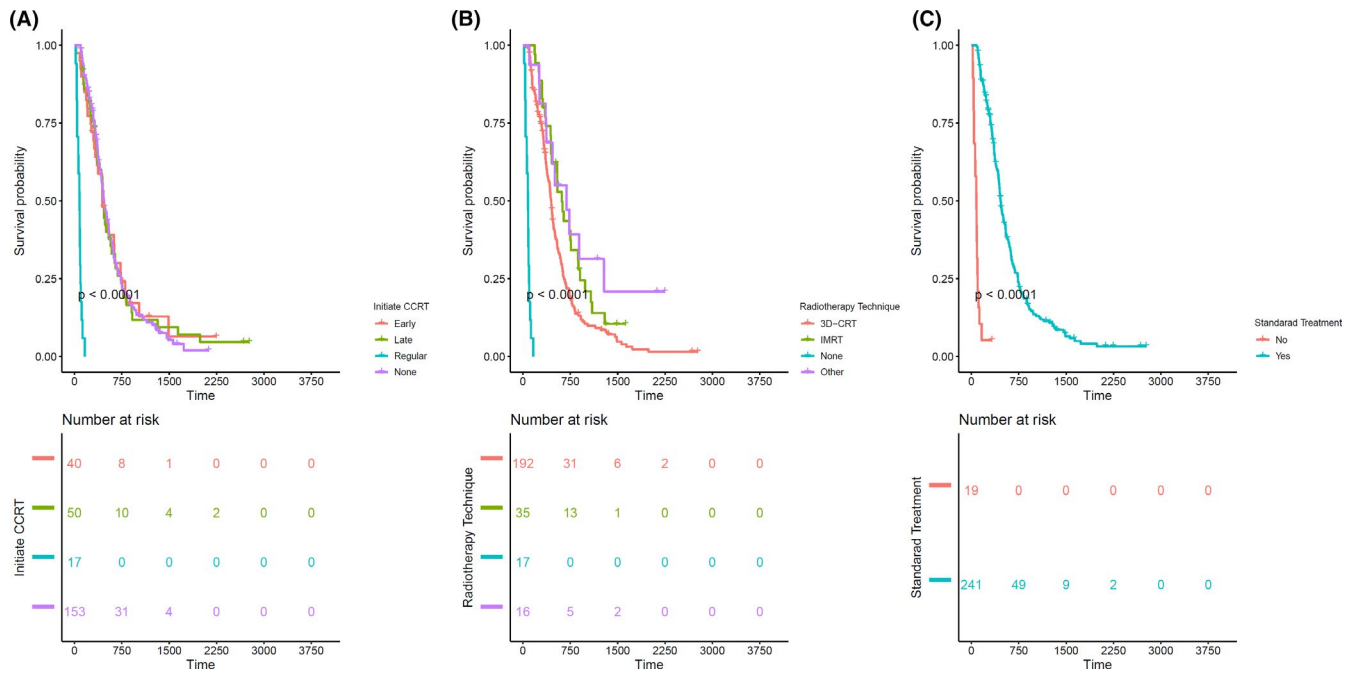
There is no doubt, that the CoxPH models are the most pervasive survival models in medical analysis, because of their simple execution and informative explanation. As described earlier (Equation 1), the CoxPH model is a linear combination of the covariate  $(\beta_1 Z_1 + \dots + \beta_p Z_p)$ . If denote  $c_k(c_1, \dots, c_k)$  the possibly censored event time for individual  $k$ , the corresponding partial likelihood<sup>16</sup> is defined by Equation (2).

$$L_{\text{Cox}} = \prod_{k=1}^n \left( \frac{e^{\beta_k Z_k(c_k)}}{\sum_{j \in R_k} e^{\beta_j Z_j(c_k)}} \right)^{D_k} \quad (2)$$

where  $R_k$  refers to the set of individuals at risk at event time, and if individual  $k$  is an observed event time,  $D_k = 1$  otherwise  $D_k = 0$ .

### 2.2.2 | Random survival forest

Random survival forest, an extension of Bierman's random forest method in survival analysis, is a non-linear and non-parametric model.<sup>4</sup> RSF is determined based on an ensemble tree, where a tree is grown by applying B bootstrap samples randomly of each data. Almost 37% of the data are



**FIGURE 4** Kaplan–Meier curves on the 260 GBM patients are represented for treatment factors, with pairwise comparisons using the log-rank test and risk table. (A) Shows a statistically significant difference in the survival of patients who did not start any standard treatment and patients who initiated standard treatment at regular time with a  $p$  value  $< 0.0001$ . (B) Indicate a statistically significant difference in the survival of patients who did not receive any radiotherapy and patients who received radiotherapy with a  $p$  value  $< 0.0001$ . (C) Present a statistically significant difference in the survival of patients receiving standard radiotherapy with concomitant chemotherapy (Commonly temozolomide) compared with those who did not receive such treatment with a  $p$  value  $< 0.0001$ . 3D-CRT, three-dimensional conformal radiotherapy; GBM, glioblastoma; I\_CCRT, initiate concurrent chemoradiation therapy; IMRT, intensity-modulated radiation therapy; ST, standard treatment

excluded in each bootstrap (in-bag) sample, which implies out-of-bag data. Subsequently, variables with suitable criteria with maximum log-rank risk tests are nominated and randomly selected to dichotomize each node of a tree.

This process iteratively is continued until it met the stopping criteria. The ensemble CHF is determined by averaging over the CHF of each tree from nodes with no further split (terminal nodes). Eventually, the prediction error of ensemble CHF is estimated by the concordance index ( $c$ -index).

### 2.2.3 | DeepSurv

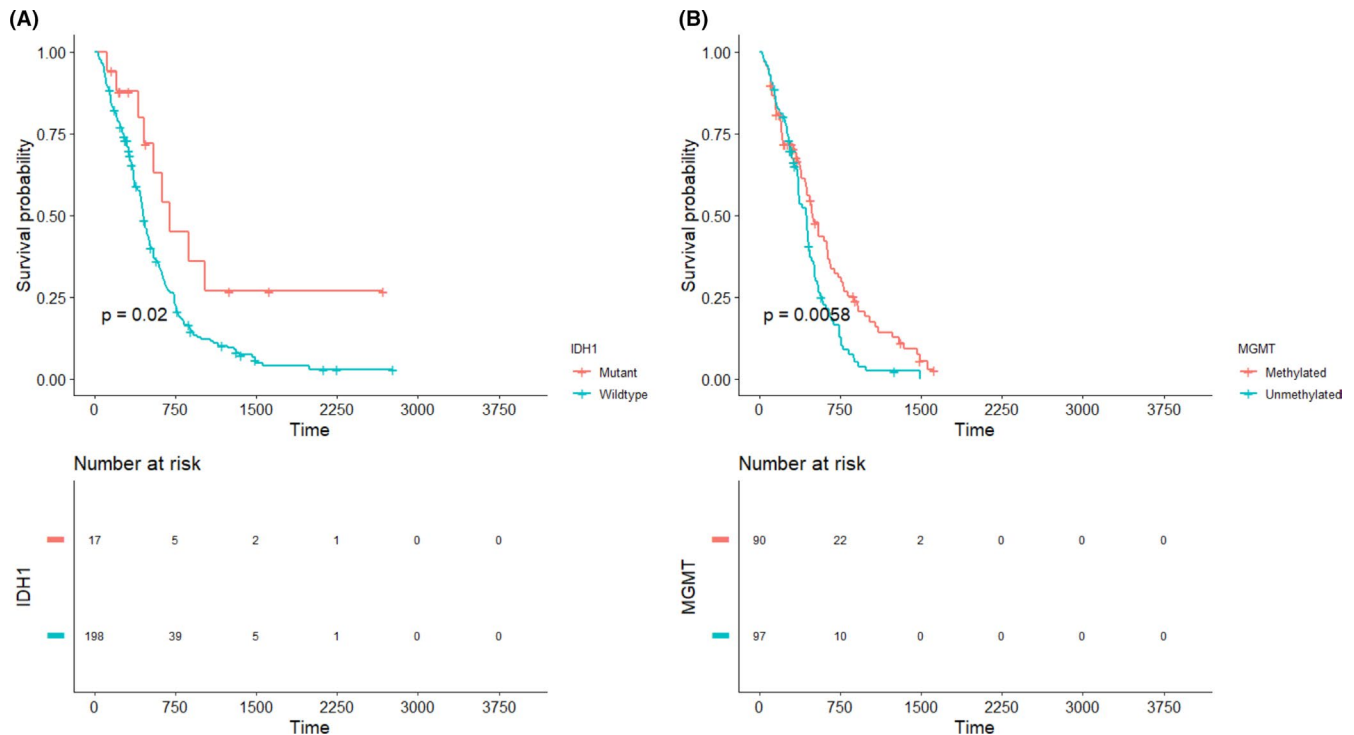
DeepSurv, a deep feed-forward neural network, is a non-linear extension of the CoxPH model.<sup>9</sup> The risk function ( $h_{\theta}(z)$ ) of DeepSurv is estimated by the network output and is parameterized using the weight of the network ( $\theta$ ). The loss function of the network (Equation 3) is computed by taking a negative log over the partial likelihood of  $L_{\text{COX}}$  from Equation (2), with extra modification.

where  $R_k$  refers to the set of individuals at risk at event time, and  $D_k = 1$  if individual  $k$  is an observed event time, otherwise  $D_k = 0$ .

DeepSurv has employed more advanced training methods such as rectified linear unit (ReLU) function, dropout, batch normalization, etc., using TensorFlow (an open-source Python library) to improve efficiency.

Deep neural networks have demonstrated remarkable performance on many machine learning applications.<sup>17</sup> However, their performance is highly affected by the appropriate configuration of the model hyperparameters to yield the minimum value of the loss function and the best value for the model parameters. Lately, Bayesian optimization has demonstrated promising results by providing more powerful and intelligent tools for assessing search spaces.<sup>18</sup> In contrast to random search or grid search, Bayesian optimization is sequential model-based optimization, in which the mean and variance of the model are sequentially updated to the last observation.<sup>19</sup>

$$\text{loss}(\theta) := -\frac{1}{n(D=1)} \sum_{k:D_k=1} \log \left( \sum_{j \in R_k} \exp(\beta_k Z_k(c_k)) - \beta_k Z_k(c_k) \right) + \lambda * \|\theta_2^2\| \quad (3)$$



**FIGURE 5** Kaplan–Meier curves on 260 patients are represented for molecular markers (IDH1 and MGMT), using the log-rank test pairwise comparisons. (A) Indicates that IDH1 (wild-type) is statistically significantly related to poor prognosis with a  $p$  value  $< 0.02$ . (B) Shows methylated MGMT is related to a better prognosis with a  $p$  value  $< 0.00058$ . IDH1, isocitrate dehydrogenase 1; MGMT, O-6-methylguanine methyltransferase

In this work, the CoxPH and the RSF were performed using the R packages, survival, and randomForestSRC, respectively, while DeepSurv by an open-source Python package. Hyperopt, a TensorFlow Python package,<sup>20</sup> was employed for Bayesian hyperparameter optimization.

### 2.3 | Statistical analysis

Typical metrics such as the root mean squared error are inappropriate for survival model analysis since survival data usually incorporate censoring data. In general, censoring occurs in survival data due to missing patients' follow-up or patients alive more than the study time. In this work, variables prognosis was assessed by the univariate and multivariate CoxPH regression model, the log-rank test, and the hazard ratio. The performance of the survival models was evaluated by the  $c$ -index.<sup>21</sup> Extra information about these metrics is available in Data S1. Survival curves were plotted using the Kaplan–Meier methods. All statistical analysis were performed by the R language (<https://www.r-project.org/>, version 3.6.2), using the survival, survminer, and survivalROC packages.

## 3 | RESULTS

The Kaplan–Meier curves of the demographic information of patients, tumor variables, and post-surgical treatment characteristics, as well as molecular markers (e.g., IDH1 and MGMT), are presented in Figures 2–5, respectively, to assess the concurrent prognostic effect of inclusion criteria in this study.

The relevant variables to the prognosis of the GBM patients were rated by univariate CoxPH analysis for both clinical and molecular variables, as are represented in Tables S2 and S3, respectively.

The multivariate analysis was performed using statistically significant variables in univariable analysis (Table 2).

Henceforth, the entire eligible dataset was randomly divided into the training set (70%) and into the testing set (30%), which was repeated 10 times to ensure all data were examined.

The survival outcome of the training set (to building the predictive model) and testing set (to assess the prediction model accuracy) was not significantly different. The  $c$ -index was computed in each iteration on the train and test datasets. The ultimate  $c$ -index for three predictive survival models, that is, CoxPH, RSF, and DeepSurv, was obtained by averaging across indicators.

**TABLE 2** Multivariate analysis over statistically significant univariate variables for overall survival in patients with glioblastoma using Cox proportional hazard regression models with concordance index = 0.69 and log-rank = 9.412e-19

Characteristic	Overall survival		
	Factors (n)	Hazard ratio (95% CI)	p value
<b>Demographic</b>			
Age	40 < Middle ≤ 65 (153)	Reference	
	Elderly > 65 (79)	1.70 (1.229–2.34)	0.018*
	Young ≤ 40 (28)	0.58 (0.343–0.99)	0.034*
KPS	Bad ≤ 70 (60)	Reference	
	70 < Good ≤ 90 (151)	0.96 (0.660–1.39)	0.827
	Well > 90 (49)	0.80 (0.503–1.28)	0.352
Race	Asian (5)	Reference	
	Black (12)	2.93 (0.73–11.74)	0.13
	White (243)	2.92 (0.888–9.57)	0.078
<b>Tumor</b>			
Location	Frontal (76)	Reference	
	Occipital (22)	0.79 (0.471–1.33)	0.381
	Multiple <sup>a</sup> (39)	1.58 (1.16–2.46)	0.012*
	Parietal (50)	1.06 (0.705–1.60)	0.771
	Temporal (73)	1.25 (0.859–1.81)	0.045
<b>Post-surgery treatment</b>			
Initiate CCRT	Early (40)	Reference	
	Late (50)	1.03 (0.593–1.60)	0.92
	None (17)	8.84 (0.994–78.68)	0.05
	Regular (153)	0.88 (0.583–1.34)	0.56
Radiotherapy type	3D-CRT (192)	Reference	
	IMRT (35)	0.68 (0.433–1.035)	0.082
	None (17)	Reference	
	Other <sup>b</sup> (16)	0.5 (0.266–0.95)	0.035*
Standard treatment	No (19)	Reference	
	Yes (241)	0.35 (0.0469–2.75)	0.32

Abbreviations: 3D-CRT, three-dimensional conformal radiation therapy; CCRT, concurrent chemoradiation therapy; CI, confidence interval; IMRT, intensity-modulated radiation therapy; KPS, Karnofsky performance score.

<sup>a</sup>Other relates to tumors located at more than one lobe.

<sup>b</sup>Multiple is related to remaining radiotherapy methods such as stereotactic or brachytherapy.

\*Variable with *p* value < 0.05 is considered significant.

The optimum value of the hyperparameters was selected using two main strategies which include: (i) random search and (ii) Bayesian optimization. The performance of the hyperparameter tuning was evaluated by *k*-means cross-validation (*k* = 5). A configuration with the largest validation *c*-index was determined to avoid the models' overfitting. For searching hyperparameters, 100 iterations were performed. The selected hyperparameters used in the configuration of the DeepSurv models are represented in Table S4.

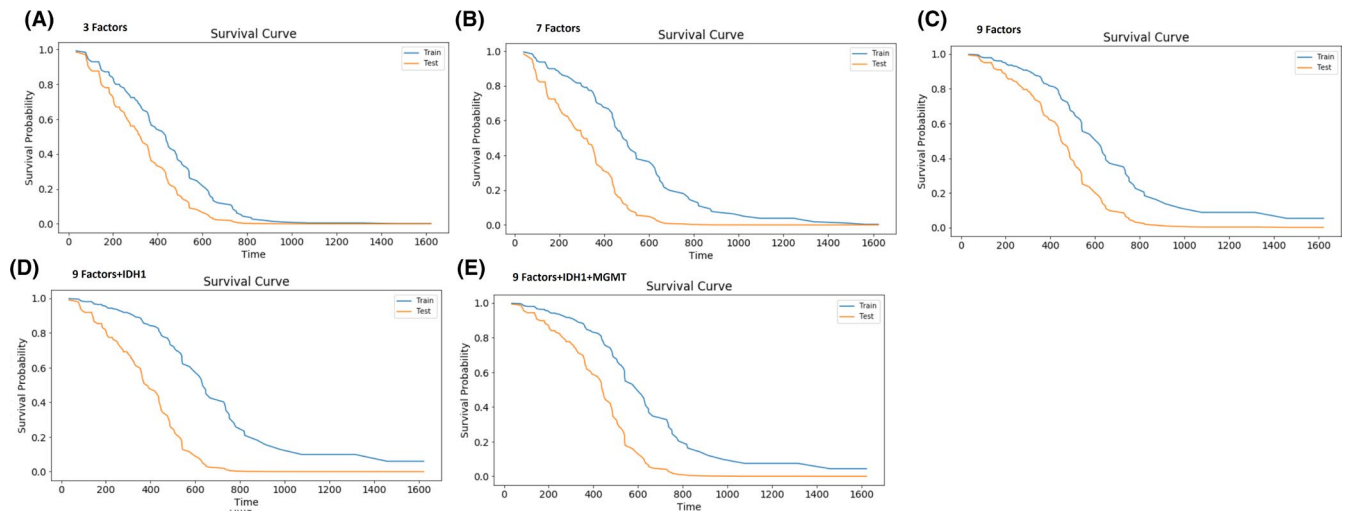
The DeepSurv model was developed with a three-layer neural network.

Our analysis was started with three multivariate statistically significant factors (i.e., age, tumor

location, and radiotherapy methods). Thereupon, statistically significant variables in univariate analysis, insignificant factors, and molecular markers were integrated into our survival analysis. Kaplan–Meier survival curves of these combinations of risk factors on both training and testing datasets were plotted for the deep learning-based survival model (DeepSurv) in Figure 6.

A comparison of the performance of survival models, that is, CoxPH, RSF, and DeepSurv, tuning with random search and Bayesian hyperparameter, at each level of increasing the variables is presented in Table 3.





**FIGURE 6** Kaplan–Meier survival curves of the deep learning-based survival model (DeepSurv, optimized with Bayesian hyperparameter) on the training and testing datasets. Panel (A) shows the survival curve for three statistically significant multivariate factors (age, tumor location, and radiotherapy types), panel (B) shows the survival of seven statistically significant univariate factors (age, tumor location, and radiotherapy types, KPS, race, initiate CCRT, and standard treatment), panel (C) shows the survival curve of all factors (age, tumor location, and radiotherapy types, KPS, race, initiate CCRT, standard treatment, gender, and laterality), panel (D) shows the survival curve of a molecular marker (IDH1) combined with nine factors, and panel (E) shows the survival curve of molecular markers (IDH1 and MGMT) combined with nine factors. CCRT, concurrent chemoradiation therapy; IDH1, isocitrate dehydrogenase 1; KPS, Karnofsky performance score; MGMT

## 4 | DISCUSSION

Though standard treatments of GBM tumors only postpone tumors growth for a while, introducing novel and personalized treatment methods may engender a new door to improve prognosis or even cure this disease. Accurate survival predictions are always desirable for physicians and patients to individualize treatment planning and avoid inessential treatments.

This study was designed to examine the influence of concurrent variables, that is, clinical, tumor, post-surgery treatment, and molecular factors, and mainly to validate and optimize the performance of deep learning-based survival models in improving the prediction accuracy of survival models.

Our results in Table S2 indicate that risk factors including race, age, KPS, tumor location, standard treatment, time to initiate CCRT, and type of radiotherapy are univariate significant for overall survival prediction ( $p$  value  $\leq 0.05$ ). However, three of them (i.e., age, tumor location, and type of radiotherapy) were statistically significantly prognosis in multivariate analysis (Table 2). In agreement with previous work,<sup>22</sup> age and tumor location factors were reported as significant multivariate covariates. Since, study results have demonstrated that univariable relations alone may not be sufficient and informative to determine important significant variables, particularly for complex datasets, using significant univariate statistical variables in multivariate analysis is the

most common approach. Multivariate analysis selects the variables that are independently most closely related to prognosis<sup>23</sup>, given that two interrelated variables are unlikely to choose both as significant variables by multivariate analysis.

Our results indicated that younger patients' survival was better than elderly patients. Though there is no consensus in standardized age cutoff,<sup>24</sup> we categorized patients into elderly  $\geq 65$ ,<sup>25,26</sup> young  $< 40$ ,<sup>27</sup> and middle ( $40 \leq \text{age} < 65$ ) ages, based on the mentioned references. Age consistently is reported in various literature as efficacious prognostic survival variable,<sup>26,28,32,33,34</sup>.

In consensus with previous works,<sup>29,30</sup> our findings also showed that tumors located at the temporal lobes or multiple lobes were associated with unfavorable prognostic factors compared with tumors located at frontal lobes. The poor prognosis of tumors located at the temporal lobes can be interpreted by the results of Kocher et al.,<sup>31</sup> in which tumors located at the temporal lobe and the parietal lobe were determined as the most vulnerable lobes for cognitive function in patients with GBM. Furthermore, patients who received non-standard treatment of GBM showed the worst prognosis, which is compatible with the benefit of current standard care of GBM patients (i.e., radiotherapy concurrent with chemotherapy, e.g., Temozolomide).

Besides, patients who underwent radiotherapy with stereotactic therapy or brachytherapy were statistically associated with a favorable prognosis. However, only 7%

TABLE 3 Comparison of the DeepSurv, RSF, and CoxPH survival models by the c-index, with different combinations of variables, on the train and test datasets

Level	Survival model	c-index									
		DeepSurv (Random search)		DeepSurv (Bayesian optimization)		RSF		CoxPH			
		Train	Test	Train	Test	Train	Test	Train	Test		
1	Three factors (A + Lo + RT)	0.670	0.602	0.672	0.610	0.631	0.597	0.629	0.598		
2	Seven factors (A + Lo + RT + KPS + I_CCRT + RT + ST + L)	0.702	0.587	0.72	0.608	0.679	0.592	0.667	0.590		
3	Nine factors (A + Lo + I_CCRT + KPS + RT + ST + L + R + G)	0.774	0.631	0.784	0.625	0.698	0.614	0.683	0.612		
4	Nine factors + IDH1	0.779	0.638	0.796	0.645	0.701	0.625	0.681	0.610		
5	Nine factors + IDH1 + MGMT	0.808	0.693	0.823	0.70	0.728	0.668	0.713	0.674		

Abbreviations: CCRT, concurrent chemoradiation therapy; c-index, concordance index; CoxPH, Cox proportional hazard; I\_CCRT=initiate CCRT; IDH1, isocitrate dehydrogenase 1; KPS, Karnofsky performance score; MGMT, O-6-methylguanin-methyltransferase; RSF, random survival forest; RT, radiotherapy technique; ST, standard treatment.

A = age, Lo = location, L = laterality, R = race, G = gender.

of cases were treated using these methods, and further studies are required to accept this factor as a biomarker. These findings are supported by the fact that ongoing improvements in medical imaging and radiation therapy techniques that have facilitated treatment volume delineation and treatment conformality.<sup>32,33</sup> For example, in contrast to the 3D-CRT, IMRT feasible further treatment conformality using several modulated beams with various intensities at different angles. Stereotactic radiotherapy allows for an even more accurate representation of treatment volumes while saving surrounding vital structures, using many beam sources

To build survival models, we started with three significant variables in our multivariate analysis. At this step, the DeepSurv (optimized with random search), CoxPH, and RSF models achieved a *C*-index of 0.67, 0.629, and 0.631, respectively. Subsequently, all variables were added in five levels to evaluate the effect of different variables combinations on survival model accuracy. After involving all variables in the analysis (fifth level), the DeepSurv models configured with random search (*c*-index = 0.808) outperformed the CoxPH (*c*-index = 0.713) and RSF (*c*-index = 0.728) survival models. Furthermore, at this level, the DeepSurv model, configured with Bayesian hyperparameter optimization, strikingly performed best among the three survival models and achieved the highest *c*-index = 0.823.

In conclusion, from our promising findings, four analytical issues were deduced. First, age, tumor location, and methods of radiotherapy were independently significant prognosis variables. Second, even insignificant prognosis variables played roles in improving the predictive accuracy of survival models.

Third, the deep learning-based survival model outperformed the Cox proportional hazard regression and random survival forest models' ability for accurate GBM survival prediction. Furthermore, the optimum hyperparameter tuning may remarkably improve the deep learning-based survival models performance. As a result, the remarkable performance of DeepSurv indicates the ability of the deep learning model in learning complex association of risk factors. The deep learning-based survival model may have a great potential to be incorporated into the treatment planning of patients with GBM in a routine oncology workflow by improving the prediction of mortality risk of GBM patients. In the future, this work will be extended to investigate the adding value of the reproducible radiomics features to the deep learning-based survival model.<sup>34</sup>

## ACKNOWLEDGMENTS

The authors are grateful to all participants in this research.

## CONFLICT OF INTEREST

There is no conflict of interest declared in this article.

## DATA AVAILABILITY STATEMENT

The data might be made available upon request, and some restrictions will apply.

## ORCID

Hajar Moradmand  <https://orcid.org/0000-0003-0186-468X>

## REFERENCES

- Ostrom QT, Cioffi G, Gittleman H, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. *Neuro Oncol*. 2019;21:v1-v100.
- Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med*. 2005;352:987-996.
- Cox DR. Regression models and life-tables. *J R Statist Soc: Ser B (Methodol)*. 1972;34:187-202.
- Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. 2008;2:841-860.
- Wang P, Li Y, Reddy CK. Machine learning for survival analysis: a survey. 2017. Accessed August 01, 2017. <https://ui.adsabs.harvard.edu/abs/2017arXiv170804649W>
- Faraggi D, Simon R. A neural network model for survival data. *Stat Med*. 1995;14:73-82.
- Nezhad MZ, Sadati N, Yang K, Zhu D. A deep active survival analysis approach for precision treatment recommendations: application of prostate cancer. *Expert Syst Appl*. 2019;115:16-26.
- Ren K, Qin J, Zheng L, et al. Deep recurrent survival analysis. 2018. Accessed September 01, 2018. <https://ui.adsabs.harvard.edu/abs/2018arXiv180902403R>
- Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol*. 2018;18:24.
- Kim YJ, Lee HJ, Kim KG, Lee SH. The effect of CT scan parameters on the measurement of CT radiomic features: a lung nodule phantom study. *Comput Math Methods Med*. 2019;2019:8790694.
- Liu K, Xia W, Qiang M, et al. Deep learning pathological microscopic features in endemic nasopharyngeal cancer: prognostic value and potential role for individual induction chemotherapy. *Cancer Med*. 2020;9:1298-1306.
- Ferguson SD, Hodges TR, Majd NK, et al. A validated integrated clinical and molecular glioblastoma long-term survival-predictive nomogram. *Neuro-Oncol Adv*. 2020;3:vdaa146.
- Senders JT, Staples P, Mehrdash A, et al. An online calculator for the prediction of survival in glioblastoma patients using classical statistics and machine learning. *Neurosurgery*. 2020;86:e184-e192.
- Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015;19:A68-A77.
- Puchalski RB, Shah N, Miller J, et al. An anatomic transcriptional atlas of human glioblastoma. *Science*. 2018;360:660-663.
- Lin DY, Wei LJ. The robust inference for the cox proportional hazards model. *J Am Stat Assoc*. 1989;84:1074-1078.

17. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*. 2018;15:20170387.
18. Feurer M, Springenberg JT, Hutter F. Initializing Bayesian hyperparameter optimization via meta-learning. Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. Austin, Texas, AAAI Press; 2015.
19. Smith LN. A disciplined approach to neural network hyperparameters: part 1 -- learning rate, batch size, momentum, and weight decay. 2018. Accessed March 01, 2018. <https://ui.adsabs.harvard.edu/abs/2018arXiv180309820S>
20. Putatunda S, Rama K. A Comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost. Proceedings of the 2018 International Conference on Signal Processing and Machine Learning. Shanghai, China, Association for Computing Machinery; 2018.
21. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med*. 1984;3:143-152.
22. Mineo JF, Bordron A, Baroncini M, et al. Prognosis factors of survival time in patients with glioblastoma multiforme: a multivariate analysis of 340 patients. *Acta Neurochir (Wien)*. 2007;149:245-252. discussion 252-253.
23. Alexopoulos EC. Introduction to multivariate regression analysis. *Hippokratia*. 2010;14:23-28.
24. Lin Z, Yang R, Li K, et al. Establishment of age group classification for risk stratification in glioma patients. *BMC Neurol*. 2020;20:310.
25. Ewelt C, Goepfert M, Rapp M, Steiger HJ, Stummer W, Sabel M. Glioblastoma multiforme of the elderly: the prognostic effect of resection on survival. *J Neurooncol*. 2011;103:611-618.
26. Ladomersky E, Scholtens DM, Kocherginsky M, et al. The coincidence between increasing age, immunosuppression, and the incidence of patients with glioblastoma. *Front Pharmacol*. 2019;10:200.
27. Lee Y, Scheck AC, Cloughesy TF, et al. Gene expression analysis of glioblastomas identifies the major molecular basis for the prognostic benefit of younger age. *BMC Med Genomics*. 2008;1:52.
28. Walid MS. Prognostic factors for long-term survival after glioblastoma. *Perm J*. 2008;12:45-48.
29. Krex D, Klink B, Hartmann C, et al. Long-term survival with glioblastoma multiforme. *Brain*. 2007;130:2596-2606.
30. Natsume A, Aoki K, Ohka F, et al. Genetic analysis in patients with newly diagnosed glioblastomas treated with interferon-beta plus temozolomide in comparison with temozolomide alone. *J Neurooncol*. 2020;148:17-27.
31. Kocher M, Jockwitz C, Caspers S, et al. Role of the default mode resting-state network for cognitive functioning in malignant glioma patients following multimodal treatment. *Neuroimage Clin*. 2020;27:102287.
32. Mann J, Ramakrishna R, Magge R, Wernicke AG. Advances in radiotherapy for glioblastoma. *Front Neurol*. 2018;8:748.
33. Minniti G, Niyazi M, Alongi F, Navarria P, Belka C. Current status and recent advances in reirradiation of glioblastoma. *Radiat Oncol*. 2021;16:36.
34. Moradmand H, Aghamiri SMR, Ghaderi R. Impact of image preprocessing methods on reproducibility of radiomic features in multimodal magnetic resonance imaging in glioblastoma. *J Appl Clin Med Phys*. 2020;21:179-190.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

**How to cite this article:** Moradmand H, Aghamiri SMR, Ghaderi R, Emami H. The role of deep learning-based survival model in improving survival prediction of patients with glioblastoma. *Cancer Med*. 2021;10:7048–7059. <https://doi.org/10.1002/cam4.4230>