frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 30 November 2023
DOI 10.3389/fmed.2023.1271687

OPEN ACCESS

EDITED BY
Gongning Luo,
Harbin Institute of Technology, China

REVIEWED BY
Suyu Dong,
Northeast Forestry University, China
Shaodong Cao,
The Fourth Hospital of Harbin Medical University, China

*CORRESPONDENCE
Guolin Ma
maguolin1007@qq.com
Chuanchen Zhang
zhangchuanchen666@163.com

†These authors share first authorship

RECEIVED 02 August 2023
ACCEPTED 15 November 2023
PUBLISHED 30 November 2023

CITATION
Zhang D, Luan J, Liu B, Yang A, Lv K, Hu P, Han X, Yu H, Shmuel A, Ma G and Zhang C (2023) Comparison of MRI radiomics-based machine learning survival models in predicting prognosis of glioblastoma multiforme.
Front. Med. 10:1271687.
doi: 10.3389/fmed.2023.1271687

Comparison of MRI radiomics-based machine learning survival models in predicting prognosis of glioblastoma multiforme

Di Zhang[1†], Jixin Luan[2,3†], Bing Liu[2,3], Aocai Yang[2,3], Kuan Lv[4], Pianpian Hu[4], Xiaowei Han[5], Hongwei Yu[3], Amir Shmuel[6,7], Guolin Ma[3]* and Chuanchen Zhang[1]*

[1]Department of Radiology, Liaocheng People's Hospital, Shandong First Medical University & Shandong Academy of Medical Sciences, Liaocheng, Shandong, China, [2]China-Japan Friendship Hospital (Institute of Clinical Medical Sciences), Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China, [3]Department of Radiology, China-Japan Friendship Hospital, Beijing, China, [4]Peking University China-Japan Friendship School of Clinical Medicine, Beijing, China, [5]Department of Radiology, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China, [6]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada, [7]Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada

Objective: To compare the performance of radiomics-based machine learning survival models in predicting the prognosis of glioblastoma multiforme (GBM) patients.

Methods: 131 GBM patients were included in our study. The traditional Cox proportional-hazards (CoxPH) model and four machine learning models (SurvivalTree, Random survival forest (RSF), DeepSurv, DeepHit) were constructed, and the performance of the five models was evaluated using the C-index.

Results: After the screening, 1792 radiomics features were obtained. Seven radiomics features with the strongest relationship with prognosis were obtained following the application of the least absolute shrinkage and selection operator (LASSO) regression. The CoxPH model demonstrated that age (HR = 1.576, $p = 0.037$), Karnofsky performance status (KPS) score (HR = 1.890, $p = 0.006$), radiomics risk score (HR = 3.497, $p = 0.001$), and radiomics risk level (HR = 1.572, $p = 0.043$) were associated with poorer prognosis. The DeepSurv model performed the best among the five models, obtaining C-index of 0.882 and 0.732 for the training and test set, respectively. The performances of the other four models were lower: CoxPH (0.663 training set / 0.635 test set), SurvivalTree (0.702/0.655), RSF (0.735/0.667), DeepHit (0.608/0.560).

Conclusion: This study confirmed the superior performance of deep learning algorithms based on radiomics relative to the traditional method in predicting the overall survival of GBM patients; specifically, the DeepSurv model showed the best predictive ability.

KEYWORDS

glioblastoma multiforme, radiomics, machine learning, survival models, prognosis

1 Introduction

Glioblastoma multiforme (GBM) is the most common and least prognostic primary tumour of the central nervous system, with a 5-year survival rate of 6–22% based on a combination of age at diagnosis and other risk factors (1). Prognostic models that include only the patient's age, ethnicity, whether or not they receive radiotherapy, and risk factors such as the size, location and histopathological composition of the tumour often fail to predict overall survival (OS) accurately (2, 3). Therefore, identifying risk factors for GBM prognosis and developing appropriate predictive models are essential for the individualized and precise treatment of GBM patients.

Radiomics, which transforms digital medical images into mineable high-dimensional features and builds statistical models to analyze the features, has been widely used in tumour diagnosis, prognosis prediction, and treatment selection (4). Studies have shown that GBM radiomics information is closely related to patient prognosis and recurrence (5, 6). Zhang et al. (7) developed and validated a radiomics nomogram model to determine GBM survival probabilities in a non-invasive manner, achieving superior accuracy in both the training and test set. Survival analysis (also known as time-effect analysis) methods have been widely used in medical research, such as clinical efficacy trials and disease prognosis analysis. The Cox proportional-hazards model (Cox-PH) is the most well-known method used to determine the association between clinical predictor variables and the risk of mortality events. The CoxPH model is based on the assumption of a linear combination of event risk and variables; however, it is likely to be too simplistic to fit the actual disease progression.

Machine learning is a branch of artificial intelligence that has a wide range of applications in diagnosing and prognostic assessing GBM (5, 8). Compared to CoxPH models, machine learning can identify clinically significant risks with some marginal variables that can significantly improve the model's performance (9). Deep learning (DL) is a frontier area of machine learning algorithms. Deep learning-based features are mainly extracted through convolutional neural networks (CNN), and feature learning algorithms are derived from the data itself and are more targeted to specific studies (10), and are widely used in imaging diagnosis, disease staging and prognosis, which can effectively improve outcome prediction (11–13). The Deepsurv model is a deep learning technique applied to a non-linear cox proportional risk network (14). Studies have shown that the DeepSurv model can obtain patient risk factors from multiple parameters and has achieved good predictive performance in assessing different patients, such as lung cancer and nasopharyngeal carcinoma (15, 16). Previous deep-learning algorithms that have been applied to assess the prognosis of GBM patients used traditional clinical prognostic risk factors and did not incorporate radiomics features (17). To our knowledge, no study has been conducted on the prognosis of GBM patients using radiomics combined with machine learning. Therefore, this study aimed to construct: (1) a traditional CoxPH model, (2) a tree-based SurvivalTree model, (3) an RSF model based on ensemble learning, (4) a DeepSurv, and (5) a DeepHit model based on deep learning for predicting the overall survival of GBM patients based on GBM radiomics and clinical data. Following the construction of these five models, we compared their performance.

2 Materials and methods

2.1 Clinical case data

According to the proposed inclusion criteria, (1) clinical information of The Cancer Genome Atlas (TCGA) for GBM was downloaded from the TCGA database[1] and (2) Magnetic Resonance Imaging (MRI) data were obtained from the Cancer Imaging Archive (TCIA),[2] and a total of 262 patients were enrolled. Then, 131 patients were excluded due to (1) the lack of fluid-attenuated inversion recovery (FLAIR) sequences from TCIA ($n = 114$) and (2) MRI sequences acquired with severe motion or artefacts that may have induced bias in the subsequent analysis ($n = 17$). A total of 131 patients with GBM were subsequently retrospectively enrolled in our study. In this retrospective study, the requirement for informed consent was waived, as the relevant patient data in the TCGA were publicly available. We followed the relevant policies of the TCGA and TCIA in the acquisition and use of data. The flow chart for this study is shown in Figure 1.

2.2 Image acquisition and segmentation

Using ITK-SNAP[3] software to segment the FLAIR images of patients in 3D, the segmentation process is shown in Figure 2. The FLAIR scan parameters were as follows: thickness = 4 ~ 5.5 mm, TR/TE = 9,000 ~ 12,500/140 ~ 157 ms, slice gap = 4 ~ 6.5 mm, flip angle = 80 ~ 90°. The area of interest covered the entire tumour and edema region, and all feature extraction methods were implemented using the Cancer Imaging Phenomics Toolkit (CaPTk www.cbica.upenn.edu/captk). To confirm the reproducibility of the features, 30 patients were randomly selected, two people performed the Region Of Interest (ROI) segmentation, and the intraclass correlation coefficient (ICC) of the two ROIs was calculated (18). A threshold of ICC > 0.8 was set for considering a good agreement between the two neuro-radiologists. Features that achieved ICC higher than this thereshold were considered as showing reproducibility. The calculated features all contain first-order statistical features and statistical-based texture features, such as grey-level co-occurrence matrices (GLCM), grey-level dependence matrix (GLDM), neighbourhood grey-tone difference matrices (NGTDM), grey-level run-length matrices (GLRLM), and grey-level size zone matrices (GLSZ), grey-level size zone matrices (GLSZM) (19, 20).

2.3 Establishing radiomics signature and data cleaning

The least absolute shrinkage and selection operator (LASSO) method was used to select key features from the dataset significantly associated with prognosis. The selected features were linearly combined according to their respective coefficient weights to construct

1 https://tcga-data.nci.nih.gov/
2 https://wiki.cancerimagingarchive.net/
3 https://www.itk-snap.org/

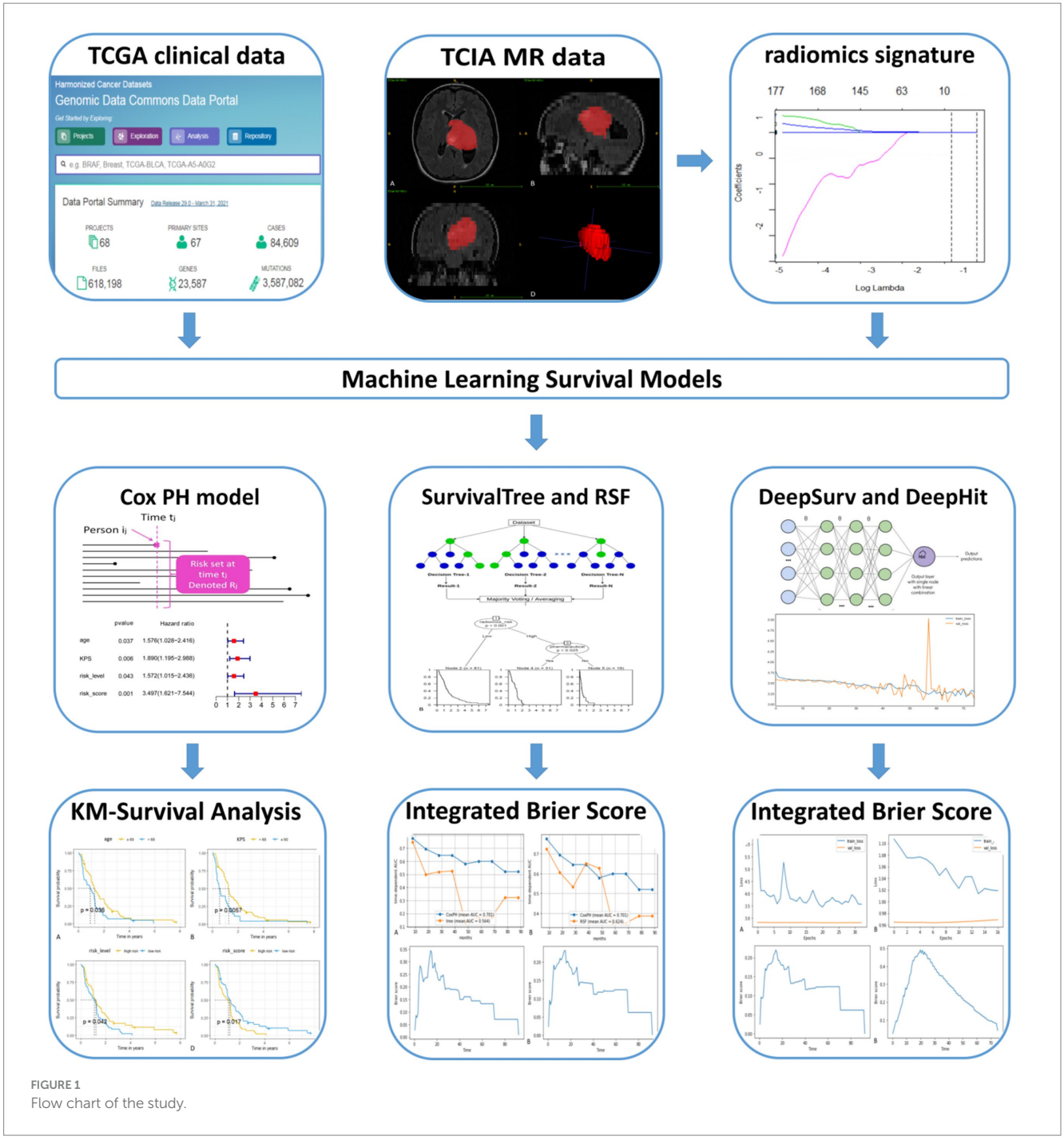

FIGURE 1
Flow chart of the study.

a radiomics signature, calculate the risk score for each patient, and determine the risk level. Subsequently, all collected data were classified as numerical or subtypes according to the input features. The missing data imputation was performed using the k-nearest neighbor (KNN) algorithm (Supplementary Table S1).

2.4 Feature engineering

According to Subtype, one-hot coding was performed to convert different categories of risk factors into categorical variables. This resulted in two new features called Subtype_Mesenchymal and Subtype_Proneural.

2.5 Construction of the model

2.5.1 CoxPH model

For the CoxPH model, proportional risk assumptions were made using the CoxPHFitter function. Filter-based feature selection was performed using Cox regression to select features significantly associated with prognosis in GBM patients. All comparisons were

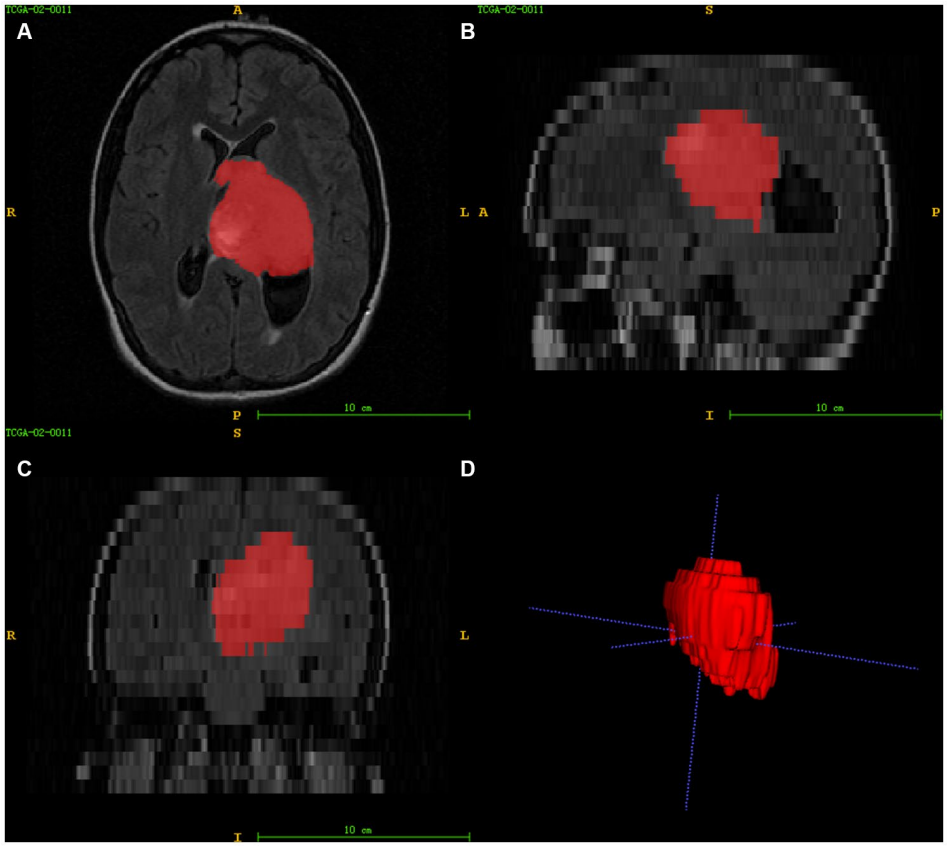

FIGURE 2
Image segmentation **(A–C)** represent the axial, sagittal, and coronal views of the images, respectively, and **(D)** shows the 3D reconstruction results of the ROI.

performed at the 95% confidence level, with $p < 0.05$ indicating statistical significance.

2.5.2 SurvivalTree model

SurvivalTree is based on classification and regression trees (CART) (21). The model is based on the tree structure, and the tree building mainly includes tree generation and pruning. Simple dichotomous classification problems can better perform the prognostic grouping of the method.

2.5.3 RSF model

Random Survival Forest is a combination of random forest (RF) and survival analysis methods. The model calculates a cumulative risk function for each tree by selecting a subset of variables at each node and splitting the node tree based on survival time and event state, and finally calculates the mean of the integrated cumulative risk function to predict the error (22).

2.5.4 DeepSurv model

DeepSurv is a feed-forward deep neural network for CoxPH models to model a nonlinear representation of the risk of clinical events based on input features. The model architecture includes network inputs from patient data, fully connected and hidden layers, and an output layer with linearly activated individual nodes for estimating the logarithmic risk function in the CoxPH model (14). DeepSurv can make predictions without specifying interaction terms, and in addition, the model's hyperparameters can vary depending on the model's performance.

2.5.5 DeepHit model

The DeepHit model was initially designed to analyze the competing risks of multiple events (23). In the present study, we considered only one event: patient survival. Therefore, we can use a simplified DeepHit model to analyze our data. We can obtain an estimated probability value with the softmax layer of the model.

2.6 Model training and evaluation

After data preprocessing, the data was divided into 70% training data and 30% test data. The hyperparameters of the models were selected via random search. The performance of the models was compared using Harrell's concordance index (C-index) and brier scores. C-index was used to estimate the proportion of random individuals with the same survival time ranking as their accurate survival time, with a C-index value of 1 indicating perfect discrimination and when 0.5 indicating random prediction. The brier score represents the mean squared difference between the observed patient status and the predicted probability of survival, with scores ranged from 0 (worst) to 1 (best). The overall estimate of the brier score for all available times is called the Integrated Brier Score (IBS). In practice, models with IBS below 0.25 are considered valuable. In

addition, for the SurvivalTree and RSF models, we also used the receiver operating characteristic curves (ROC) over time and calculated the area under the curve (AUC) values to evaluate the model performance.

2.7 Statistical analysis

Statistical analysis was performed using R 3.6.0 and the model construction was performed using Python 3.7. The R packages used are as follows: glmnet package for LASSO logistic regression, gplots and pheatmap packages for heat map analysis. The Python packages were used are as follows: CoxPH analysis using the lifelines package, SurvivalTree and RSF using the scikit-survival package, feature importance ranking using the permutation_importance function; DeepSurv and DeepHit using the Pytorch-based pycox package. The comparison of patients between training and test set was performed for continuous variables with a t-test or Mann–Whitney test. The chi-square test was performed for subtype variables. All statistics were two-tailed, and *p*-values less than 0.05 were considered statistically significant.

3 Results

3.1 Clinical characteristics of patients

The clinical characteristics of the patients in the training and test set are shown in Table 1. There were no statistically significant differences in patient age, sex, race, radiation, pharmaceutical, survival status or survival month between the training and test set ($p = 0.071–1.000$).

3.2 Radiomics feature extraction and construction of radiomics signature

In this study, 1792 radiomics features were obtained based on T2-FLAIR images from the TICA database, using CaPTk software. The 1792 features were brought into the LASSO cox regression model to screen the optimal radiomics features. We screened the optimal radiomics features in the full dataset using the LASSO Cox regression model with ten-fold cross-validation (24). We obtained seven radiomics features (three signal intensity features and four texture features) that were most closely related to the prognosis. A radiomics signature was constructed based on the linear combination of the screened seven radiomics features and their corresponding Cox regression coefficient products. The radiomics signature we constructed is described by a formula in the Supplementary Material.

3.3 Correlation between radiomics signature and clinical information

The correlation between the radiomics signature and clinical information was evaluated using heat map analysis (Supplementary Figure S1). The results show that "GLCM_Contrast_Variance" has a high correlation with survival status, mostly in red color.

3.4 CoxPH model

The univariate cox analysis showed that age (HR = 1.576, $p = 0.037$), KPS score (HR = 1.890, $p = 0.006$), radiomics risk score (HR = 3.497, $p = 0.001$), and radiomics risk level (HR = 1.572, $p = 0.043$) were prognostic factors for overall survival in GBM (Table 2), and the univariate analysis forest plot is shown in Figure 3; multivariate cox analysis showed that KPS score (HR = 1.864, $p = 0.008$), radiomics risk score (HR = 3.370, $p = 0.003$) were prognostic factors for overall survival of GBM (Table 2). In the training and test set, the C-index of the CoxPH model was divided into 0.663 and 0.635, with an overall C-index of 0.662, and for predicting 1-year, 3-year, and 5-year survival, the brier score was 0.225, 0.080, and 0.040, respectively, and the IBS was 0.102 (Table 3). The KM survival curves for variables that were significant for the univariate survival analysis are shown in Figure 4.

3.5 SurvivalTree and RSF model

GBM survival prediction models based on the SurvivalTree and RSF tree algorithms were built using the training set and validated in the test set. Figure 5 shows the AUC values of the CoxPH model, the SurvivalTree model and the RSF model over time. As can be seen from the graph, the CoxPH model has the highest AUC value of 0.701, and the SurvivalTree model has the lowest AUC of 0.564.

In the training and test set, the C-index of the SurvivalTree model was divided into 0.702 and 0.655, and the overall C-index was 0.564. For predicting 1-year, 3-year, and 5-year survival, the brier scores were 0.225, 0.080, and 0.040, respectively, and the combined brier score was 0.192. In the training and test set, the C-index of the RSF model was divided into 0.735 and 0.667, and the overall C-index was 0.642; for predicting 1-year, 3-year, and 5-year survival, the brier scores were 0.214, 0.143, and 0.124, respectively, and the IBS was 0.152 (Table 3). The IBS plots of the two models are shown in Figure 5.

The ranked importance of SurvivalTree and RSF model features are shown in Figure 6 and Supplementary Table S2; from the table, we can see that KPS, radiation and risk score are more important for the model. For both models, radiation is the most important feature, if radiation is removed from the model, the C-index of both will decrease by 0.145 and 0.101, respectively.

3.6 Deep learning model

DeepSurv and DeepHit survival prediction models based on deep learning algorithms were built using the training set and validated in the test set. In the training and test sets, the DeepSurv model had a C-index of 0.882 and 0.732, an overall C-index of 0.691, and a brier score of 0.203, 0.139, and 0.124 for predicting 1-year, 3-year, and 5-year survival, respectively, with a combined brier score of 0.116. In the training and test set, the DeepHit model had a C-index of 0.608 and 0.560, an overall C-index of 0.617, and a brier score of 0.347, 0.330, and 0.146 for predicting 1-year, 3-year, and 5-year survival, respectively, with an IBS of 0.261 (Table 3). The IBS plots for the two models are shown in Figure 7.

TABLE 1 Demographics of patients enrolled in the training set and test set.

Variables		Total (n = 131)	Training set (n = 91)	Test set (n = 40)	p
Age					0.220
	≤60	73 (56%)	47 (52%)	26 (65%)	
	>60	58 (44%)	44 (48%)	14 (35%)	
Sex					0.979
	female	44 (34%)	30 (33%)	14 (35%)	
	male	87 (66%)	61 (67%)	26 (65%)	
Race					0.462
	white	20 (15%)	12 (13%)	8 (20%)	
	others	111 (85%)	79 (87%)	32 (80%)	
KPS					0.645
	≤60	93 (71%)	63 (69%)	30 (75%)	
	>60	38 (29%)	28 (31%)	10 (25%)	
Subtype					0.742
	Classical	36 (27%)	24 (26%)	12 (30%)	
	Proneural	49 (37%)	36 (40%)	13 (32%)	
	Mesenchymal	46 (35%)	31 (34%)	15 (38%)	
CIMP_status					0.773
	G-CIMP	116 (89%)	81 (89%)	35 (88%)	
	Non G-CIMP	15 (11%)	10 (11%)	5 (12%)	
Radiation					1.000
	no	102 (78%)	71 (78%)	31 (78%)	
	yes	29 (22%)	20 (22%)	9 (22%)	
Pharmaceutical					0.454
	no	101 (77%)	68 (75%)	33 (82%)	
	yes	30 (23%)	23 (25%)	7 (18%)	
Survival status					0.071
	alive	16 (12%)	8 (9%)	8 (20%)	
	dead	115 (88%)	83 (91%)	32 (80%)	
Survival months[#]		12.27 (5.5, 19.9)	13.13 (5, 22.09)	11.71 (6.88, 17.62)	0.581

[#]Continuous variables; median (range).

TABLE 2 Univariate and multivariate cox analysis of overall survival of GBM patients.

Variables	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% CI)	p value	Hazard ratio (95% CI)	p value
Age	1.576 (1.028–2.416)	0.037	1.452 (0.943–2.235)	0.090
KPS	1.890 (1.195–2.988)	0.006	1.864 (1.175–2.956)	0.008
Risk level	1.572 (1.015–2.436)	0.043	1.041 (0.580–1.850)	0.090
Risk score	3.497 (1.621–7.544)	0.001	3.370 (1.499–7.573)	0.003

4 Discussion

Precision treatment of GBM can slow down tumour growth and help improve patient prognosis. Previous studies on GBM have used deep learning for diagnostic and prognostic assessment of tumours (17, 25). To our knowledge, this is the first study to use machine learning and radiomics approaches to assess the prognosis of GBM patients. In this study, by constructing radiomics prognostic labels, using different machine learning models and comparing the performance with the traditional CoxPH model, the results show that the DeepSurv deep learning model shows superior predictive power compared to the traditional CoxPH model.

While traditional radiography focuses on the visual presentation of images, radiomics focuses on the relationship between image

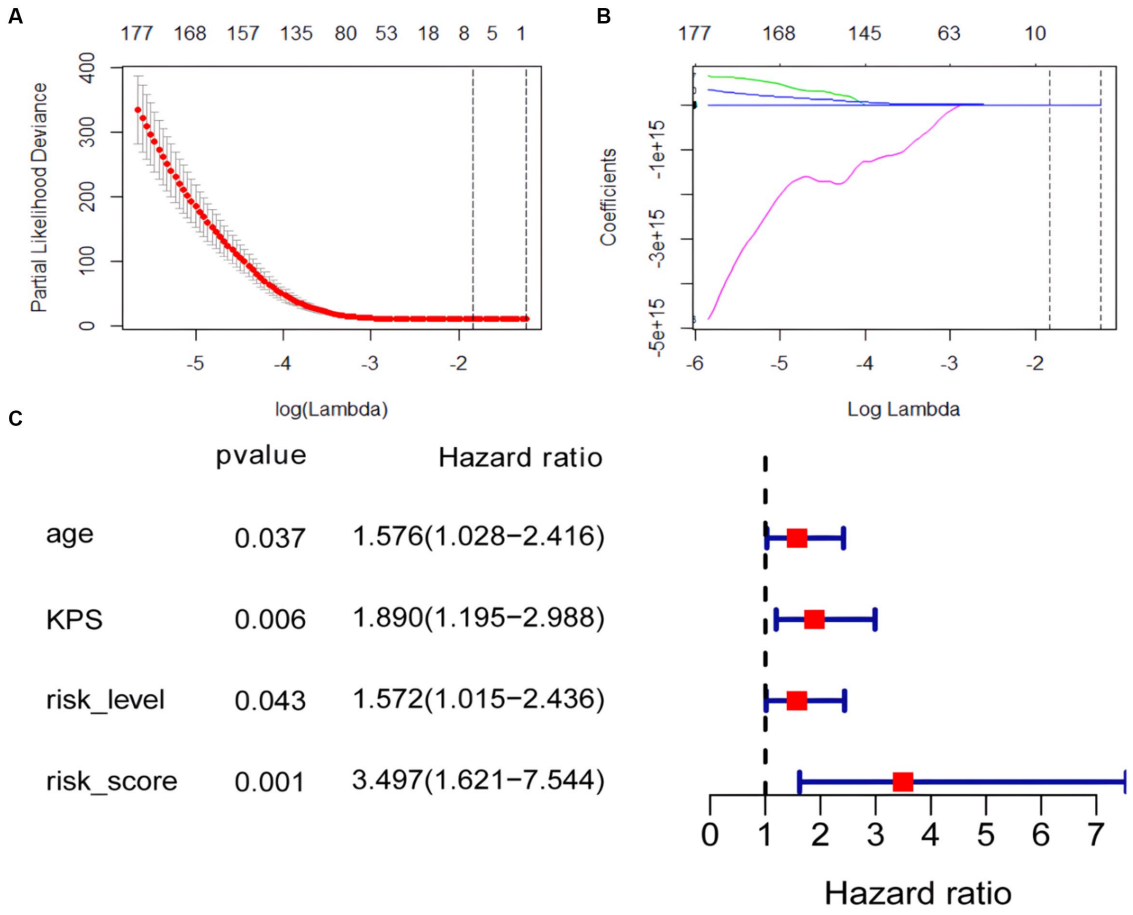

FIGURE 3
Coefficient convergence of LASSO Cox model for screening radiomics features and forest plot of univariate cox analysis. **(A)** The LASSO Cox model used tenfold cross-validation to select the optimal parameters. **(B)** The convergence of the coefficients of radiomics features under the parameters corresponding to the left figure, with each curve in the panel representing the trajectory of a feature coefficient. **(C)** Forest plot of univariate cox analysis.

TABLE 3 Hyperparameters, C-index and IBS results for the five models.

Model	C-index		Hyperparameters	C-index	Brier score			IBS
	Training set	Test set			1-year	3-year	5-year	
CoxPH	0.663	0.635	none	0.662	0.225	0.080	0.040	0.102
Survival Tree	0.702	0.655	max_depth:5,min_samples_leaf:2,min_samples_split:12,n_estimators = 10	0.564	0.263	0.190	0.133	0.192
RSF	0.735	0.667	max_features:sqrt,min_samples_leaf = 2,min_samples_split = 4,n_estimators = 10	0.642	0.214	0.143	0.124	0.152
DeepSurv	0.882	0.732	num_nodes = [32,32],out_features = 1,dropout = 0.2,learning rate = 0.005	0.691	0.203	0.139	0.124	0.116
DeepHit	0.608	0.560	num_nodes = [32,32],out_features = labtrans.out_features,dropout = 0.1,learning rate = 0.001	0.617	0.348	0.330	0.146	0.261

phenotypes and biological features and has been widely used in tumour diagnosis and prognosis evaluation (4). Studies have shown that FLAIR sequences are more advantageous in showing the extent of tumour borders and edema and that 90% of GBM recurrence occurs in the peritumoral edema area and has been shown to correlate with the prognosis of GBM (26). The FLAIR sequence was superior in showing the extent of the tumour border and edema. Some progressive patients showed no significant enhancement on the contrast scan but showed a high signal on the FLAIR sequence (27). Therefore, it is important to explore the prognostic evaluation of non-contrast FLAIR sequences in GBM. In order to construct a radiomics prognostic signature, we used the LASSO cox regression model to reduce 1792

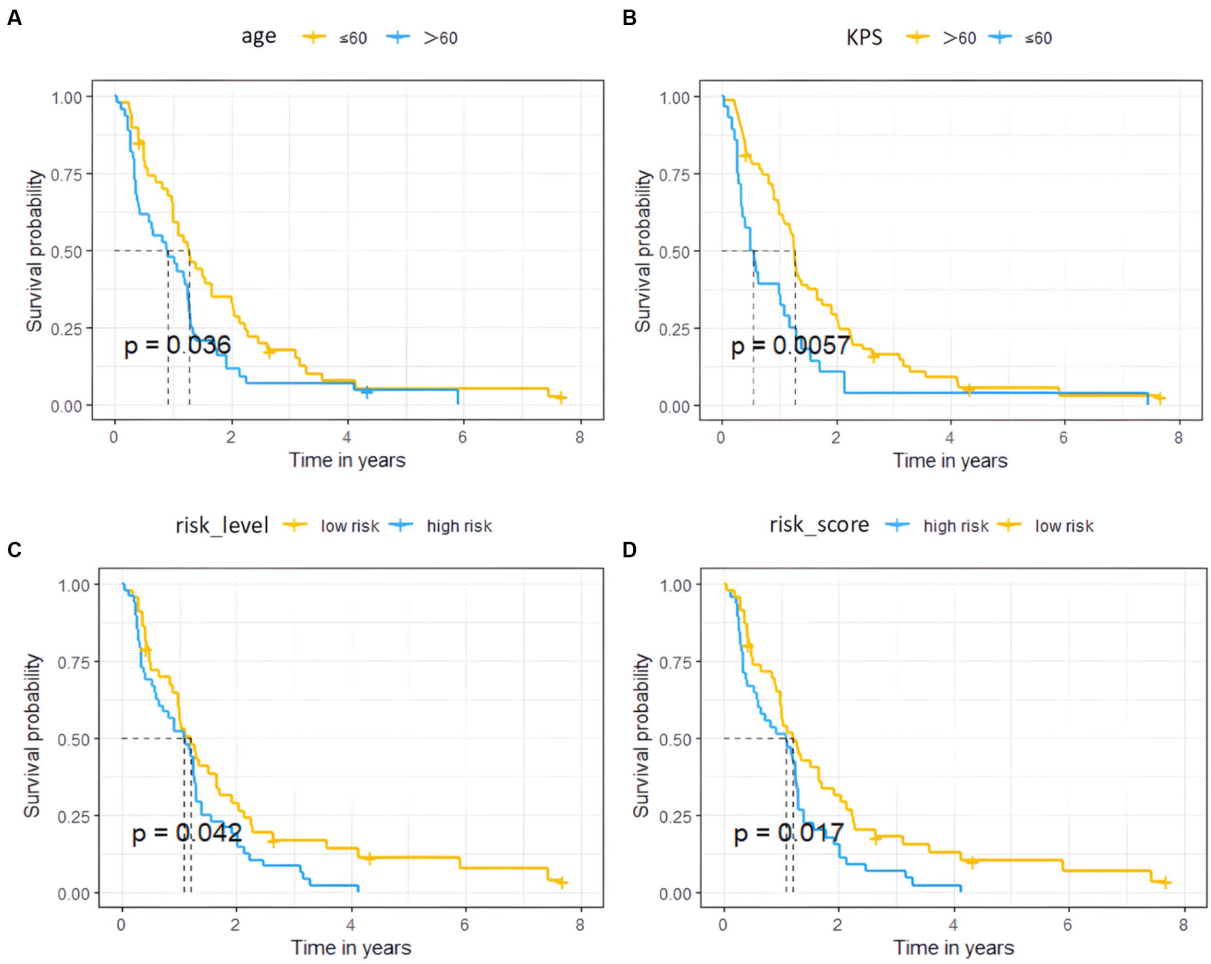

FIGURE 4
Survival curves of the high and low risk groups by univariate Cox analysis. **(A–D)** represent age, KPS, radiomics risk level, and radiomics risk score, respectively.

features to 7 potential predictive features. The results of this study showed that the seven radiomics features obtained in the FLAIR sequence were strongly associated with GBM survival, and these features indicated grey-scale heterogeneity of GBM. In addition, the radiomics risk score was shown to be an independent prognostic factor for GBM by cox univariate and multivariate analyses. The radiomics risk score was likewise the more important feature in the tree model-based feature importance ranking, suggesting that our constructed radiomics risk score can be used as a prognostic marker for GBM.

The CoxPH model is a classic approach to survival analysis and event prediction; however, the model is semi-parametric and assumes that the risk of an event is linearly related to the variables. Recently, tree-based models have received increasing attention from researchers in addressing the identification of multidimensional interactions. SurvivalTree is similar to decision trees because it is constructed by the recursive splitting of tree nodes. Compared to CoxPH, SurvivalTree is more relaxed in its requirements for survival information and does not require survival times to satisfy a specific distribution (21). RSF is a combination of random forest and SurvivalTree. The advantage of the RSF model is that it is not constrained by the assumptions of proportional risk and log-linearity (22). Also, it can prevent the overfitting problem of its algorithm through two random sampling processes (28). In our study, the SurvivalTree and the RSF model achieved a C-index of 0.70 or higher in the training set. However, as the survival tree model has fewer parameters available for adjustment and is not an integrated algorithm, it has a lower overall C-index. The IBS results for both models also showed that the RSF performed better. In addition, the AUC values for the cumulative survival times of the two models indicate a significant difference between the first and second half of the time horizon, with higher AUC values for the model in the first half of the time horizon and lower AUC values in the second half of the time horizon. Therefore, the models are most effective in predicting mid-term mortality.

Deep learning models can learn and infer higher-order nonlinear combinations between patient clinical outcomes and predictor variables in an entirely data-driven manner and have been shown to outperform standard survival analysis, with one advantage being the ability to discern complex relationships between clinical outcomes and predictor variables without prior feature selection (14). In this study, the DeepSurv model achieved the highest C-index in both the training and test set. At the same time, the overall C-index also indicated that the model was superior, suggesting that the deep learning-based

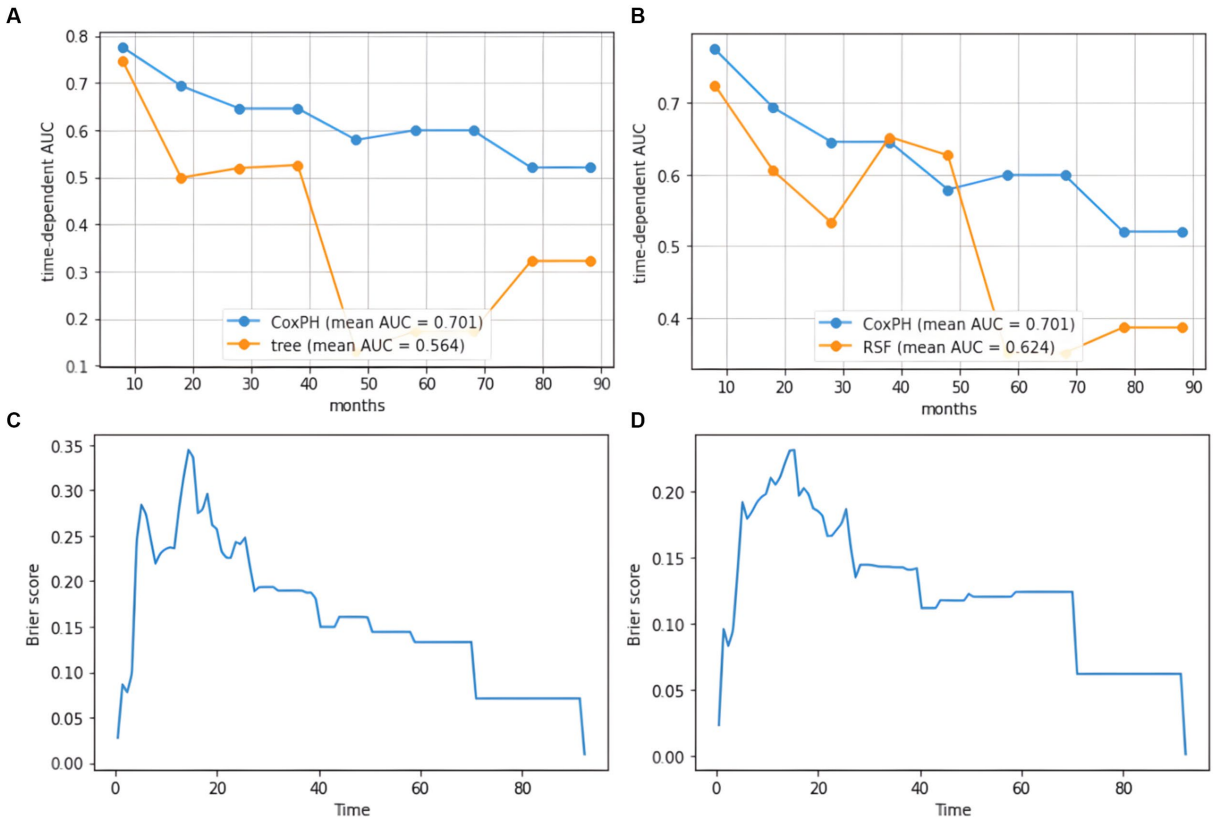

FIGURE 5
AUC plot of cumulative survival time and IBS diagram based on the tree model. **(A)** AUC results for the SurvivalTree model, **(B)** AUC results for the RSF model, **(C)** IBS results for the SurvivalTree model, **(D)** IBS results for the RSF model.

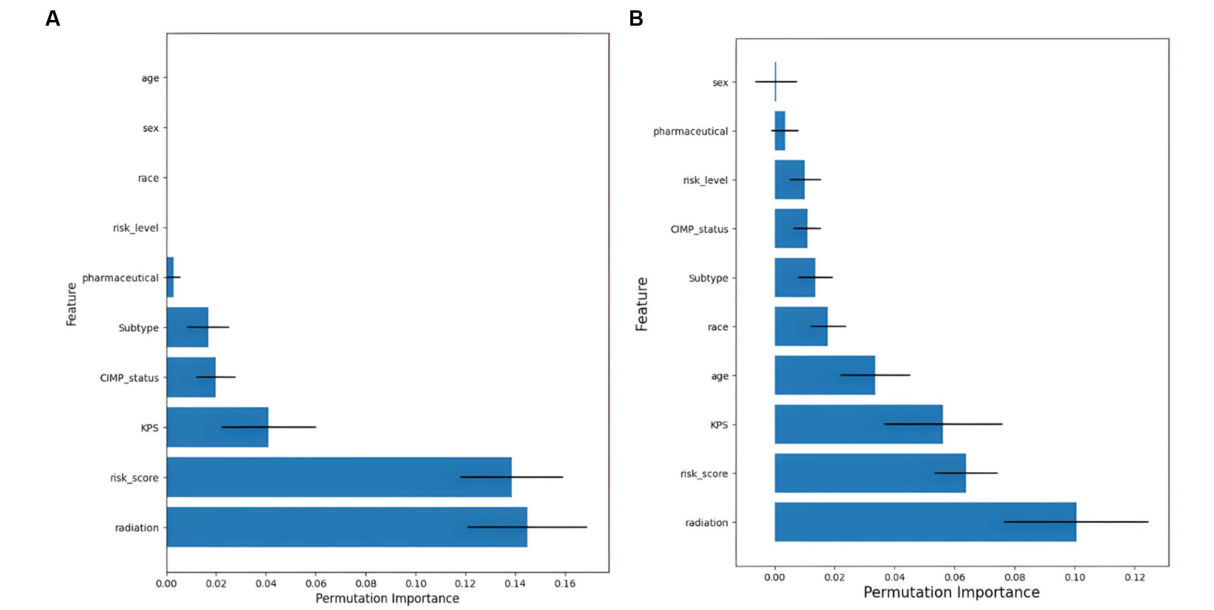

FIGURE 6
Feature importance ranking results. **(A)** Results of feature importance ranking for the SurvivalTree model, and **(B)** results of feature importance ranking for the RSF model.

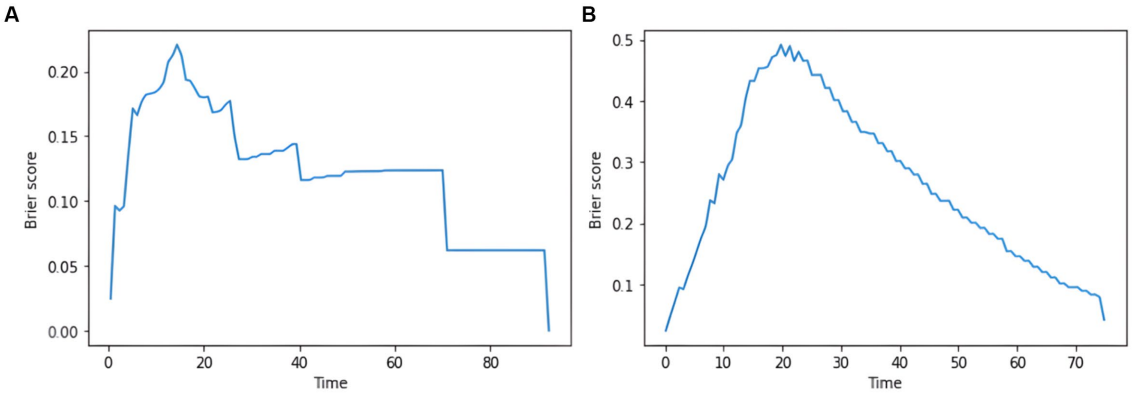

FIGURE 7
plot of the integrated brier score based on the deep learning model. **(A)** IBS results for the DeepSurv model. **(B)** IBS results for the DeepHit model.

survival model outperformed the CoxPH and RSF models in predicting GBM survival. Previous deep learning prognostic models based on clinical risk factors achieved a C-index of 0.823 and 0.700 in the training and test set, respectively (17); the present study achieved 0.882 and 0.667 in the training and test set, indicating the superior performance of the prognostic model based on radiomics features. Another deep learning model constructed in this study is DeepHit, which can directly learn the distribution of first death times and performs better in dealing with multiple competing risks (28). However, since the ending of this study is a dichotomous variable and there are no multiple competing risks, the performance of this model was not improved by hyperparameter tuning, and this model may not apply to our data structure.

There are limitations to this study. First, MRI images were collected retrospectively from the TCIA database, and the heterogeneity of different imaging parameters generated by different devices and field strengths could not be controlled. In addition, there was a relatively low number of patients in this study. Some patients also had incomplete clinical risk factors. Second, a large amount of redundant information in the sequence images leads to a considerable workload and subjectivity in manual segmentation. A more advanced approach is to use deep learning models such as CNN to learn features directly from images, which reduces the presence of subjectivity between the raters. Finally, to construct prognostic models, our study only extracted features from FLAIR images. In constructing the models, it did not make use of structural images or functional MRI techniques.

5 Conclusion

In conclusion, based on the TCGA and TCIA databases combined with a radiomics approach, this study confirmed that the DeepSurv model based on deep learning achieves better performance in GBM patient data compared to the CoxPH model. Based on the above-optimized model, a personalized treatment recommendation system for GBM can be developed to predict patient prognosis accurately.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary materials, further inquiries can be directed to the corresponding authors.

Ethics statement

The study is based on the data available in the public domain to use; therefore, no ethics statement is required for this work.

Author contributions

DZ: Data curation, Methodology, Software, Writing – original draft, Writing – review & editing. JL: Methodology, Software, Validation, Writing – original draft, Writing – review & editing. BL: Investigation, Software, Validation, Visualization, Writing – original draft. AY: Data curation, Project administration, Writing – original draft. KL: Data curation, Validation, Writing – original draft. PH: Data curation, Formal analysis, Writing – review & editing. XH: Conceptualization, Data curation, Methodology, Writing – original draft. HY: Conceptualization, Formal analysis, Investigation, Writing – original draft. AS: Writing – review & editing. GM: Funding acquisition, Supervision, Writing – review & editing, Writing – original draft. CZ: Funding acquisition, Methodology, Supervision, Writing – review & editing, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the National Natural Science Foundation of China (No. 61976110, 81971585 and 82271953), Guangzhou Science and Technology Planning Project (No. 202103010001).

Acknowledgments

The authors thank Zeshan Yao for English editing. The authors thank the TCGA platform (https://www.cancer.gov/tcga/) and the TCIA platform (https://www.cancerimagingarchive.net/) for making their data sets publicly available.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2023.1271687/full#supplementary-material

References

1. Ostrom QT, Patil N, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS. Cbtrus statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013–2017. *Neuro-Oncology*. (2020) 22:iv1–iv96. doi: 10.1093/neuonc/noaa200

2. Ferguson SD, Hodges TR, Majd NK, Alfaro-Munoz K, al-Holou WN, Suki D, et al. A validated integrated clinical and molecular glioblastoma Long-term survival-predictive nomogram. *Neurooncol Advances*. (2020) 3:vdaa146. doi: 10.1093/noajnl/vdaa146

3. Senders JT, Staples P, Mehrtash A, Cote DJ, Taphoorn MJB, Reardon DA, et al. An online calculator for the prediction of survival in glioblastoma patients using classical statistics and machine learning. *Neurosurgery*. (2020) 86:E184–92. doi: 10.1093/neuros/nyz403

4. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036

5. Lao J, Chen Y, Li Z-C, Li Q, Zhang J, Liu J, et al. A deep learning-based Radiomics model for prediction of survival in glioblastoma Multiforme. *Sci Rep*. (2017) 7:10353. doi: 10.1038/s41598-017-10649-8

6. Ammari S, Sallé de Chou R, Balleyguier C, Chouzenoux E, Touat M, Quillent A, et al. A predictive clinical-Radiomics nomogram for survival prediction of glioblastoma using Mri. *Diagnostics*. (2021) 11:2043. doi: 10.3390/diagnostics11112043

7. Zhang X, Lu H, Tian Q, Feng N, Yin L, Xu X, et al. A Radiomics nomogram based on multiparametric Mri might stratify glioblastoma patients according to survival. *Eur Radiol*. (2019) 29:5528–38. doi: 10.1007/s00330-019-06069-z

8. Bathla G, Priya S, Liu Y, Ward C, le NH, Soni N, et al. Radiomics-based differentiation between glioblastoma and primary central nervous system lymphoma: a comparison of diagnostic performance across different Mri sequences and machine learning techniques. *Eur Radiol*. (2021) 31:8703–13. doi: 10.1007/s00330-021-07845-6

9. Waljee AK, Higgins PDR. Machine learning in medicine: a primer for physicians. *Am J Gastroenterol*. (2010) 105:1224–6. doi: 10.1038/ajg.2010.173

10. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med*. (2019) 25:24–9. doi: 10.1038/s41591-018-0316-z

11. Xia W, Hu B, Li H, Shi W, Tang Y, Yu Y, et al. Deep learning for automatic differential diagnosis of primary central nervous system lymphoma and glioblastoma: multi-parametric magnetic resonance imaging based convolutional neural network model. *J Magn Reson Imaging*. (2021) 54:880–7. doi: 10.1002/jmri.27592

12. González G, Ash SY, Vegas-Sánchez-Ferrero G, Onieva Onieva J, Rahaghi FN, Ross JC, et al. Disease staging and prognosis in smokers using deep learning in chest computed tomography. *Am J Resp Crit Care Med*. (2018) 197:193–203. doi: 10.1164/rccm.201705-0860oc

13. Kim DW, Lee S, Kwon S, Nam W, Cha I-H, Kim HJ. Deep learning-based survival prediction of Oral Cancer patients. *Sci Rep*. (2019) 9:6994. doi: 10.1038/s41598-019-43372-7

14. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Deepsurv KY. Personalized treatment recommender system using a cox proportional hazards deep neural network. *BMC Med Res Methodol*. (2018) 18:24. doi: 10.1186/s12874-018-0482-1

15. Kim YJ, Lee H-J, Kim KG, Lee SH. The effect of Ct scan parameters on the measurement of Ct Radiomic features: a lung nodule phantom study. *Comput Math Method Med*. (2019) 2019:1–12. doi: 10.1155/2019/8790694

16. Liu K, Xia W, Qiang M, Chen X, Liu J, Guo X, et al. Deep learning pathological microscopic features in endemic nasopharyngeal Cancer: prognostic value and Protentional role for individual induction chemotherapy. *Cancer Med*. (2019) 9:1298–306. doi: 10.1002/cam4.2802

17. 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) 10:7048–59. doi: 10.1002/cam4.4230

18. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med*. (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012

19. Osadebey ME, Pedersen M, Arnold DL, Wendel-Mitoraj KE. Blind blur assessment of Mri images using parallel multiscale difference of Gaussian filters. *Biomed Eng Online*. (2018) 17:76. doi: 10.1186/s12938-018-0514-4

20. Guang D. An entropy interpretation of the logarithmic image processing model with application to contrast enhancement. *IEEE Trans Image Process*. (2009) 18:1135–40. doi: 10.1109/tip.2009.2016796

21. Nunn ME, Fan J, Su X, Levine RA, Lee H-J, McGuire MK. Development of prognostic indicators using classification and regression trees for survival. *Periodontol*. (2011) 58:134–42. doi: 10.1111/j.1600-0757.2011.00421.x

22. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. (2008) 2:841–60. doi: 10.1214/08-aoas169

23. Lee C, Yoon J, Mvd S. Dynamic-Deephit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Trans Biomed Eng*. (2020) 67:122–33. doi: 10.1109/tbme.2019.2909027

24. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med*. (2007) 26:5512–28. doi: 10.1002/sim.3148

25. Pasquini L, Napolitano A, Tagliente E, Dellepiane F, Lucignani M, Vidiri A, et al. Deep learning can differentiate Idh-mutant from Idh-wild Gbm. *J Pers Med*. (2021) 11:290. doi: 10.3390/jpm11040290

26. Lemée J-M, Clavreul A, Menei P. Intratumoral heterogeneity in glioblastoma: Don't forget the Peritumoral brain zone. *Neuro-Oncology*. (2015) 17:1322–32. doi: 10.1093/neuonc/nov119

27. Grossman R, Shimony N, Shir D, Gonen T, Sitt R, Kimchi TJ, et al. Dynamics of Flair volume changes in glioblastoma and prediction of survival. *Ann Surg Oncol*. (2016) 24:794–800. doi: 10.1245/s10434-016-5635-z

28. Lee C, Light A, Saveliev ES, van der Schaar M, Gnanapragasam VJ. Developing machine learning algorithms for dynamic estimation of progression during active surveillance for prostate Cancer. *NPJ Digit Med*. (2022) 5:110. doi: 10.1038/s41746-022-00659-w