



Research Paper

Deep learning features from diffusion tensor imaging improve glioma stratification and identify risk groups with distinct molecular pathway activities



Jing Yan^{a,b,1}, Yuanshen Zhao^{c,1}, Yinsheng Chen^{d,1}, Weiwei Wang^{b,e}, Wenchao Duan^f, Li Wang^e, Shenghai Zhang^c, Tianqing Ding^c, Lei Liu^c, Qiuchang Sun^c, Dongling Pei^f, Yunbo Zhan^f, Haibiao Zhao^f, Tao Sun^f, Chen Sun^f, Wenqing Wang^f, Zhen Liu^f, Xuanke Hong^f, Xiangxiang Wang^f, Yu Guo^f, Wencai Li^e, Jingliang Cheng^a, Xianzhi Liu^f, Xiaofei Lv^{g,*}, Zhi-Cheng Li^{c,h,*}, Zhenyu Zhang^{b,f,*}

^a Department of MRI, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

^b Glioma Multidisciplinary Research Group, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

^c Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

^d Department of Neurosurgery/Neuro-oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

^e Department of Pathology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

^f Department of Neurosurgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

^g Department of Medical Imaging, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

^h National Innovation Center for Advanced Medical Devices, Shenzhen, China

ARTICLE INFO

Article History:

Received 21 April 2021

Revised 1 September 2021

Accepted 3 September 2021

Available online xxx

Keywords:

Glioma

Deep learning

Diffusion tensor imaging

Prognosis

Pathway

ABSTRACT

Background: To develop and validate a deep learning signature (DLS) from diffusion tensor imaging (DTI) for predicting overall survival in patients with infiltrative gliomas, and to investigate the biological pathways underlying the developed DLS.

Methods: The DLS was developed based on a deep learning cohort (n = 688). The key pathways underlying the DLS were identified on a radiogenomics cohort with paired DTI and RNA-seq data (n=78), where the prognostic value of the pathway genes was validated in public databases (TCGA, n = 663; CGGA, n = 657).

Findings: The DLS was associated with survival (log-rank $P < 0.001$) and was an independent predictor ($P < 0.001$). Incorporating the DLS into existing risk system resulted in a deep learning nomogram predicting survival better than either the DLS or the clinicomolecular nomogram alone, with a better calibration and classification accuracy (net reclassification improvement 0.646, $P < 0.001$). Five kinds of pathways (synaptic transmission, calcium signaling, glutamate secretion, axon guidance, and glioma pathways) were significantly correlated with the DLS. Average expression value of pathway genes showed prognostic significance in our radiogenomics cohort and TCGA/CGGA cohorts (log-rank $P < 0.05$).

Interpretation: DTI-derived DLS can improve glioma stratification by identifying risk groups with dysregulated biological pathways that contributed to survival outcomes. Therapies inhibiting neuron-to-brain tumor synaptic communication may be more effective in high-risk glioma defined by DTI-derived DLS.

Abbreviations: DLS, deep learning signature; DTI, diffusion tensor imaging; WHO, World Health Organization; LGG, lower-grade gliomas; GBM, glioblastoma; M.D, mean diffusivity; FA, fractional anisotropy; AD, axial diffusivity; RD, radial diffusivity; CNN, convolutional neural networks; RNA-seq, RNA sequencing; TCIA, The Cancer Imaging Archive; TCGA, The Cancer Genome Atlas; CGGA, China Cancer Genome Atlas; GSA, Genome Sequence Archive; DEGs, differentially expressing genes; CAM, Class activation map; FDR, false discovery rate; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSA, gene set variation analysis; NRI, net reclassification improvement; AIC, Akaike information criterion; HR, hazard ratio; CI, confidence interval; NBTSC, neuron-to-brain tumor synaptic communication

* Corresponding authors at: Department of Neurosurgery, First Affiliated Hospital of Zhengzhou University, Jian she Dong Road 1, Zhengzhou, Henan, 450052, China; Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China; Department of Medical Imaging, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, 651 Dongfeng Road East, Guangzhou, Guangdong, 510060, China.

E-mail addresses: lvxf@sysucc.org.cn (X. Lv), zc.li@siat.ac.cn (Z.-C. Li), fcczhangzy1@zzu.edu.cn (Z. Zhang).

¹ Jing Yan, Yuanshen Zhao and Yinsheng Chen contributed equally to this work.

<https://doi.org/10.1016/j.ebiom.2021.103583>

2352-3964/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Funding: A full list of funding bodies that contributed to this study can be found in the Acknowledgements section.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Research in context

Evidence before this study

Recently, machine learning has been applied in extracting imaging features for prediction of clinical outcomes in glioma. Recent studies have shown that deep convolutional neural networks (CNN) can achieve state-of-the-art performance in tumor detection and diagnosis. However, there lack diffusion tensor imaging (DTI)-based CNN model for survival prediction of glioma patients, and little work has been done regarding biological underpinnings of deep CNN features. We searched published literatures on PubMed and Web of Science with the following terms: “(deep learning) AND diffusion tensor imaging AND (survival OR prognosis) AND glioma”, without date restriction or limitation to English language publications. This search did not identify any previous publications investigating the prognostic values of deep learning signature (DLS) based on diffusion tensor imaging (DTI) on glioma.

Added value of this study

In the current study, The DLS was developed from DTI for survival prediction based on a training cohort (n = 381) and a tuning cohort (n = 96), and validated on an internal validation cohort (n = 99), an external validation cohort (n = 77), and a public TCIA cohort (n = 35). Incorporating the DLS into existing risk system resulted in a deep learning nomogram predicting survival better than either the DLS or the clinicomolecular nomogram alone. Furthermore, five kinds of pathways (synaptic transmission, calcium signaling, glutamate secretion, axon guidance, and glioma pathways) underlying the DLS were identified on a radiogenomics cohort with paired DTI and RNA-seq data (n=78). Average expression value of pathway genes showed prognostic significance in our radiogenomics cohort and validated in public databases (TCGA, n = 663; CGGA, n = 657).

Implications of all the available evidence

This study demonstrated DTI-derived DLS, which associated with dysregulated pathways, was an independent prognostic factor conferring incremental value over clinicomolecular factors in survival prediction. DTI-derived DLS provides a noninvasive approach to stratify glioma patients and offers molecular signatures to inform personalized treatment. Therapies inhibiting neuron-to-brain tumor synaptic communication may be more effective in high-risk glioma defined by DTI-derived DLS.

across individuals [5,6]. Hence, accurate prediction of clinical outcomes can provide social benefit and information for optimizing personalized treatment of glioma patients.

Although conventional MRI can demonstrate anatomic parameters such as size, shape, and morphological features of the tumor, it is limited in delineating microscale tumor infiltration. Diffusion tensor imaging (DTI) is a promising imaging approach to detect microstructural tissue changes of the whole tumor by assessing the water diffusion in vivo. DTI has been demonstrated sensitivity to tumor infiltration that is not evident on conventional MRI. Various DTI metrics such as mean diffusivity (MD), fractional anisotropy (FA), axial diffusivity (AD), and radial diffusivity (RD) have been shown predictive of tumor progression and survival outcomes in LGG and GBM [7–11]. However, earlier works mainly focused on semiquantitative DTI metrics or histogram analysis, which may not get the utmost out of all the information embedded in such images.

Recently, machine learning methods have been applied in extracting imaging features for prediction of clinical outcomes in glioma [6,12–17]. Specifically, there are two most popular imaging-based machine learning approaches: handcrafted radiomics analysis and convolutional neural networks (CNN). Radiomics features extracted from MRI have shown predictive of survival in glioma [13,14]. However, handcrafted radiomics features are constricted by current understanding of medical imaging and therefore may limit the potential of the prediction model. CNNs improved the handcrafted radiomics pipeline by automatically learning discriminative features directly from images. Recent studies have shown that deep CNNs can achieve state-of-the-art performance in tumor detection and diagnosis, compared with other machine learning approaches and even human experts [18,19]. However, few studies have focused on the prognostic value of DTI-based CNN in survival prediction for glioma patients.

Notwithstanding its predictive power, the data-driven nature of CNN has led to its inherent lack of biological interpretability of the learned deep features. In contrast with conventional biomarkers driven by biological hypotheses, the biological meaning of the deep CNN features that are predictive of patient outcomes remains unclear. Without biological basis, such black box-like property of deep CNN becomes a clear obstacle towards its wide application in practice. A few pioneer studies have initially revealed the connections between radiomics features and underlying gene expression patterns [6,20], but to our knowledge little work has been done regarding biological underpinnings of deep CNN features used for survival prediction in glioma. Therefore, this study hypothesized that deep CNN features learned from DTI were predictive of survival outcomes in glioma patients, and might be genetically driven by different biological pathways that contributed to cancer prognosis. To this end, the aims of this multicenter study were to develop and validate a deep learning model from DTI for predicting survival of glioma patients, and to uncover the biological meaning of the prognostic deep CNN features by identifying their underlying biological pathways using paired DTI and RNA sequencing (RNA-seq) data.

2. Methods

2.1. Study design

This study was a part of the registered clinical trial “MR Based Survival Prediction of Glioma Patients Using Artificial Intelligence”

1. Introduction

Gliomas are primary brain tumors originating from glial or precursor cells [1]. The newest World Health Organization (WHO) classification of CNS tumors has classified gliomas into four grades, and WHO II-IV gliomas are considered as infiltrative gliomas [2,3]. Notably, precise prediction of the clinical outcomes of infiltrative gliomas is challenging [3]. As for lower-grade gliomas (LGG, WHO II or III), some relapse or progress to WHO IV glioblastoma (GBM) after treatment within several months, while some others remain indolent for several years [4]. On the other hand, the heterogeneity of GBM also leads to largely varied prognosis

(ClinicalTrials.gov ID: [NCT04215211](https://clinicaltrials.gov/ct2/show/study/NCT04215211)). This study was approved by the Human Scientific Ethics Committee of the First Affiliated Hospital of Zhengzhou University (No. 2019-KY-176) and the Sun Yat-Sen University Cancer Center (B2019-085-01). The overall design of our study included two steps: prognostic deep CNN modeling and radiogenomics profiling, as illustrated in Fig. 1. First, an imaging-based deep learning signature (DLS) was developed from DTI for survival prediction based on training/tuning cohorts and validated on an internal validation cohort and two external validation cohorts. Then, the key biological pathways underlying the DLS were identified based on a radiogenomics dataset with both DTI and RNA-seq, where the prognostic value of the pathway genes was validated in three public cohorts.

2.2. Study cohorts

Informed consents were obtained from patients whose fresh tumor specimens were used for RNA-seq. For the rest patients, informed consents were waived by the Committee due to the retrospective and anonymous nature of this study. There were three datasets in this study: a deep learning dataset (n = 688) with DTI imaging for training and validating the DLS, an independent radiogenomics analysis dataset (n = 78) with paired DTI and RNA-seq for identifying biological pathways underlying the deep learning features, and a public radiogenomics validation dataset (n = 1320) with only RNA-seq data for further validating the prognostic value of the DLS-associated pathway genes. These datasets were collected from two local institutions the First Affiliated Hospital of Zhengzhou University (FAHZZU) and Sun Yat-Sen University Cancer Center (SYSUCC) between January 2012 and December 2018 and three public databases The Cancer Imaging Archive (TCIA), The Cancer Genome Atlas (TCGA), and China Cancer Genome Atlas (CGGA). The inclusion criteria are summarized in Supplementary A1 and the patient enrolment process is shown in Fig. 2. Specifically, the deep learning dataset comprised five cohorts: a (1) training cohort (n = 381, from FAHZZU) and a (2) tuning cohort (n = 96, from FAHZZU) used to develop the DLS, an (3) internal validation cohort (n = 99, from FAHZZU) and an (4)

external validation cohort (n = 77, from SYSUCC) and a (5) public validation cohort (n = 35, from TCIA) used to validate the DLS. Note that the training, tuning, and internal validation cohorts were randomly selected from the FAHZZU patient set, where the clinical parameters among these cohorts were balanced. The radiogenomics analysis dataset comprised 78 patients from FAHZZU (not included in the deep learning dataset) with paired DTI and RNA-seq data. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive in National Genomics Data Center under accession number HRA000802 (<https://bigd.big.ac.cn/gsa-human/browse/HRA000802>). The public radiogenomics validation dataset contains RNA-seq data only, including a LGG dataset of 509 lower-grade gliomas patients from TCGA, a GBM dataset of 154 GBM patients from TCGA, and a glioma dataset of 657 patients from CGGA. Detailed information on RNA sequencing, Detection of IDH mutation, Image acquisition and preprocessing was described in Supplementary A2-A5 and Supplementary Table S1.

2.3. Deep learning signature building

A deep CNN model was used to perform the survival analysis. The architecture of the deep CNN was a ResNet-34-based network [21], as illustrated in Fig. 1. The network input were axial slices cropped from the four registered maps FA, MD, AD and RD. To ensure that only slices within tumors were used as input, a 3D bounding box containing just the entire tumor was derived based on the delineated tumor contours for each patient. To represent the entire tumor in general, 4 equally-spaced axial slices within the tumor were extracted from each of the 4 DTI maps. Then, the 4 slices were cropped into small ones using the bounding box. Finally, 16 cropped slices per patient were automatically generated from the 4 DTI maps and used as a single sample for a 3D tumor in both training and validation. The network was trained from scratch on the training cohort (n = 381, 6096 images) while optimized on the tuning set (n = 96, 1536 images). The network output was the predicted risk score regarding the overall survival of the input patients, which was used

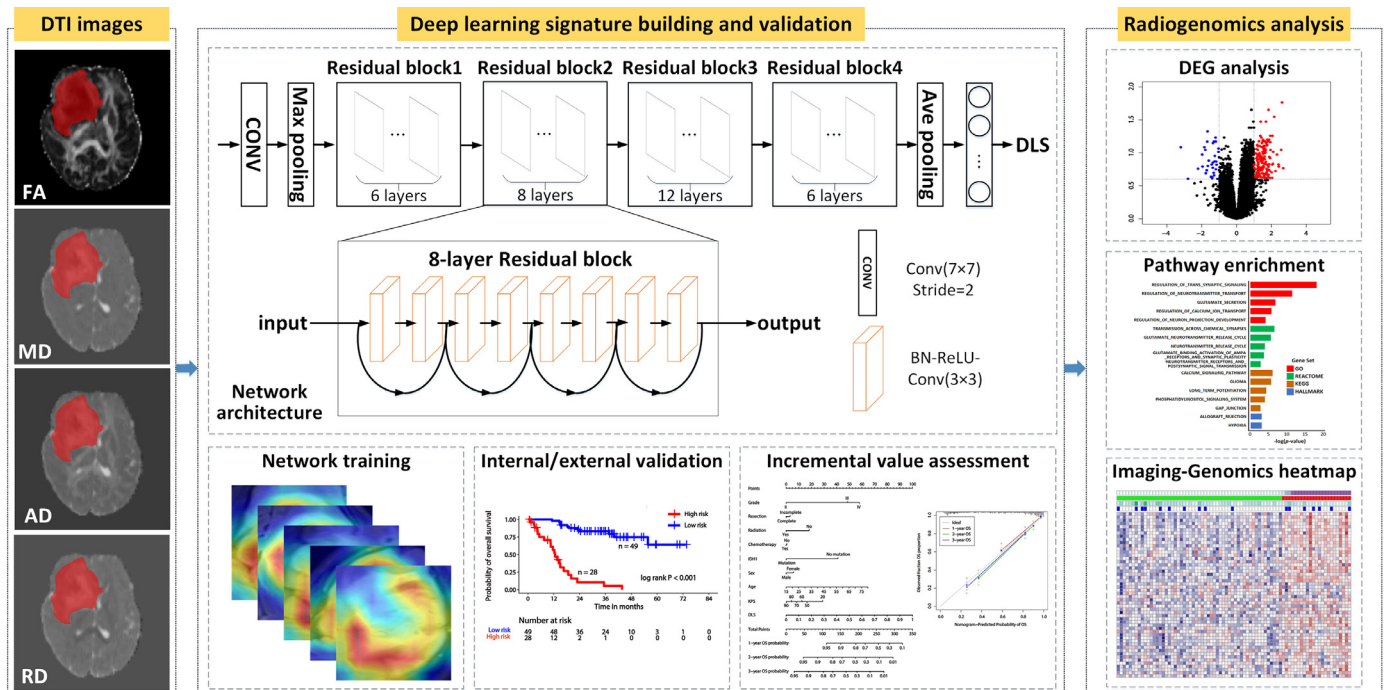


Fig. 1. The overview of the study design, including the deep learning signature (DLS) development and validation, and the radiogenomics analysis.

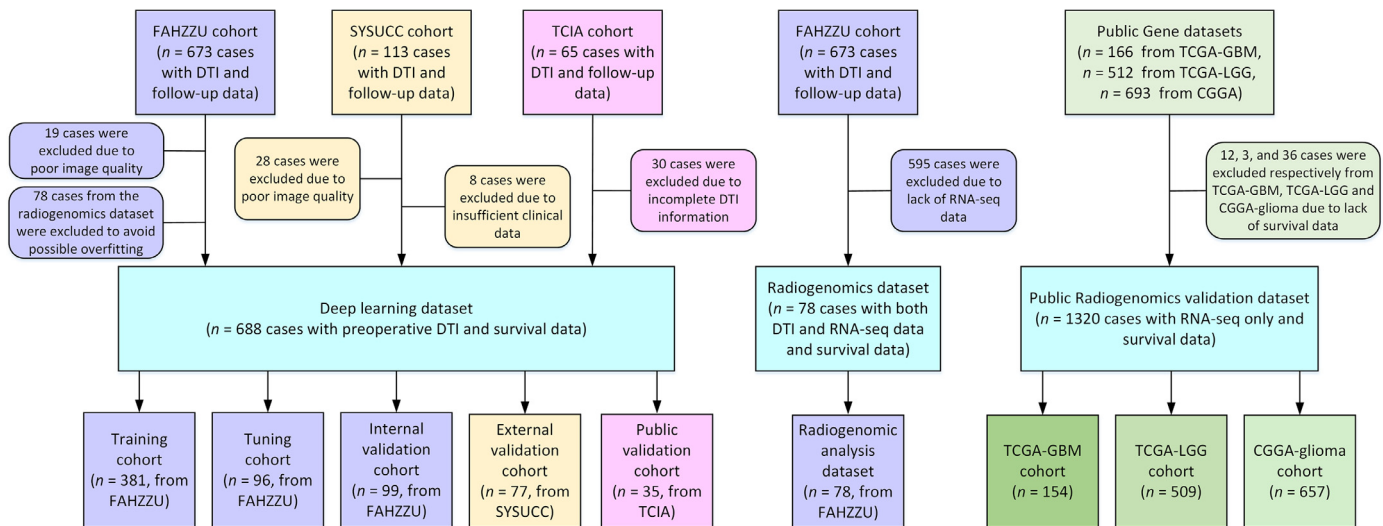


Fig. 2. Patient enrollment process for the three datasets.

as the DLS for survival prediction. The details of the network training can be found in Supplementary A6.

2.4. Identification of biological pathways associated with DLS

Based on the radiogenomics analysis dataset with both MRI and RNA-seq, the possible biological pathways underlying the DLS were identified. First, differentially expressing genes (DEGs) between the high- and low-risk subgroups stratified by the DLS were identified with an R package DESeq2. Then, significant DEGs with false discovery rate (FDR) < 0.25 and $|\log_2(\text{Fold Change})| > 0.10$ were analyzed to enrich overrepresented pathways with an R package clusterProfiler based on four annotated databases: Gene Ontology (GO) Biological Process, Kyoto Encyclopedia of Genes and Genomes (KEGG), Hallmark, and Reactome. FDR < 0.05 was considered as significant enrichment. Then, a gene set variation analysis (GSVA) was performed for each enriched pathway to calculate a patient-specific GSVA score that quantified the pathway activity [22]. A Pearson correlation was used to assess if the pathway GSVA score was significantly associated (FDR < 0.01) with the DLS. Finally, the significantly correlated pathways were used to biologically annotate the DLS.

2.5. Statistics

Validation of the DLS: Statistical analysis was performed using R version 3.6.1. P-value < 0.05 was considered significant. The patient and tumor characteristics between training and validation cohorts were assessed by Wilcoxon test or Chi-square test. The association of the DLS with survival was first assessed in the training cohort and then validated in the tuning, internal validation, external validation, and public TCIA cohorts by using Kaplan-Meier analysis. According to a DLS cutoff value determined using the X-tile on the training cohort [23], patients were stratified into low-risk and high-risk subgroups. The cutoff was applied to the tuning, internal, external and public TCIA cohorts. A weighted log-rank test (the G-rho rank test, $\rho = 1$) was used to validate the significant difference in the survival between the risk subgroups. The assessment of the DLS as an independent prognostic factor was performed by integrating clinical risk factors such as gender (female or male), age, grade (II, III or IV), pre-operative KPS, extent of resection (complete or incomplete), radiation therapy (yes or no), chemotherapy (yes or no), IDH status (mutated or wild-type) into the multivariate Cox proportional hazard model.

Incremental prognostic value of the DLS: To demonstrate the incremental prognostic value of the DLS over the clinicomolecular risk

factors for individualized assessment of survival, both a clinicomolecular nomogram and a deep learning nomogram was constructed based on the training cohort. The clinicomolecular nomogram consisted of age, gender, KPS, grade, extent of resection, radiation therapy, chemotherapy and IDH mutation. The deep learning nomogram was built by incorporating the DLS into the clinicomolecular nomogram based on Cox analysis. Then, the incremental prognostic value of the DLS was assessed by comparing the performance of the two nomograms in terms of discrimination, calibration, reclassification and clinical usefulness. First, the Harrell C-indices of the DLS and the two nomograms were calculated as the discriminative measure. Then, the calibration curves of the two nomograms were plotted to validate the agreement between the predicted and observed outcomes. The net reclassification improvement (NRI) was calculated to assess the performance improvement added by the DLS. The Akaike information criterion (AIC) was computed to assess the potential risk of model overfitting. The decision curve analysis was performed to validate the clinical usefulness of the prediction models.

Prognostic value of the DLS-correlated pathway genes: To further demonstrate the DLS-pathway-prognosis linkage, the collective prognostic value of the DLS-correlated pathways was assessed by Cox regression. Specifically, the association of the average expression value of the genes contained in the DLS-correlated pathways and the patient survival was assessed by using Kaplan-Meier analysis. A cutoff value was determined by using X-tile tool on the radiogenomics analysis cohort and was used to stratify patients into two subgroups. This cutoff was consistently applied to the three public RNA-seq datasets including the TCGA-LGG, TCGA-GBM, and CGGA-glioma.

2.6. Role of funding source

All sources of funding have been declared as an acknowledgment at the end of the manuscript. The funders did not play any role in research design, data collection, data analysis, interpretation, report writing and implementation supervision. All authors confirmed that they had full access to all the data in the study and accepted responsibility to submit for publication.

3. Results

3.1. Patient characteristics

According to the selection criteria, a total of 688 patients were included in the deep learning dataset for DLS training and validation.

As shown in Supplementary Table S2, there was no significant difference in survival between training cohort and validation cohorts (Mean survival: training cohort, 25.2 months; tuning cohort, 26.3 months; internal validation cohort, 26.4 months; external validation cohort, 27.3 months; public TCIA cohort, 22.6 months, log-rank P-value > 0.05). The distribution of clinical characteristics (grade, gender, age, KPS, chemotherapy, radiation, extent of resection, IDH mutation) was also balanced between the training and validation cohorts (Chi-square or Wilcoxon P-value > 0.05).

3.2. Association of the DLS with survival

The C-index for the DLS was 0.825 in the training cohort, 0.745 in the tuning cohort, 0.746 in the internal validation cohort, 0.794 in the external validation cohort, and 0.789 in the public

TCIA cohort. The optimum cutoff value was 0.14, which divided the patients into a high-risk subgroup (DLS ≥ 0.14) and a low-risk subgroup (DLS < 0.14). The results of Kaplan-Meier analysis were shown in Fig. 3a-f. Significant association of DLS with survival was found in the training cohort (log-rank P < 0.001; hazard ratio [HR] = 11.850, 95% confidence interval [CI]: 7.931, 17.700), and was confirmed in the tuning cohort (log-rank P < 0.001; HR = 6.623, 95% CI: 3.168, 13.840), the internal validation cohort (log-rank P < 0.001; HR = 4.471, 95% CI: 2.204, 9.071), the external validation cohort (log-rank P < 0.001; HR = 8.340, 95% CI: 6.540, 18.430), the public TCIA cohort (log-rank P < 0.001; HR = 10.180, 95% CI: 1.551, 39.790), and the radiogenomics analysis dataset (log-rank P < 0.001; HR = 8.154, 95% CI: 2.104, 21.600). The DLS was identified as an independent risk factor (HR = 9.169, 95% CI: 6.888, 12.200, P < 0.001).

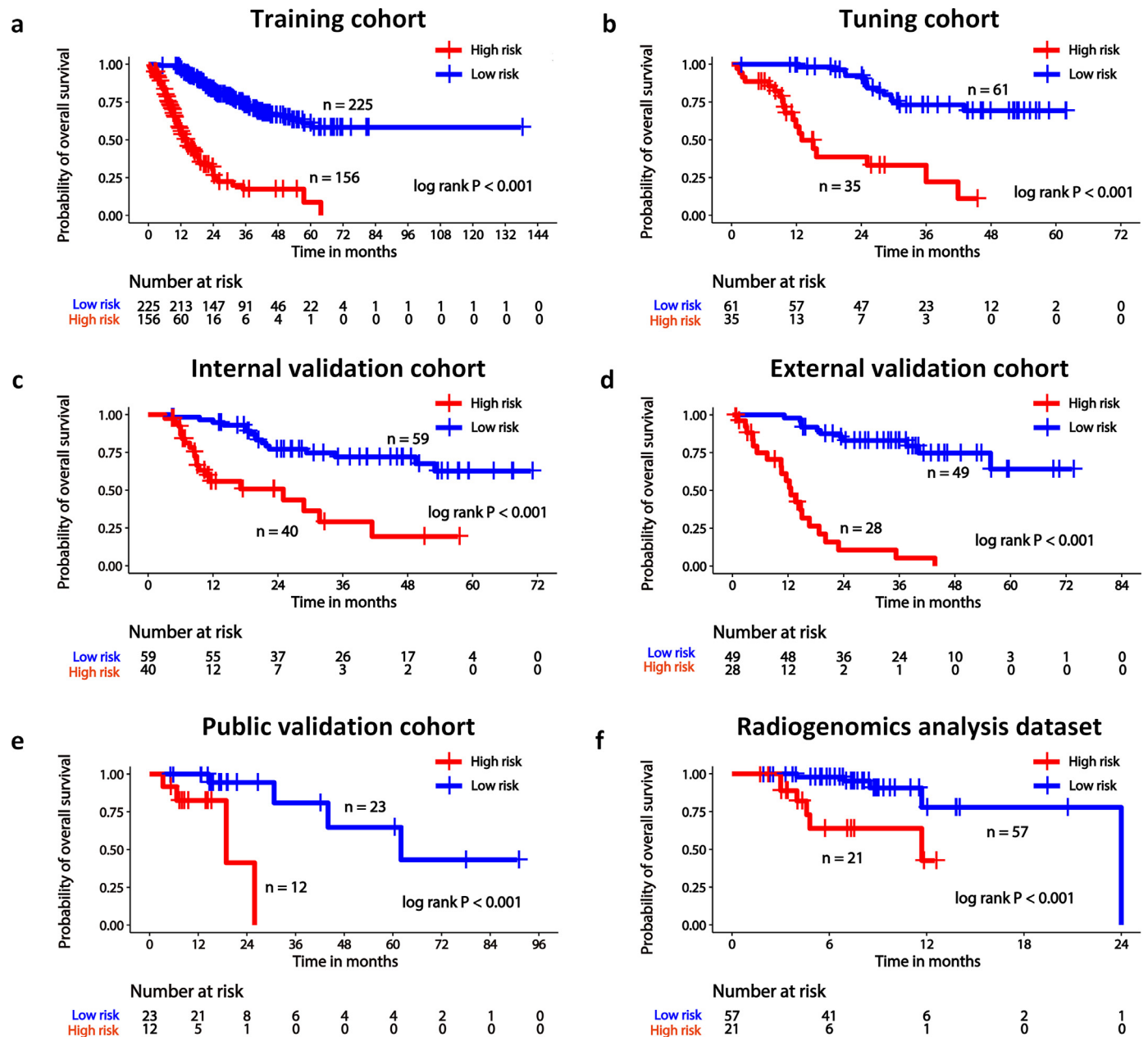


Fig. 3. Kaplan-Meier analysis according to the deep learning signature (DLS) for overall survival in the training (a), tuning (b), internal validation (c), external validation (d), and public validation (e) cohorts, as well as the radiogenomics analysis dataset (f). Significant associations of DLS with overall survival were demonstrated. The numbers of patients at risk for each time interval are shown in the bottom of each plot.

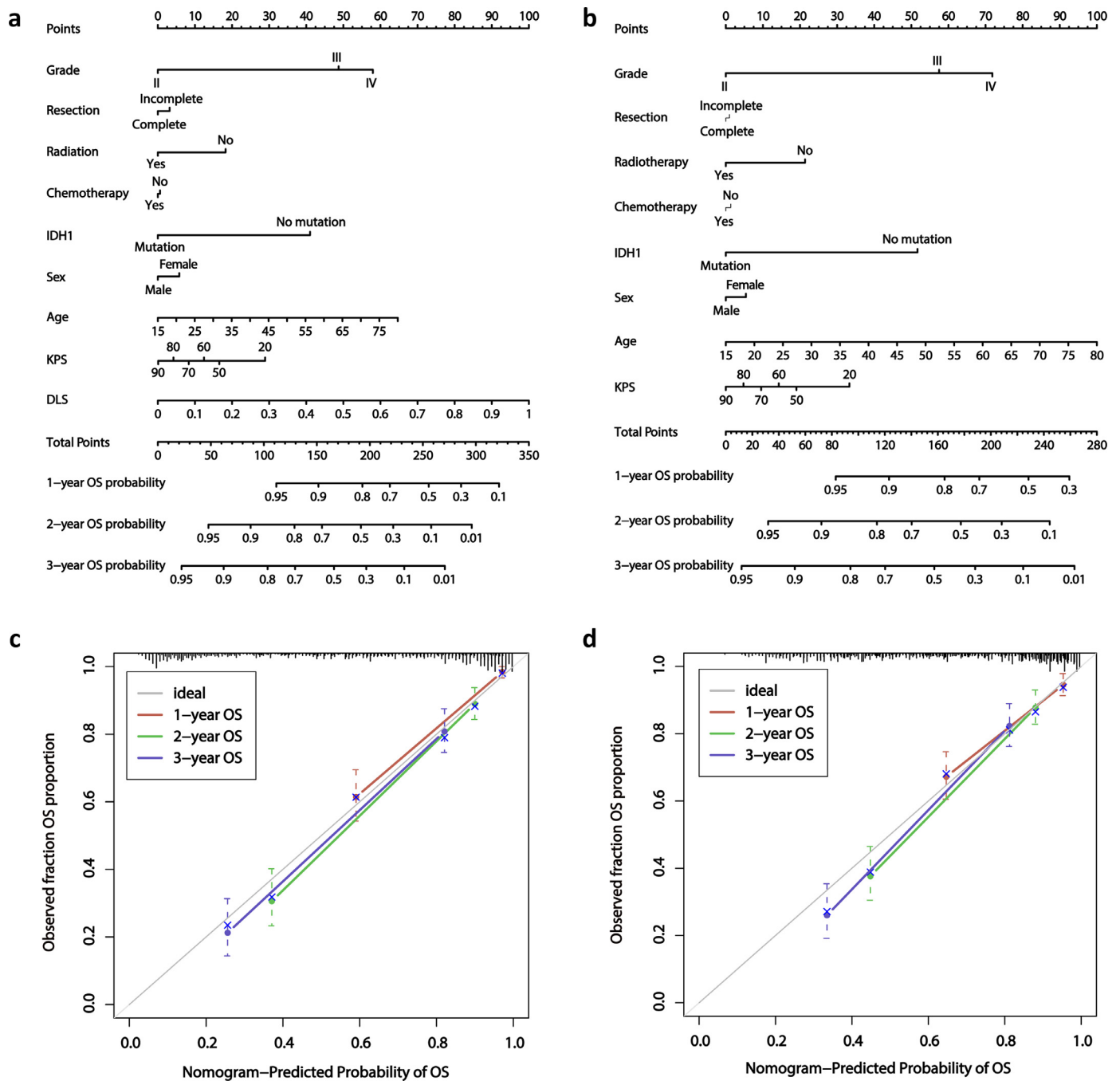


Fig. 4. The deep learning nomogram (a) and the clinicomolecular nomogram (b) for predicting the 1-, 2-, and 3-year overall survival outcomes, along with the calibration curves for evaluation of the deep learning nomogram (c) and the clinicomolecular nomogram (d).

3.3. Assessment of the incremental prognostic value of the DLS

The clinicomolecular nomogram and deep learning nomogram for individual survival prediction were shown in Fig. 4a-b, respectively. The C-indices and AIC values for the two nomograms and the DLS were summarized in Table 1. The clinicomolecular nomogram achieved a C-index of 0.805 in the training cohort, 0.838 in the tuning cohort, 0.791 in the internal validation cohort, and 0.771 in the external validation cohort. Integrating the DLS into the clinicomolecular nomogram resulted in an improved C-index of 0.835 in the training cohort, 0.890 in the tuning cohort, 0.840 in the internal validation cohort, and 0.903 in the external validation cohort. The deep learning nomogram had lower AIC values,

indicating its better reliability against overfitting. The calibration curves for both nomograms for the probability of 1-, 2-, or 3-year death were shown in Fig. 4c-d, respectively. The calibration curve of the deep learning nomogram demonstrated better agreement between the predicted and observed survival. Incorporating the DLS into the clinicomolecular nomogram generated a total NRI of 0.646 (95% CI: 0.552, 0.773, $P < 0.001$) regarding the survival prediction, indicating an improved classification performance of the resulted deep learning nomogram. The decision curves showed in Supplementary Figure S1 validated the clinical usefulness of the prediction models, indicating that the deep learning nomogram added more benefit than either the clinicomolecular nomogram or the DLS.

Table 1.

The C-indices and Akaike information criterion (AIC) values for survival prediction using the imaging-based deep learning signature (DLS), the clinicomolecular (CM) nomogram and the deep learning (DL) nomogram in the training, tuning, internal validation and external validation cohorts, respectively.

Model	Index	Training	Tuning	Internal validation	External validation
DLS	C-index	0.825 (0.794, 0.856)	0.745 (0.659, 0.831)	0.746 (0.675, 0.817)	0.794 (0.725, 0.863)
	AIC	1450	251	278	206
CM nomogram	C-index	0.805 (0.732, 0.810)	0.838 (0.774, 0.903)	0.791 (0.710, 0.871)	0.771 (0.714, 0.896)
	AIC	1471	239	273	227
DL nomogram	C-index	0.835 (0.806, 0.865)	0.890 (0.845, 0.935)	0.840 (0.785, 0.895)	0.903 (0.859, 0.946)
	AIC	1404	221	261	194

3.4. Identification of biological pathways associated with DLS

In the radiogenomics analysis cohort (44 male and 34 female, age range: 18–72 years, median age: 48 years) with both DTI and RNA-seq, 207 DEGs differentially expressed between risk subgroups stratified by the DLS were identified, as listed in Supplementary Table S3 and shown by a volcano plot in Fig. 5a. The enrichment analysis based on the DEGs identified the key biological pathway, as shown in Fig. 5b. A complete list of enriched pathways with FDR < 0.01 was provided in Supplementary Table S4. The DLS was found to be significantly correlated with misadjusted GO annotations and signaling pathways related to chemical synaptic transmission/neurotransmitter transport, calcium transport/signaling, glutamate secretion/glutamate binding activation of AMPA receptors, neuron projection development/axon guidance, and glioma pathways, as shown in Fig. 6a and Supplementary Table S5. The average expression value of these DLS-related pathway genes succeeded to stratify the radiogenomics analysis cohort into two risk subgroups (log-rank P = 0.018, HR = 2.741, 95% CI: 1.017, 7.388) with a cutoff value of 29.31. The prognostic power of these DLS-related genes were further confirmed on the TCGA-LGG dataset (log-rank P < 0.001, HR = 1.036, 95% CI: 1.015, 1.058), the TCGA-GBM dataset (log-rank P = 0.025, HR = 2.105, 95% CI: 1.998, 2.213), and the CGGA-glioma dataset (log-rank P = 0.008, HR = 1.056, 95% CI: 1.008, 1.103), as shown by the Kaplan-Meier curves in Fig. 5c. To further reveal the DLS-pathways-survival linkage, the class activation maps (CAMs) of the DLS with corresponding FA, MD, AD and RD images of four representative patients classified into high- and low-risk subgroups were presented in Fig. 6b. These CAMs indicated that the proposed deep CNN model could highlight certain risky regions that may be relevant to tumor prognosis while suppress other less relevant regions. The heatmap-like display allowed assessing the region of risk with potential prognostic value on each DTI-derived map such as FA, MD, AD and RD. Furthermore, we found that higher mean FA and lower mean MD, AD and RD within the highlighted regions could be found in the high-risk subgroup than those in the low-risk subgroup, as shown by the boxplots in Fig. 6c. Moreover, the results of DEGs in the radiogenomics dataset (n = 78) showed the expressions of representative genes such as SNAP25 and KIF5A (core genes of chemical synaptic transmission/neurotransmitter transport pathways), PRKCB and CAMK2A (core genes of calcium signaling and glioma pathways) in the low-risk subgroups were significantly lower than those in the high-risk subgroup, as shown by the boxplots in Fig. 6d.

4. Discussion

In this multicenter study, we developed and validated a deep learning prognostic signature using DTI metrics for improving the survival prediction of glioma patients, and further revealed the key biological pathways underlying the deep imaging features. The major findings of our study were that (1) DTI-derived DLS can offer incremental prognostic value beyond traditional clinical parameters and IDH mutation status in prediction of overall survival for glioma

patients, and (2) prognostic deep features learned from DTI metrics were associated with biological pathways involved in synaptic transmission, calcium transport, glutamate secretion/binding activation of AMPA receptors, neuron projection development/axon guidance, and glioma pathways.

Several studies have presented radiomics models to predict the clinical outcomes of gliomas from MRI, such as the radiomics analysis on 233 LGG patients [15] and the radiomics model on 217 GBM patients [14]. Deep learning approach improved the handcrafted radiomics pipeline by learning discriminative image features on its own. However, deep CNNs usually require a large set of labelled training images before they can achieve acceptable performance. For example, 730 patients with gastric cancer were recruited to develop a deep learning predictor from CT imaging for prediction of lymph node metastasis [24]. In another study, preoperative MRI from 1163 patients were collected to train and validate a CNN for renal tumor classification [19]. So far, studies are still limited regarding imaging-based CNN for prediction of glioma survival. Lao et al. developed a machine learning model combining both radiomics and deep features from standard MRI for survival prediction in GBM on a small cohort of 112 patients [16]. Yoon HG et al. trained a deep CNN from 88 GBM patients for survival prediction and tested its performance in 30 patients [17]. Previous studies built their prediction model using conventional MR sequences (T1, T1c, T2, Flair) with limited sample size, while the use of CNN for survival prediction on DTI has not been investigated. Our study enrolled 688 glioma patients with preoperative DTI to develop and validate a deep learning signature named DLS for survival prediction. Our results demonstrated integrating the DLS with existing risk factors resulted in an improved accuracy in survival prediction.

Previous studies have shown that higher FA and lower MD within the tumor conferred poor prognosis in both GBM and LGG [9,11]. Increased FA and decreased MD values could reflect the increased cell proliferation or cellularity [8,25]. Consistent with previous studies, our results showed higher mean FA and lower mean MD, AD and RD within the highlighted regions of the CAMs of deep learning features could be found in the high-risk subgroup than that in the low-risk subgroup. Furthermore, we demonstrated that the localizability of the deep features in our approach for low- and high-risk classification, and the most discriminative regions of the CAMs were mainly in the tumor margin and edema areas as illustrated in Fig. 6b. Thus, we deduced that the tumor margin and edema subregions with increased FA and decreased MD, AD and RD may indicated a more infiltrative tumor habitat.

Notably, the underlying biological interpretations of the imaging-based models developed by artificial intelligence should be elucidated before translation into clinical practice [26]. In this study, a radiogenomic analysis combining DTI and transcriptomic data demonstrated the CNN-learned imaging phenotypes of gliomas were significantly associated with key genes and dysregulated signaling pathways. We identified 207 DEGs across the low-risk and high-risk subgroups derived from the DTI-based DLS, and the prognostic values of some DEGs in human cancers has been revealed previously. For

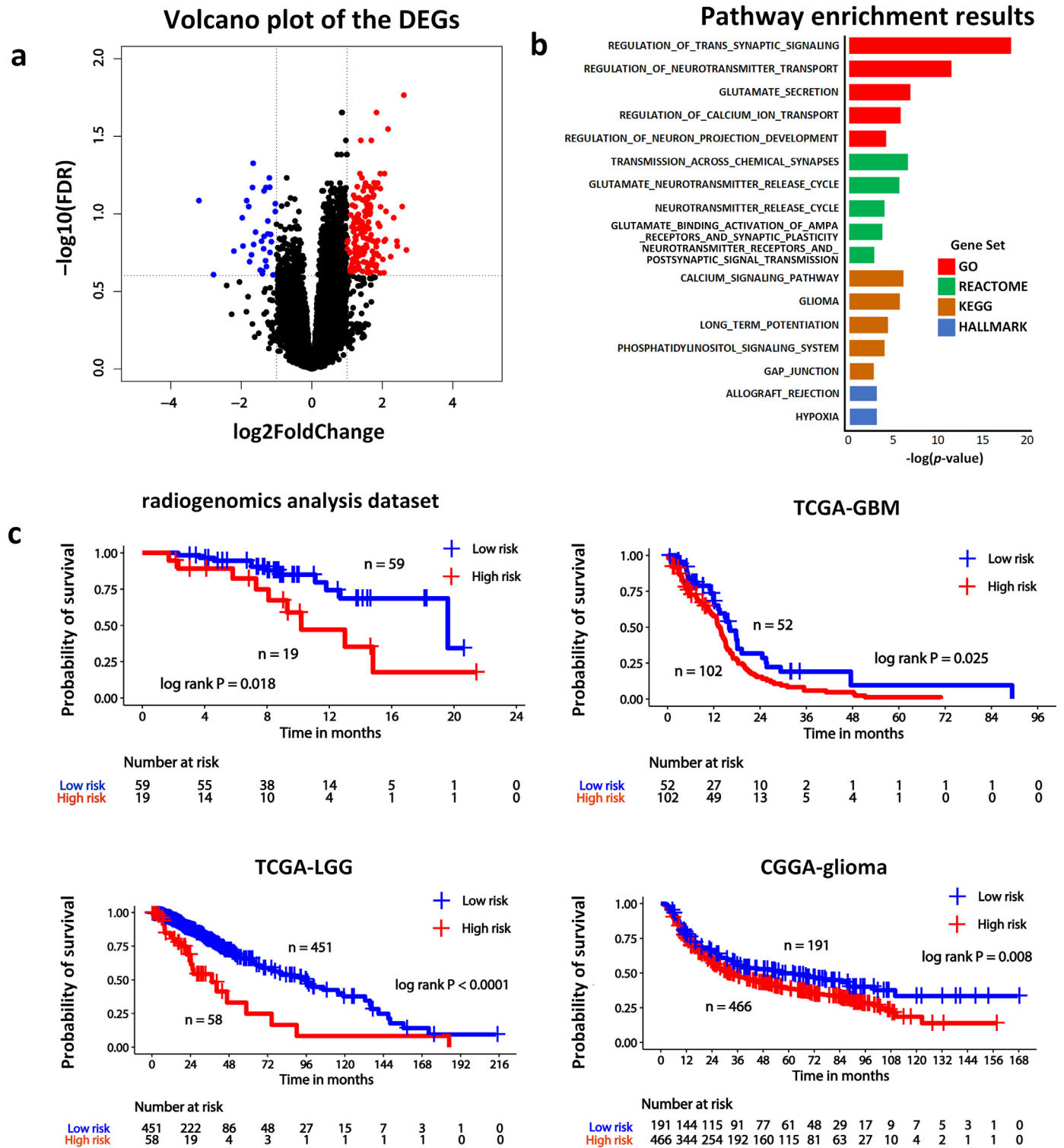


Fig. 5. A summary of the deep learning signature (DLS)-associated key genes and pathways along with the assessment of their prognostic significance. (a) Volcano plot of the differentially expressed genes (DEGs) between risk subgroups stratified by the DLS in radiogenomics analysis dataset. The red and green dots represent DEGs that were upregulated and downregulated, respectively. (b) Key enriched pathways in Gene Ontology (GO) Biological Process (red), Reactome (green), Kyoto Encyclopedia of Genes and Genomes (KEGG, brown), and Hallmark (blue) databases. (c) Kaplan-Meier curves based on the average expression value of the genes contained in the DLS-correlated pathways for overall survival prediction in the radiogenomics analysis dataset, TCGA-GBM cohort, TCGA-LGG cohort, and CGGA-glioma cohort. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

instance, high expression of SNAP25 [27] and KIF5A [28] were demonstrated to be associated with worse prognosis in colon cancer and bladder cancer. In addition, PRKCB and CAMK2A were revealed as prognostic oncogenes, and their high expressions predict poor prognosis in GBMs [29]. Similarly, our results showed the expressions of

SNAP25, KIF5A, PRKCB and CAMK2A in the low-risk subgroups were significantly lower than those in the high-risk subgroup in our radiogenomic cohort. To further investigate the prognostic values of the pathway genes associated with the DLS-based risk subgroups, the mean expression of the genes within the enriched pathways was

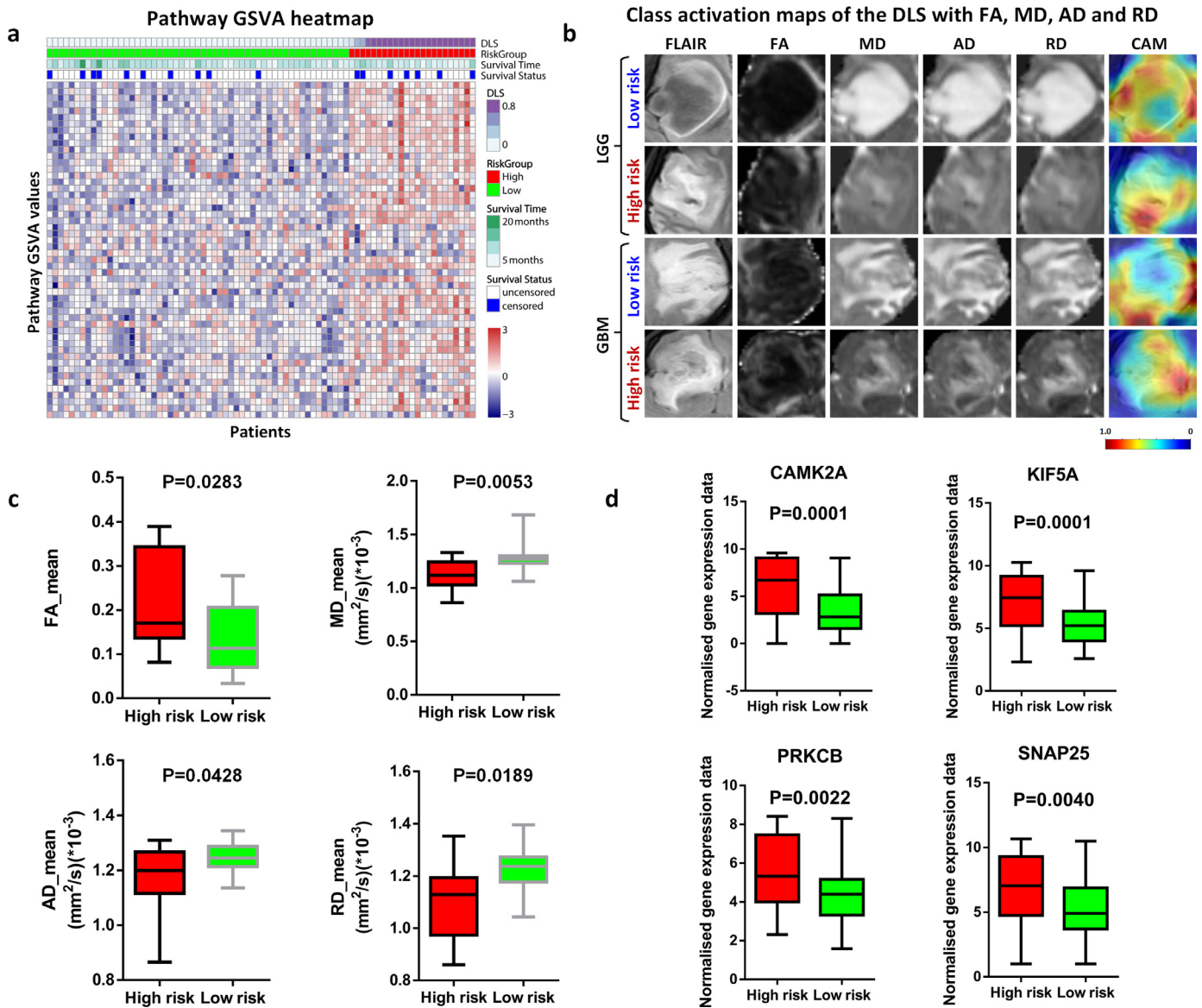


Fig. 6. A summary of the imaging-transcriptomics-prognosis associations. (a) A heatmap of the gene set variation analysis (GSVA) score of enriched pathways significantly correlated with the deep learning signature (DLS). 78 glioma patients with paired DTI and RNA-seq are shown on the x-axis, and 72 enriched pathways significantly correlated with the DLS are shown on the y-axis, which are also displayed in the Supplementary Table S5. (b) DTI maps and corresponding class activation maps (CAMs) of the DLS in two GBM patients and two LGG patients classified into low-risk subgroup (the first row, LGG, overall survival = 56.2 months, DLS = 0.002170; the third row, GBM, overall survival = 69.6 months, DLS = 0.000095) and high-risk subgroups (the second row, LGG, overall survival = 27.3 months, DLS = 0.999827; the fourth row, GBM, overall survival = 3.0 months, DLS = 1). (c) Boxplots of the mean value of FA, MD, AD and RD within the highlighted regions of CAMs in the high- and low-risk subgroups. (d) Boxplot of the expression of four representative genes CAMK2A, KIF5A, PRKCB and SNAP25 in the high- and low-risk subgroups.

found to be significantly associated with overall survival of the patients in our radiogenomics cohort, and this association was confirmed externally in TCGA and CGGA cohorts. These results demonstrated the prognostic values of DLS-associated key genes may contribute to the prognosis in glioma patients.

Our imaging-transcriptomic analysis revealed that high-risk phenotype defined by deep features from DTI is significantly associated with misadjusted GO annotation and signaling pathways related to chemical synaptic transmission/neurotransmitter transport, calcium transport/signaling, glutamate secretion/glutamate binding activation of AMPA receptors, neuron projection development/axon guidance, and glioma pathways, while these GO annotation and signaling pathways were negatively associated with low-risk deep imaging phenotype. DTI is a method that provides quantitative information about microscopic water diffusion characteristics along different orientations, which is highly anisotropic in the white matter whereas

isotropic in grey matter [30,31]. The anisotropic water diffusion is related to the ordered arrangement of the myelinated fibers in the white matter, and water molecules preferentially diffuse along the length of the neuronal axons [30,31]. Hence, DTI has been considered to be a powerful imaging tool for measuring macroscopic axonal organization in nervous system tissues [30]. Combining the biological properties of DTI and our radiogenomics findings, we propose two potential explanations for the biological mechanisms underlying in the prediction model using DTI features. The first one is the neuronal activity-related glioma progression, which is a remarkable mechanism recently found [32,33]. Venkatesh et al. [32] and Venkataramani et al. [33] suggested that glutamate-induced neuronal hyperexcitation transmits through axon and stimulates chemical synapses on glioma cells. AMPA receptors of glioma cells that are stimulated by glutamate propagates calcium signaling and further promote tumor cell growth and invasion. Thus, the synapse, calcium, glutamate and axon-related

GO annotation and signaling pathways revealed by this study indicate the deep features from DTI may reflect the glioma progression by glutamatergic neuron-to-brain tumor synaptic communication (NBTSC) [34], and this hypothesis is potentiated by the DTI's imaging capability on neuronal axons. In this perspective, NBTSC-inhibiting therapies [34] may be more effective in high-risk glioma defined by DTI-derived DLS. The second potential explanation is canonical pathways associated with gliomas such as KEGG glioma pathway, WNT signaling pathway and HIF-1 signaling pathway. These signaling pathways have been well investigated in glioma carcinogenesis and were revealed significantly related to deep features from DTI that confers prognostic significance in this study.

The present study has limitations. First, the retrospective nature of the design renders the study subject to inherent biases and confounders, although we included a relatively large sample size of cases in 2 institutions and TCIA, as well as adjusted for putative prognostic factors of gliomas. Second, deep learning features extracted by black-box-like networks are nameless and graphically obscure, which is a prominent obstacle lies in the way of translating deep learning model into clinical practice. Although we have attempted to unravel the biological basis of our presented model using radiogenomic analysis, much more should be done for explaining the biological mechanisms for deep features with prognostic significance. Third, the tumor regions of interest were drawn by only one radiologist and confirmed by a neurosurgeon, where bias might occur in the manual tumor delineation. In future we will employ automatic algorithms to achieve accurate and repeatable tumor segmentation. Fourth, as diffuse glioma is considered as not a focal but a whole brain disease, it is a reasonable hypothesis that whole-brain DTI features might better characterize the tumor invasion and thus be predictive of patient prognosis. Therefore, our future exploration also includes a whole-brain DTI model for survival prediction.

In conclusion, we proposed a deep learning model using pre-operative DTI images, which performed with robustness and generalizability to predict the clinical outcomes of glioma patients. Remarkably, we demonstrated certain deep features are associated with distinct signalling pathways that confer prognostic significance in glioma patients.

Contributors

Research conception: Zhenyu Zhang, Jing Yan, Zhicheng Li, Xianzhi Liu, Jingliang Cheng, Wencai Li, Xiaofei Lv and Yinsheng Chen; Data processing, drafting of manuscript: Zhenyu Zhang, Jing Yan, Zhicheng Li, Yuanshen Zhao, Weiwei Wang, Li Wang, Shenghai Zhang, Tianqing Ding, Lei Liu and Qiuchang Sun; Data acquisition: Dongling Pei, Wenchao Duan, Yunbo Zhan, Haibiao Zhao, Tao Sun, Chen Sun, Wenqing Wang, Zhen Liu, Xuanke Hong, Yu Guo and Xiangxiang Wang.

Data verification: Dongling Pei, Wenchao Duan, Yunbo Zhan, Haibiao Zhao

All authors have read and approved the final version of the manuscript.

Data sharing

The RNA-seq data used for radiogenomics analysis in our study have been deposited into GSA under accession number HRA000802 (<https://bigd.big.ac.cn/gsa-human/browse/HRA000802>). The remaining data and materials used to support the findings of this study are available from the corresponding authors upon request.

Declaration of Competing Interest

The authors have declared that no competing interest exists.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. U20A20171, 82102149, 61901458, 81702465, 61571432 and U1804172 and U1904148), the Science and Technology Program of Henan Province (No. 182102310113, 192102310123 and 192102310050), the Key Program of Medical Science and Technique Foundation of Henan Province (No. SBGJ202002062), the Joint Construction Program of Medical Science and Technique Foundation of Henan Province (No. LHGJ20190156), the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2018364), and Guangdong Key Project (2018B030335001).

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.ebiom.2021.103583](https://doi.org/10.1016/j.ebiom.2021.103583).

References

- Ostrom QT, Patil N, Cioffi G, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013–2017. *Neuro Oncol* 2020;22 (12 Suppl 2): iv1–iv96.
- Louis DN, Ohgaki H, Wiestler OD, et al. WHO classification of tumors of the central nervous system. Lyon: IARC Press; 2016.
- Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. *N Engl J Med* 2015;372 (26):2499–508.
- van den Bent MJ, Wefel JS, Schiff D, et al. Response assessment in neuro-oncology (a report of the RANO group): assessment of outcome in trials of diffuse low-grade gliomas. *Lancet Oncol* 2011;12(6):583–93.
- Sottoriva A, Spiteri I, Piccirillo SG, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A* 2013;110(10):4009–14.
- Itakura H, Achrol AS, Mitchell LA, et al. Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. *Sci Transl Med* 2015;7(303):303ra138.
- Saksena S, Jain R, Narang J, et al. Predicting survival in glioblastomas using diffusion tensor imaging metrics. *J Magn Reson Imaging* 2010;32(4):788–95.
- Jamjoom AA, Rodriguez D, Rajeb AT, et al. Magnetic resonance diffusion metrics indexing high focal cellularity and sharp transition at the tumour boundary predict poor outcome in glioblastoma multiforme. *Clin Radiol* 2015;70(12):1400–7.
- Heiland DH, Simon-Gabriel CP, Demerath T, et al. Integrative diffusion-weighted imaging and radiogenomic network analysis of glioblastoma multiforme. *Sci Rep* 2017;7:43523.
- Li C, Wang S, Yan JL, et al. Intratumoral heterogeneity of glioblastoma infiltration revealed by joint histogram analysis of diffusion tensor imaging. *Neurosurgery* 2019;85(4):524–34.
- Lin H, Xu Y, Chen L, et al. Multiparametric and multiregional diffusion features help predict molecule information, grade and survival in lower-grade gliomas: a feasibility study. *Br J Radiol* 2019;92(1103):20190324.
- Macyszyn L, Akbari H, Pisapia JM, et al. Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques. *Neuro Oncol* 2016;18(3):417–25.
- Kickingeder P, Neuberger U, Bonekamp D, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. *Neuro Oncol* 2018;20(6):848–57.
- Bae S, Choi YS, Ahn SS, et al. Radiomic MRI phenotyping of glioblastoma: improving survival prediction. *Radiology* 2018;289(3):797–806.
- Qian Z, Li Y, Sun Z, et al. Radiogenomics of lower-grade gliomas: a radiomic signature as a biological surrogate for survival prediction. *Aging (Albany NY)* 2018;10(10):2884–99.
- Lao J, Chen Y, Li ZC, et al. A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. *Sci Rep* 2017;7(1):10353.
- Yoon HG, Cheon W, Jeong SW, et al. Multi-parametric deep learning model for prediction of overall survival after postoperative concurrent chemoradiotherapy in glioblastoma patients. *Cancers (Basel)* 2020;12(8):2284.
- Kermayn DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018;172(5):1122–31.
- Xi IL, Zhao Y, Wang R, et al. Deep learning to distinguish benign from malignant renal lesions based on routine MR imaging. *Clin Cancer Res* 2020;26(8):1944–52.
- Beig N, Bera K, Prasanna P, et al. Radiogenomic-based survival risk stratification of tumor habitat on Gd-T1w MRI is associated with biological processes in glioblastoma. *Clin Cancer Res* 2020;26(8):1866–76.
- He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 770–8.

- [22] Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
- [23] Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* 2004;10:7252–9.
- [24] Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol* 2020;31(7):912–20.
- [25] Kinoshita M, Hashimoto N, Goto T, et al. Fractional anisotropy and tumor cell density of the tumor core show positive correlation in diffusion tensor magnetic resonance imaging of malignant brain tumors. *Neuroimage* 2008;43(1):29–35.
- [26] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin* 2019;69(2):127–57.
- [27] Zou J, Duan D, Yu C, et al. Mining the potential prognostic value of synaptosomal-associated protein 25 (SNAP25) in colon cancer based on stromal-immune score. *PeerJ* 2020;8:e10142.
- [28] Tian DW, Wu ZL, Jiang LM, et al. KIF5A promotes bladder cancer proliferation *in vitro* and *in vivo*. *Dis Markers* 2019;2019:4824902.
- [29] Zhang Y, Xu J, Zhu X. A 63 signature genes prediction system is effective for glioblastoma prognosis. *Int J Mol Med* 2018;41(4):2070–8.
- [30] Beaulieu C. The basis of anisotropic water diffusion in the nervous system—a technical review. *NMR Biomed* 2002;15(7–8):435–55.
- [31] Mori S, Zhang J. Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron* 2006;51(5):527–39.
- [32] Venkatesh HS, Morishita W, Geraghty AC, et al. Electrical and synaptic integration of glioma into neural circuits. *Nature* 2019;573(7775):539–45.
- [33] Venkataramani V, Tanev DI, Strahle C, et al. Glutamatergic synaptic input to glioma cells drives brain tumour progression. *Nature* 2019;573(7775):532–8.
- [34] Venkataramani V, Tanev DI, Kuner T, et al. Synaptic input to brain tumors: clinical implications. *Neuro Oncol* 2021;23(1):23–33.