Heliyon 10 (2024) e33760

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Exploring the prognostic value and biological pathways of transcriptomics and radiomics patterns in glioblastoma multiforme

Jixin Luan [a,b,1], Di Zhang [c,1], Bing Liu [a,b], Aocai Yang [a,b], Kuan Lv [a,d], Pianpian Hu [a,d], Hongwei Yu [a,b], Amir Shmuel [e,f], Chuanchen Zhang [c,**], Guolin Ma [a,b,*]

[a] Department of Radiology, China-Japan Friendship Hospital, Beijing, China
[b] China-Japan Friendship Hospital (Institute of Clinical Medical Sciences), Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
[c] Department of Radiology, Liaocheng People's Hospital, Shandong First Medical University & Shandong Academy of Medical Sciences, Shandong, China
[d] Peking University China-Japan Friendship School of Clinical Medicine, Beijing, China
[e] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada
[f] Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada

ARTICLE INFO

Keywords:
GBM
Transcriptomics
Radiomics
Co-expression networks
WGCNA

ABSTRACT

Objectives: To develop a multi-omics prognostic model integrating transcriptomics and radiomics for predicting overall survival in patients with glioblastoma multiforme (GBM), and investigate the biological pathways of radiomics patterns.

Materials and methods: Transcription profiles of GBM patients and normal controls were used to obtain differentially expressed mRNAs and long non-coding RNAs (lncRNAs). Radiomics features were extracted from magnetic resonance imaging (MRI). Least absolute shrinkage and selection operator (LASSO) Cox regression was employed to select survival-associated features for the construction of transcriptomics and radiomics signatures. Genes associated with GBM prognosis were identified through the analysis of lncRNA-mRNA co-expression networks and Weighted Gene Co-expression Network Analysis (WGCNA), and their biological pathways were investigated using Genomes enrichment analysis. Transcriptomics, radiomics, and clinical data were integrated to evaluate the multi-omics prognostic model's performance.

Results: LASSO Cox regression yielded 21 survival-related features, including 19 transcriptomics features and 2 radiomics features. Based on transcriptomics and radiomics signature, GBM patients were classified as high-risk or low-risk. The genes obtained from the co-expression network screen were associated with microtubule binding, while those from the WGCNA screen were associated with growth factor receptor binding. In the training set, the AUC values for the multi-omics model and clinical model were 0.964 and 0.830, respectively, while in the validation set,

* Corresponding author.Department of Radiology, China-Japan Friendship Hospital, No. 2 East Yinghua Road, Chaoyang District, Beijing 100029, China.

** Corresponding author. Department of Radiology, Liaocheng People's Hospital, Shandong First Medical University & Shandong Academy of Medical Sciences, No. 67, Dongchang West Road, Liaocheng District, Shandong Province, 252000, China.

E-mail addresses: zhangchuanchen666@163.com (C. Zhang), maguolin1007@qq.com (G. Ma).

[1] These authors share first authorship.

https://doi.org/10.1016/j.heliyon.2024.e33760
Received 15 September 2023; Received in revised form 20 June 2024; Accepted 26 June 2024
Available online 27 June 2024

they were 0.907 and 0.787. The multi-omics prognostic model outperformed the clinical prognostic model.

Conclusions: The co-expression network and WGCNA methods revealed genes associated with multiple biological pathways in GBM. The multi-omics prognostic model demonstrated excellent performance and indicated significant potential for clinical application.

Abbreviations

GBM	Glioblastoma Multiforme
AUC	Area Under the Curve
lncRNAs	Long intergenic Non-coding RNAs
FLAIR	fluid attenuated inversion recovery
FDR	False Discovery Rate
TR/TE	Repetition Time/Echo Time
ROIs	Regions Of Interest
ICC	Intraclass Correlation Coefficient
EN2	Engrailed-2
MGMT	O6-methylguanine-DNA Methyltransferase
G0S2	G0/G1 Switch Gene 2
PPP3CB	Protein Phosphatase 3 Catalytic Subunit Beta
GLDM	Grey Level Dependency Matrix
GLSZM	Grey Level Size Zone Matrix
GLCM	Grey Level Co-dependency Matrix
GLRLM	Grey Level Run Length Matrix
53BP1	P53-binding Protein 1
EGFR	Epidermal growth factor receptor
Rap:	Ras-related Protein
BIN3	Bridging Integrator-3

1. Introduction

Glioblastoma multiforme (GBM) represents the most aggressive form of primary malignant brain tumor in the adult population, exhibiting a 5-year survival rate spanning from 6 % to 22 %. This survival rate is contingent upon the individual's age and the existence of additional risk factors [1]. Prognostic models that consider only factors such as patient age, ethnicity, radiotherapy receipt, and tumor characteristics often fall short in accurately predicting overall survival [2–4]. Thus, identifying additional prognostic factors is critical to developing more effective predictive models for GBM prognosis.

Transcriptomics and radiomics data integration, namely multi-omics patterns, can provide meaningful knowledge in GBM prognosis. A transcriptome refers to all RNAs transcribed from a specific tissue or cell during a particular developmental period. It has been shown that long intergenic non-coding RNAs (lncRNAs) are closely related to GBM [5]. Researchers developed a prognostic signature for GBM patients based on six lncRNAs, with significant differences in survival analysis ($P < 0.01$) [6]. Radiomics can reflect the molecular function of tumors by extracting image features [7]. Contrast-enhanced axial T1-weighted imaging (T1WI) and fluid attenuated inversion recovery (FLAIR) imaging can guide the prognosis of GBM patients [8]. The use of contrast agents containing gadolinium in patients with renal impairment can result in nephrogenic systemic fibrosis [9]. Therefore, extracting radiomics features from FLAIR sequences is crucial for determining the prognosis of GBM. Some researchers have developed radiomics-based prognostic models that outperform traditional clinical models [10,11]. However, these studies only focus on imaging features and do not explore molecular mechanisms. Integrating multi-omics data identifies molecular interactions and enhances understanding of disease mechanisms. The construction of lncRNA-mRNA co-expression networks is an important way to analyse the function of transcriptomics molecules [12]. Weighted gene co-expression network analysis (WGCNA) explores the association between gene modules and sample phenotypes in gene networks [13]. Some scholars have used WGCNA to study the relationship between bladder cancer radiomics features and molecular features, identifying associations between radiomics risk classes and several biological pathways, such as bladder cancer angiogenesis [14].

In this study, we identified differentially expressed mRNAs and lncRNAs, obtained prognosis-related genes by co-expression network and WGCNA method, and analyzed their biological functions by Gene Oncology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) to provide a theoretical basis for the pathogenesis and therapeutic targets of GBM. The project also aims to provide individualised survival probabilities for each GBM patient to guide clinical decision-making.

2. Methods

2.1. Study population

For the study, we applied the following inclusion criteria: (1) we selected only samples that had matching The Cancer Genome Atlas (TCGA) transcriptome and the Cancer Imaging Archive (TCIA) MRI data; (2) MRI data from TCIA had to meet high-quality standards, being free of artifacts; and (3) samples needed to have complete clinical indicator information. There were 62 samples identified, including 57 GBM patients and 5 controls, which were then randomized 6:4 for training and validation. As the patient data were obtained from publicly accessible information in the TCGA database, informed consent was not required. Fig. 1 shows the flow chart for the study.

2.2. Screening for differential mRNAs and differential lncRNAs

The "limma" package was utilized to identify mRNAs and lncRNAs with differential expression. Based on the whole genome expression profile of RNA-seq data, the screening criteria were $|\text{FoldChange}| \geq 2$ and false discovery rate (FDR) ≤ 0.01.

2.3. Segmentation of images and radiomics features

ITK-SNAP software was utilized to perform 3D segmentation of FLAIR images from patient scans. Radiomics features extraction was conducted via the Pyradiomics extractor. To ensure reproducibility, two neuro-radiologists of varying experience levels (reader 1: 5

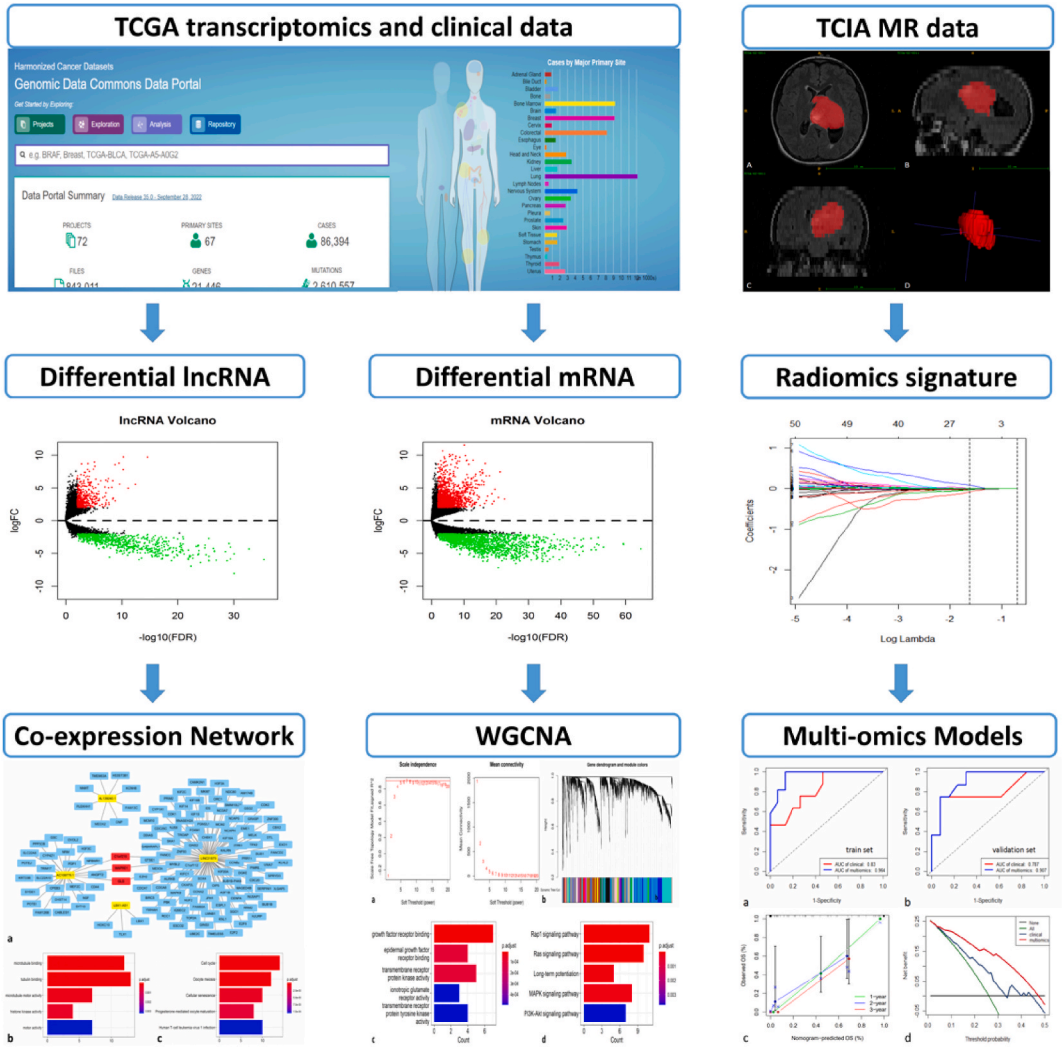

Fig. 1. Study flow chart for our analysis.

years; reader 2: 7 years) independently conducted Region Of Interests (ROI) segmentations on a subset of 30 randomly selected samples. For evaluating the consistency of measurements between the two readers, we computed the intraclass correlation coefficient (ICC). An ICC threshold of >0.75 was set to indicate satisfactory agreement. Only features with an ICC exceeding this threshold were considered reproducible.

2.4. Construction of radiomics and transcriptomics signature

Univariate Cox analysis was conducted for differentially expressed mRNAs, lncRNAs, and radiomics features. Factors with $P < 0.05$ were selected, and LASSO Cox regression was then applied to identify those significantly associated with prognosis for the construction of transcriptomics and radiomics signatures. Risk scores were computed for each patient to investigate the effectiveness of the transcriptomics or radiomics signature for clinical outcome prediction. A low-risk and high-risk group of patients was further classified based on their transcriptomics and radiomics risk scores. Survival curves for both transcriptomics and radiomics were plotted to illustrate the survival outcomes of patients in these risk categories. Additionally, a multi-omics nomogram integrating transcriptomics and radiomics will provide clinicians with a user-friendly tool for precise prediction of GBM patient survival.

2.5. WGCNA and gene function analysis

After preprocessing the data, we performed variance analysis to select the top 50 % of variant genes for subsequent WGCNA. Pearson's correlation coefficients were then calculated to construct a network that adheres to the scale-free network standard by selecting an optimal soft threshold beta. To evaluate the relationship between significant genes and radiomics signatures in identified modules, we calculated the correlation between gene significance and module membership. Gene modules with a significant positive correlation with radiomics risk level were subjected to GO and KEGG enrichment analysis.

2.6. Network of lncRNA-mRNA co-expression

The pearson correlation coefficient (PCC) was calculated between the lncRNAs screened by LASSO regression and the differential mRNAs screened in the expression profile data. lncRNA-mRNA pairs with $P < 0.05$ and $|cor| \geq 0.6$ were selected to construct a lncRNA-mRNA co-expression network [15].

Table 1
Baseline characteristics of GBM patients.

Group		Training set	Validation set	P
Age				0.187
	≤60	16 (45.71 %)	14 (63.64 %)	
	> 60	19 (54.29 %)	8 (36.36 %)	
Gender				0.339
	Female	13 (37.14 %)	11 (50.00 %)	
	Male	22 (62.86 %)	11 (50.00 %)	
Race				1.000
	Others	3 (8.57 %)	2 (9.09 %)	
	White	32 (91.43 %)	20 (90.91 %)	
KPS score				0.538
	> 60	25 (71.43 %)	14 (63.64 %)	
	≤60	10 (28.57 %)	8 (36.36 %)	
Subtype				0.618
	Classic	9 (25.71 %)	7 (31.82 %)	
	Non-classic	26 (74.29 %)	15 (68.18 %)	
CIMP				1.000
	Positive	3 (8.57 %)	1 (4.55 %)	
	Negative	32 (91.43 %)	21 (95.45 %)	
IDH				1.000
	Mutant	4 (11.43 %)	2 (9.09 %)	
	Wild type	31 (88.57 %)	20 (90.91 %)	
Radiotherapy				1.000
	Yes	4 (11.43 %)	2 (9.09 %)	
	No	31 (88.57 %)	20 (90.91 %)	
Pharmaceutical				0.945
	Yes	5 (14.29 %)	3 (13.64 %)	
	No	30 (85.71 %)	19 (86.36 %)	
Status				1.000
	Survival	4 (11.43 %)	1 (4.55 %)	
	Death	31 (88.57 %)	21 (95.45 %)	
Survival time	Time (years)	1.21 ± 1.11	1.67 ± 1.68	0.222

2.7. Prognostic models development and evaluation

Two prognostic models were developed: a clinical model, considering age, gender, race and other factors, and a multi-omics model, incorporating clinical factors, radiomics, and transcriptomics. The performance of the prognostic models was evaluated for their ability to discriminate, calibrate, and demonstrate clinical effectiveness [16]. Discrimination was assessed by using the area under the receiver-operating characteristic curve (AUC) and the concordance index (C-index). Furthermore, the reclassification performance was analyzed through IDI (integrated discrimination improvement) [17] and NRI (net reclassification improvement) [18]. The calibration curves were used to assess calibration, and the decision curve analysis (DCA) was used to evaluate clinical effectiveness, which calculated net benefits across various risk threshold probabilities [19].

2.8. Statistical analysis

Statistical analyses were conducted using R (Version 3.6.0) (http://www.R-project.org,2019). We used the following R packages: the glmnet package for LASSO regression, the survival package for survival analysis, the timeROC package to obtain ROC results, and DCA using the stdca package. The clusterProfiler package facilitated GO and KEGG analysis, while the WGCNA package was utilized for WGCNA analysis. Kaplan-Meier analysis was used to generate survival curves. To compare patient characteristics between the training and validation sets, continuous variables were analyzed with t-tests or Mann-Whitney tests, and P-values were reported. Categorical variables, such as subtypes, were assessed using chi-square tests. P-values <0.05 were considered significant.

3. Results

3.1. Patients clinical characteristics

A summary of the clinical characteristics of the training and validation sets can be found in Table 1. The median survival time was 1.21 years in the training set and 1.67 years in the validation set. No statistically significant differences were found in patient age, gender, race, KPS score, subtype, CpG island methylation phenotype (CIMP), isocitrate dehydrogenase (IDH), radiation, pharmaceutical, or overall survival (P = 0.187–1.000).

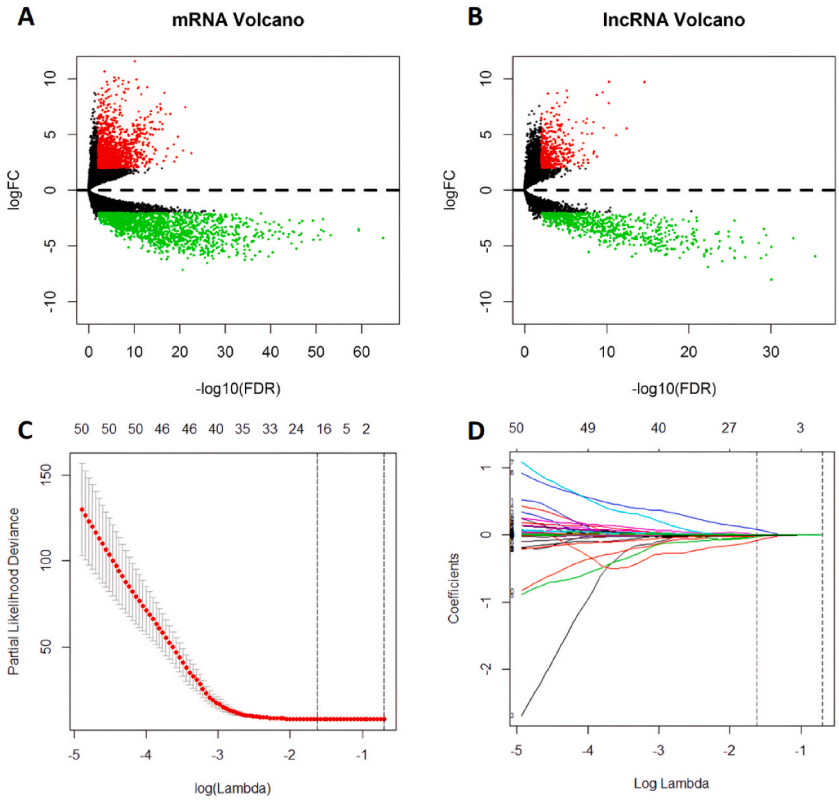

Fig. 2. Screening the mRNAs, lncRNAs and radiomics features related to prognosis and construction of transcriptomics and radiomics signature. (A) Volcano plot for differential mRNAs expression. (B) Volcano plot for differential lncRNAs expression. (C) Lambda selection by 10-fold cross-validation. (D) Processes of LASSO Cox model fitting. Each curve represents a feature.

3.2. Prognosis-related transcriptomics and radiomics factors

A total of 3129 differential mRNAs (up-regulated 1516, down-regulated 1613) and 1132 differential lncRNAs (up-regulated 411, down-regulated 721) were found in GBM patients compared to normal subjects (Fig. 2A and B). Univariate analysis identified 102 differential mRNAs and 38 differential lncRNAs as prognostic factors.

A total of 851 radiomic features were extracted, comprising 107 original features and 744 wavelet features. Univariate analysis identified six radiomic features as prognostic factors.

3.3. Construction of transcriptomics and radiomics signature

In this study, we incorporated 102 mRNAs showing differential expression, 38 differential lncRNAs, and 6 radiomic features with $P < 0.05$ from the univariate analysis into the LASSO Cox regression. The optimal prognosis-related features were screened using 10-fold cross-validation, and the best features were screened by adjusting the lambda values of different parameters to obtain the smallest bias. The screening process is shown in Fig. 2C and D.

LASSO Cox regression screened 21 features, including 2 radiomics features and 19 transcriptomics features (14 mRNAs and 5 lncRNAs), which were used as transcriptomics and radiomics signature (Table 2). The median transcriptomics and radiomics risk scores were 1.405 and 1.185.

3.4. Transcriptionomics and radiomics signature prognostic value

Univariate Cox regression analysis revealed that several factors were significantly associated with overall survival in GBM: radiation therapy ($P = 0.01$), CIMP status ($P = 0.03$), IDH mutation status ($P < 0.05$), pharmaceutical treatment ($P = 0.02$), radiomics risk level ($P = 0.02$), and transcriptomics risk level ($P < 0.05$) (Table 3). In the multivariate analysis, pharmaceutical treatment ($P < 0.05$), radiomics risk level ($P < 0.05$), and transcriptomics risk level ($P < 0.05$) appeared to be independent GBM prognostic factors (Table 3). Kaplan-Meier curves were used to compare survival rates between high-risk and low-risk groups (Fig. 3A and B). Transcriptomics signatures, radiomic signatures, and clinical factors were combined to create a multi-omics nomogram (Fig. 3C).

3.5. LncRNA-mRNA expression profiles and functional networks

The constructed lncRNA-mRNA co-expression network revealed that four lncRNAs (AC109779.1, AL139240.1, LBX1-AS1, LINC01879) were highly co-expressed with 141 mRNAs (Fig. 4A). GO analysis revealed that the mRNAs in the co-expression network were mainly involved in microtubule binding (GO:0008017) and tubulin binding (GO:0015631) (Fig. 4B), and KEGG analysis revealed that the above mRNAs were mainly involved in Cell cycle (Fig. 4C).

3.6. Molecular functional analysis in WGCNA and radiomics risk classification

With a soft threshold of $\beta = 5$, we ensured the network was scale-free (Fig. 5A), and the results generated 22 distinct gene co-expression modules in the GBM samples (Fig. 5B). A significant and positive correlation was found between the genes in the orange module and radiomics risk levels ($cor = 0.264$, $P < 0.05$). The genes in the orange module contained the genes TEK, TBX4, W5N891, CALD1 and F6TQP7. GO analysis of the genes in the orange module revealed that these genes were mainly involved in growth factor receptor binding (GO:0070851) and epithelial growth factor receptor binding (GO:0005154) (Fig. 5C), and KEGG analysis revealed that the above genes were mainly involved in various pathways such as Rap1 and Ras signaling pathway (Fig. 5D).

Table 2
Radiomics and transcriptomics signatures.

Radiomics	Coefficient	lncRNAs	Coefficient	mRNAs	Coefficient
log-sigma-3-0-mm-3D gldm LargeDependence LowGrayLevelEmphasis	−2.67e-02	AL132800.1	−1.04e-01	AC01153	3.59e-01
				BICDL1	8.85e-04
				CABP4	2.38e-03
		AL139240.1	−3.01e-01	CLEC18C	2.45e-02
				EN2	4.55e-03
				FAM9C	−6.10e-01
		AC109779.1	−2.96e-01	G0S2	1.40e-03
log-sigma-3-0-mm-3D glszm LargeArea HighGrayLevelEmphasis	−2.22e-09			GUCA1A	1.21e-02
				MYO15A	3.98e-02
		LBX1-AS1	3.35e-02	PGBD5	9.55e-04
				PPP3CB	6.39e-06
				RPE65	−2.04e-04
		LINC01879	1.09e-01	SYN3	−1.69e-03
				TMEM100	−6.92e-05

Table 3
Cox regression univariate and multivariate analyses.

Variables	Univariate HR (95 % CI)	P	Multivariate HR (95 % CI)	P
Age	1.59 (0.92–2.77)	0.10		
Gender	0.98 (0.56–1.72)	0.94		
Race	1.02 (0.37–2.84)	0.97		
KPS	1.24 (0.68–2.26)	0.48		
Subtype	0.86 (0.47–1.57)	0.62		
CIMP	9.44 (1.29–69.26)	0.03	1.99 (0.18–22.18)	0.57
IDH	10.40 (2.43–44.53)	0.00	3.77 (0.63–22.51)	0.14
Radiation	3.10 (1.30–7.41)	0.01	1.05 (0.20–5.49)	0.89
Pharmaceutical	2.49 (1.15–5.39)	0.02	6.27 (1.41–27.97)	<0.05
Radiomics risk level	1.98 (1.10–3.54)	0.02	2.01 (1.04–3.86)	<0.05
Transcriptomics risk level	19.34 (6.71–55.74)	<0.05	22.76 (7.64–67.79)	<0.05

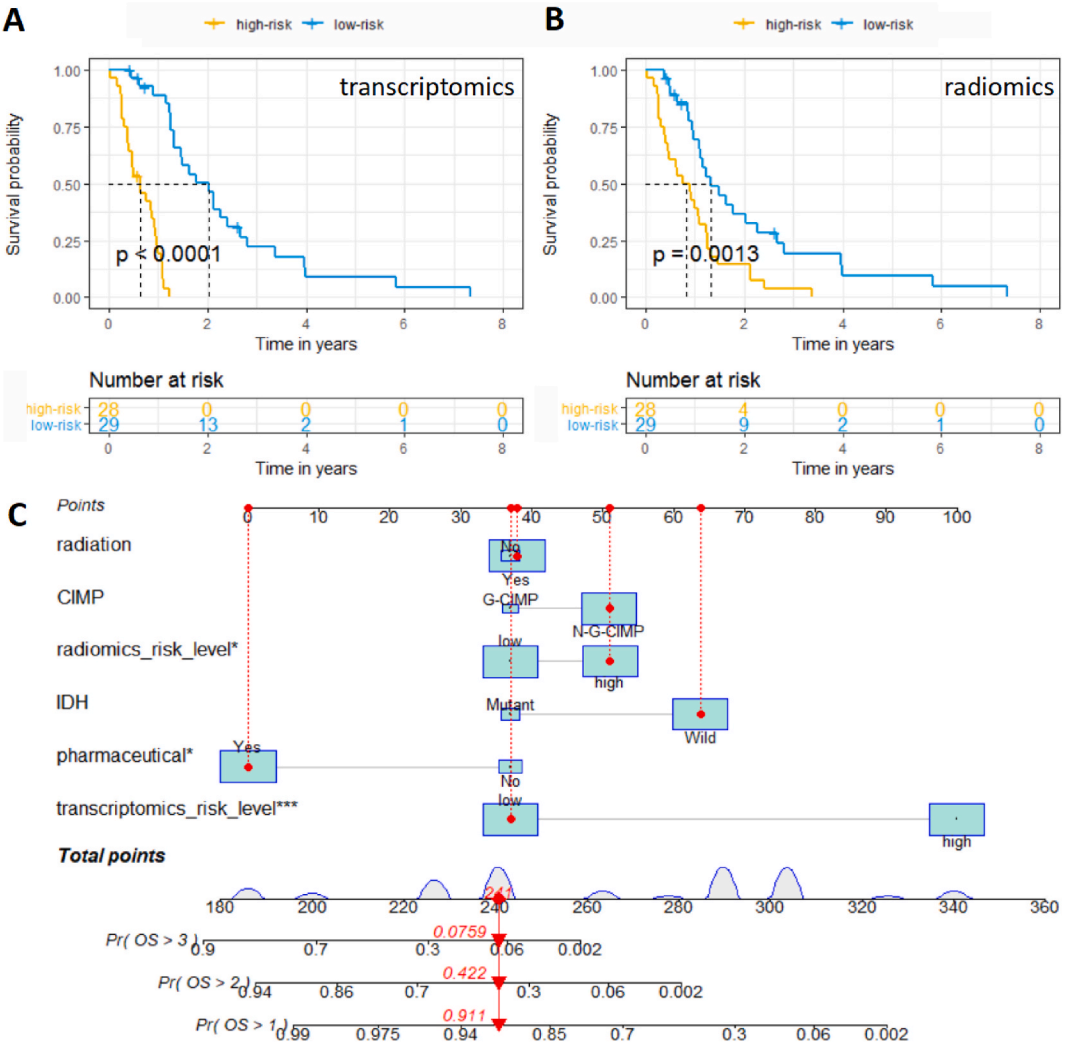

Fig. 3. Survival curves and the multi-omics nomogram. (A) Kaplan-Meier survival curves with different transcriptomics risk scores. (B) Kaplan-Meier survival curves with different radiomics risk scores. (C) The multi-omics nomogram of GBM patients.

3.7. The performances of the different prognostic models

In the training set, the AUC values for the multi-omics model and clinical model were 0.964 and 0.830, respectively, while in the validation set, they were 0.907 and 0.787. In the training set, the C-index for the multi-omics model and clinical model were 0.869 and

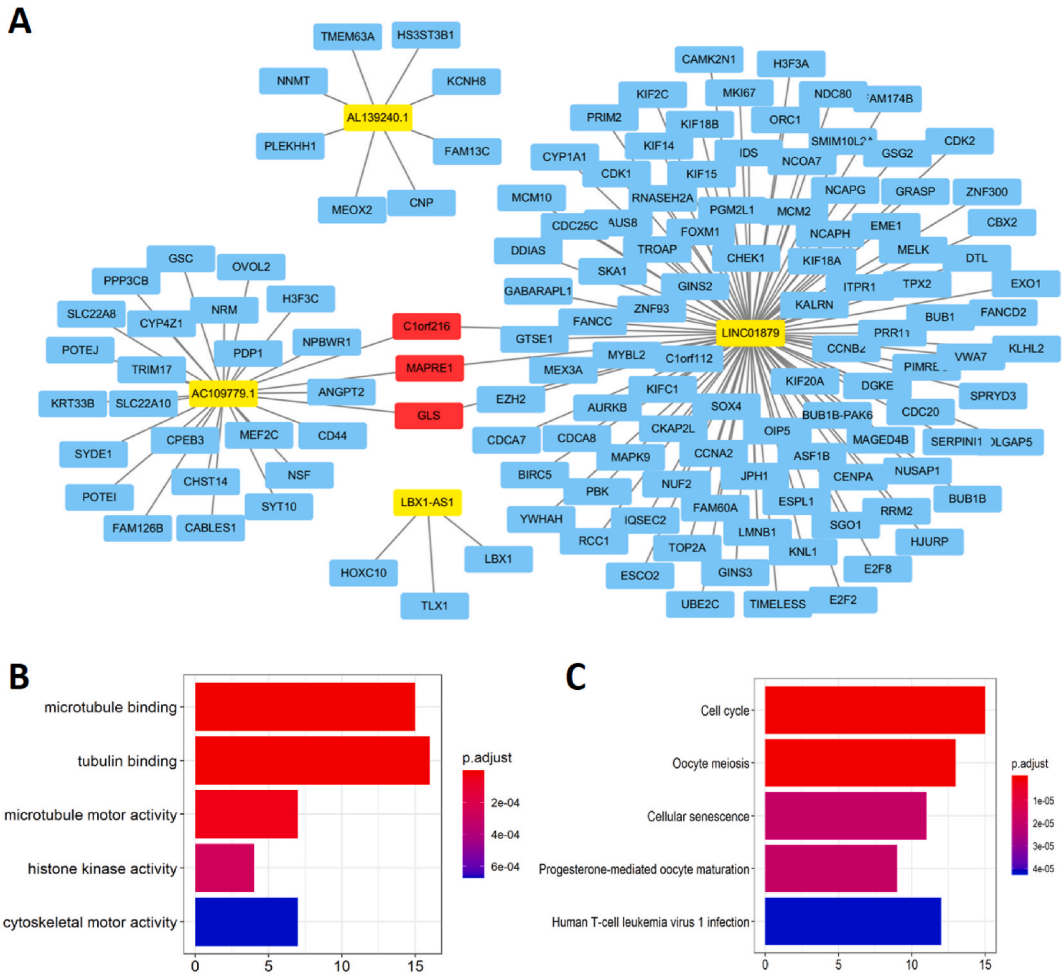

Fig. 4. LncRNA-mRNA co-expression network construction, GO and KEGG pathway enrichment analysis. (A) Yellow rectangular boxes represent lncRNAs, blue rectangular boxes represent mRNAs, red rectangular boxes represent mRNAs co-expressed to AC109779.1 and LINC01879, mRNAs or lncRNAs linked at both ends of the line have co-expression relationships. (B) The barplot of GO analysis. (C) The barplot of KEGG enrichment analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

0.784, respectively, while in the validation set, they were 0.891 and 0.751 (Fig. 6A and B). The NRI and IDI of the multi-omics model were 0.302 and 0.119 in the training set, indicating substantial improvement in reclassification. The calibration plots showed the good performance of the multi-omics model (Fig. 6C), and the multi-omics model gained more net benefits than the clinical model (Fig. 6D).

4. Discussion

This study utilized data from TCGA and TCIA to construct a transcriptomics prognostic signature comprising 19 features and a radiomics prognostic signature comprising 2 features. Additionally, co-expression networks and the WGCNA approach were applied to identify biological pathways linked to GBM. More importantly, a multi-omics prognostic model was established, and perfect model validation was performed by integrating clinical factors, enabling precise prediction of overall survival in GBM patients.

Transcriptomics of mRNA and lncRNA is closely associated with the prognosis of GBM patients [20,21]. EN2 (Engrailed-2) enhances the sensitivity of glioma to temozolomide by reducing cell proliferation and increasing apoptosis, similar to MGMT [22]. PPP3CB is a calmodulin-regulated protein phosphatase, and Lou et al. [23] used bioinformatics analysis to demonstrate that PPP3CB plays a tumor suppressive role in GBM. The lncRNAs were expressed in a variety of neural tissues or cells, including brain and retinal tissues, and were associated with the development of neurological tumors [24]. Zhang et al. [6] created a signature composed of six lncRNAs, which helped to identify high-risk patients with poor prognosis. However, few studies have reported the use of mRNAs and lncRNAs together as signature for GBM. The present study included mRNAs and lncRNAs together as signature, which can provide a more accurate prognostic assessment for GBM patients.

Radiomics is a valuable method for comprehensive assessment that reflects tumor phenotypes, genotypes and other biological features [7,25]. FLAIR sequences can reveal a high signal in some patients with progression [26]. The study's findings revealed that

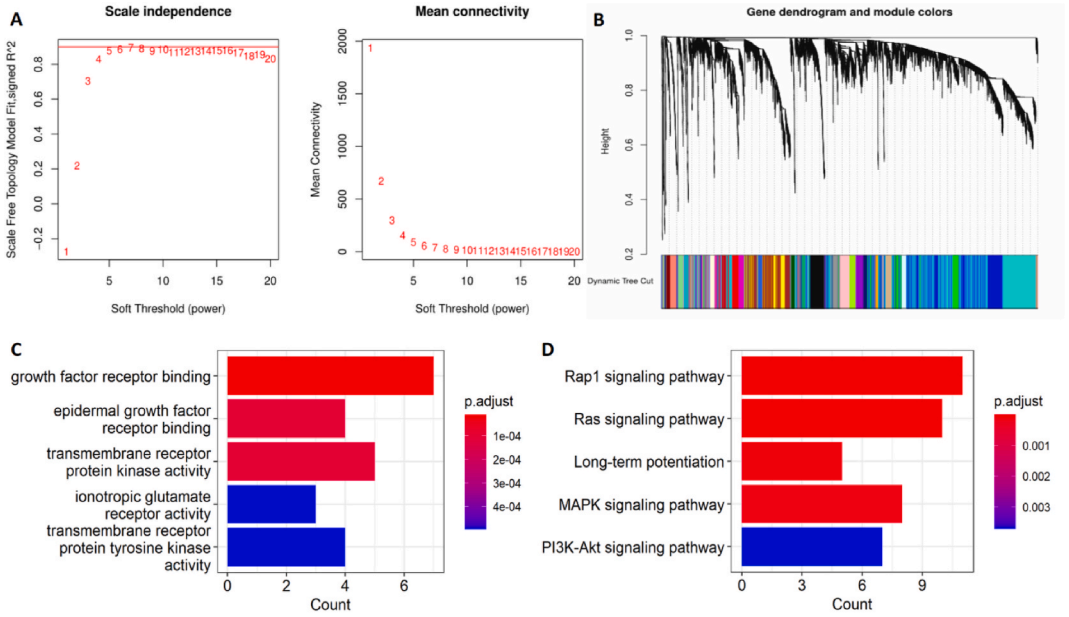

Fig. 5. WGCNA network construction, GO and KEGG pathway enrichment analysis. (A) Soft threshold determination of the WGCNA. (B) Clustering dendrograms of highly connected genes. (C) The barplot of GO analysis. (D) The barplot of KEGG enrichment analysis.

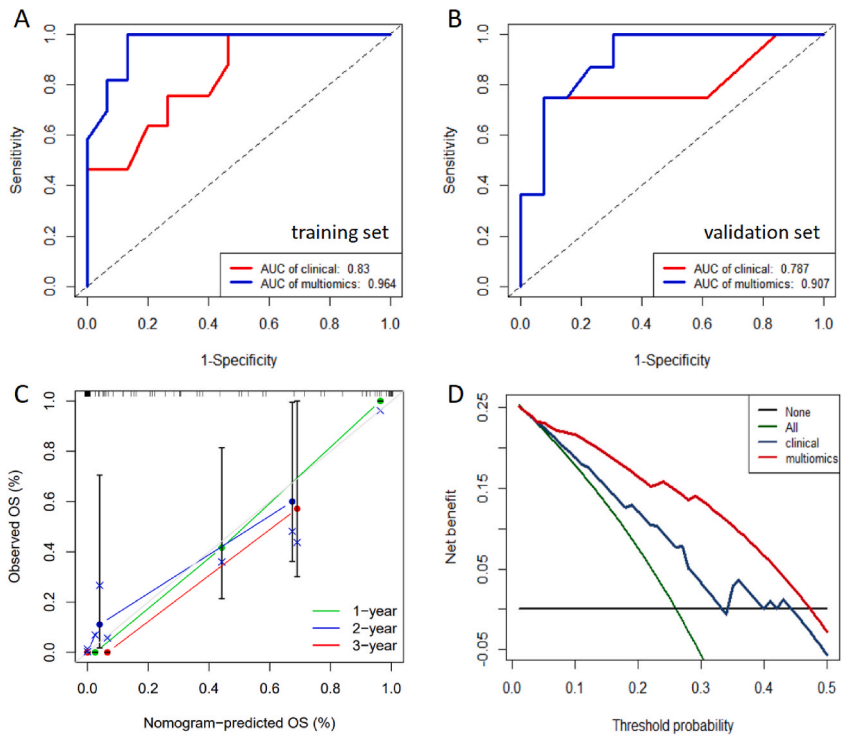

Fig. 6. The ROC curves, calibration curves, and DCA curves of the models. (A). The ROC curves for the training sets. (B). The ROC curves for the validation sets. (C) The calibration curves of the multi-omics prognostic nomogram. (D) DCA decision curves, red represents the multi-omics prognostic model and dark blue represents the clinical prognostic model. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

two radiomics features, extracted from FLAIR sequences, were strongly linked to survival in patients with GBM. These features served as indicators of the tumor's grayscale heterogeneity, providing valuable insights into its biological complexity and prognostic implications. Consistent with the findings of this study, Liu et al. [27] found that features in the Grey Level Co-dependency Matrix (GLCM) reflected regional heterogeneity and helped to determine differences between long-term and short-term groups of GBM patient survival. Similarly, Chaddad et al. [28] study of 39 GBM patients with enhanced T1WI and FLAIR sequences found that the GLCM parameters extracted from the tumor-enhanced and edematous regions correlated well with patient survival ($p < 0.01$).

In this study, GO analysis of genes within the lncRNA-mRNA co-expression network showed that they are mainly related to the microtubule binding, among others. Microtubule binding is crucial in GBM development. A novel microtubule drug has been developed which targets microtubule binding, penetrates the blood-brain barrier, selectively kills GBM cells, and has a good safety profile. KEGG analysis revealed that the above genes are associated with the cell cycle and other signalling pathways. The transcriptomics prognostic signature identified G0S2, a cell cycle gene that enhances the effectiveness of glioma radiation therapy by regulating the stability of the 53BP1 protein [29]. GO analysis of the genes identified through WGCNA indicated a strong association with growth factor receptor binding. EGFR amplification is an important oncogenic driver in GBM [30]. It has been shown that ligand-induced EGFR activation leads to proliferation and reduced invasive capacity of GBM cells through upregulation of Bridging Integrator-3 (BIN3) [31]. KEGG analysis revealed that these genes are mainly associated with signalling pathways, such as the Rap1 signaling pathway. The Ras superfamily of small G proteins encompasses the Ras-related protein (Rap) subfamily, which includes Rap1 and Rap2 [32]. The Rap1 expression is elevated in GBM patients and is associated with elevated tumor grade [33].

The multi-omics prognostic model developed in this study outperformed the clinical prognostic model. Most studies have constructed prognostic models using only clinical factors [34]. Zhou et al. [35] developed a prognostic model based on lncRNAs that predicts clinical outcomes in patients with GBM. However, the authors did not include mRNAs in their study. Chaddad et al. [36] combined a composite model of clinical factors, imaging features and genomics features, achieving the greatest AUC ($P < 0.001$). The present study innovatively incorporated radiomics and transcriptomics features together to construct a multi-omics prognostic model with superior predictive performance, with an AUC of 0.964 for predicting survival in the training set. Table 4 summarizes relevant studies conducted in recent years. This table enables a clear evaluation of the similarities and differences between our research and previous studies [34–42]. Through the data presented in the table, our model achieved high accuracy. In addition, we provided crucial biological interpretations of the radiomics features, specifically in explaining the pathological and physiological processes of GBM.

There are limitations to this study. This study retrospectively collected MRI images in the TCIA database, which has high variability in scan parameters. In addition, this study had to ensure that both transcriptomics and imaging data were available for the study population. Therefore, only 57 patients were included, which is a small sample size, as well as incomplete clinical risk factors for some patients. In this study, the prognostic model was constructed using features solely extracted from FLAIR sequences, without incorporating other functional MRI techniques. We will collect more cases, ensure scanning consistency to enhance the accuracy of prognostic predictions for GBM patients.

5. Conclusions

This study combined radiomics and transcriptomics data and used the co-expression network and WGCNA method to screen genes associated with GBM prognosis, which could provide a theoretical basis for the pathogenesis of GBM. Furthermore, a prognostic model based on multi-omics and clinical factors was developed and validated, demonstrating excellent performance and indicating significant potential for clinical application.

Ethical approval

The study utilizes publicly available data from the public domain, and the dataset used is freely accessible, therefore, no ethics statement is required for this work.

Data availability statement

The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author.

Funding

This study was supported by the Guangzhou Science and Technology Planning Project (No. 202103010001), the National Natural Science Foundation of China (No. 82271953, 81971585 and 61976110), and STI2030-Major Projects (No. 2022ZD0213300) and Capital's Funds for Health Improvement and Research (No. 2022-1-2031) and Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (No. CNLZD2303). Beijing Municipal Science and Technology Project (No. Z211100003521009).

CRediT authorship contribution statement

Jixin Luan: Writing – original draft, Software, Investigation, Formal analysis, Data curation. **Di Zhang:** Supervision, Software,

Table 4
Application of prognostic models in GBM survival prediction.

Authors	Year	Data Source	Predictors	Algorithm	Performance
Gittleman et al. [34]	2017	Internal cohort	Clinical + MGMT	Cox regression	With a C-index of 0.657
Zhou et al. [35]	2018	TCGA	Clinical + transcriptomics features	Cox regression	Training: AUC 0.73, validation: AUC 0.67
Chaddad et al. [36]	2019	TCIA and Internal cohort	Radiomics + clinical + genomics + protein expression	Random forest	With a AUC of 0.782
Zhang et al. [37]	2019	Internal cohort	Radiomics + clinical	Logistic regression	Training: C-index, 0.971, validation: C-index 0.974
Verma et al. [38]	2020	Internal cohort	Assessing features that are prognostic for progression-free survival	LASSO, Cox regression	With the highest C-index of 0.80
Choi et al. [39]	2021	Internal cohort	Radiomics + clinical + MGMT and IDH status	Deep learning	Combined overall and progression-free survival AUC 0.73 and 0.67
Ammari et al. [40]	2021	Brain Tumor Segmentation (BraTS)	Radiomics + clinical	Seven Machine learning (ML) algorithms	With the highest AUC of 0.71
Kazerooni et al. [41]	2022	Internal cohort	Radiomics + genomics + MGMT + methylation + clinical	Deep learning	With a C-index of 0.75
Hajianfar et al. [42]	2023	TCIA and Internal cohort	Radiomics + clinical	Six time-to-event ML algorithms	With the highest C-index of 0.77

Formal analysis. **Bing Liu:** Software, Resources. **Aocai Yang:** Validation, Resources. **Kuan Lv:** Data curation. **Pianpian Hu:** Data curation. **Hongwei Yu:** Software, Data curation. **Amir Shmuel:** Writing – review & editing, Validation. **Chuanchen Zhang:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Guolin Ma:** Methodology, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Zeshan Yao for English editting.

References

[1] Q.T. Ostrom, N. Patil, G. Cioffi, K. Waite, C. Kruchko, J.S. Barnholtz-Sloan, CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013-2017, Neuro Oncol. 22 (2020) 1–96. https://doi.org/10.1093/neuonc/noaa200.
[2] A.C. Tan, D.M. Ashley, G.Y. López, M. Malinzak, H.S. Friedman, M. Khasraw, Management of glioblastoma: State of the art and future directions,, CA Cancer J Clin. 70 (2020) 299–312. https://doi.org/10.3322/caac.21613.
[3] L. Khan, H. Soliman, A. Sahgal, J. Perry, W. Xu, M.N. Tsao, External beam radiation dose escalation for high grade glioma, Cochrane Database Syst Rev. (2016) Cd011475. https://doi.org/10.1002/14651858.CD011475.
[4] E. Costa, T.M. Lawson, J. Lelotte, E. Fomekong, R.G. Vaz, L. Renard, et al., Long-term survival after glioblastoma resection: hope despite poor prognosis factors, J Neurosurg Sci. 63 (2019) 251–257. https://doi.org/10.23736/s0390-5616.18.04180-2.
[5] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, Nat Rev Genet. 10 (2009) 57–63. https://doi.org/10.1038/nrg2484.
[6] X.Q. Zhang, S. Sun, K.F. Lam, K.M. Kiang, J.K. Pu, A.S. Ho, et al., A long non-coding RNA signature in glioblastoma multiforme predicts survival, Neurobiol Dis. 58 (2013) 123–131.. https://doi.org/10.1016/j.nbd.2013.05.011.
[7] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G. van Stiphout, P. Granton, et al., Radiomics: extracting more information from medical images using advanced feature analysis, Eur J Cancer. 48 (2012) 441–446. https://doi.org/10.1016/j.ejca.2011.11.036.
[8] R.E. Yoo, S.H. Choi, T.M. Kim, S.H. Lee, C.K. Park, S.H. Park, et al., Independent poor prognostic factors for true progression after radiation therapy and concomitant temozolomide in patients with glioblastoma: subependymal enhancement and low ADC value, AJNR Am J Neuroradiol. 36 (2015) 1846–1852. https://doi.org/10.3174/ajnr.A4401.
[9] T. Grobner, Gadolinium–a specific trigger for the development of nephrogenic fibrosing dermopathy and nephrogenic systemic fibrosis? Nephrol Dial Transplant. 21 (2006) 1104–1108. https://doi.org/10.1093/ndt/gfk062.
[10] Y. Tan, S.T. Zhang, J.W. Wei, D. Dong, X.C. Wang, G.Q. Yang, et al., A radiomics nomogram may improve the prediction of IDH genotype for astrocytoma before surgery, Eur Radiol. 29 (2019) 3325–3337. https://doi.org/10.1007/s00330-019-06056-4.
[11] X Jia, Y Zhai, D Song, Y Wang, S Wei, F Yang, et al., A Multiparametric MRI-Based Radiomics Nomogram for Preoperative Prediction of Survival Stratification in Glioblastoma Patients With Standard Treatment, Front Oncol. 12 (2022) 758622. https://doi.org/10.3389/fonc.2022.758622.
[12] S. Liu, R. Mitra, M.M. Zhao, W. Fan, C.M. Eischen, F. Yin, et al., The potential roles of long noncoding RNAs (lncRNA) in glioblastoma development, Mol Cancer Ther. 15 (2016) 2977–2986. https://doi.org/10.1158/1535-7163.Mct-16-0320.
[13] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinformatics 9 (2008) 559. https://doi.org/10.1186/1471-2105-9-559.
[14] P. Lin, D.Y. Wen, L. Chen, X. Li, S.H. Li, H.B. Yan, et al., A radiogenomics signature for predicting the clinical outcome of bladder urothelial carcinoma, Eur Radiol. 30 (2020) 547–557. https://doi.org/10.1007/s00330-019-06371-w.
[15] Q. Li, H. Jia, H. Li, C. Dong, Y. Wang, Z. Zou, LncRNA and mRNA expression profiles of glioblastoma multiforme (GBM) reveal the potential roles of lncRNAs in GBM pathogenesis, Tumour Biol. 37 (2016) 14537–14552. https://doi.org/10.1007/s13277-016-5299-0.
[16] A.C. Alba, T. Agoritsas, M. Walsh, S. Hanna, A. Iorio, P.J. Devereaux, et al., Discrimination and calibration of clinical prediction models: users' guides to the medical literature, JAMA. 318 (2017) 1377–1384. https://doi.org/10.1001/jama.2017.12126.
[17] K.F. Kerr, R.L. McClelland, E.R. Brown, T. Lumley, Evaluating the incremental value of new biomarkers with integrated discrimination improvement, Am J Epidemiol. 174 (2011) 364–374. https://doi.org/10.1093/aje/kwr086.

[18] M.J. Pencina, D'Agostino RbSr, R.B. D'Agostino Jr., R.S. Vasan, Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond, Stat Med. 27 (2008) 157–172. https://doi.org/10.1002/sim.2929.
[19] A.J. Vickers, A.M. Cronin, E.B. Elkin, M. Gonen, Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers, BMC Med Inform Decis Mak. 8 (2008) 53. https://doi.org/10.1186/1472-6947-8-53.
[20] A. Arimappamagan, K. Somasundaram, K. Thennarasu, S. Peddagangannagari, H. Srinivasan, B.C. Shailaja, et al., A fourteen gene GBM prognostic signature identifies association of immune response pathway and mesenchymal subtype with high risk group, PLoS One 8 (2013) e62042. https://doi.org/10.1371/journal.pone.0062042.
[21] A.P. Patel, I. Tirosh, J.J. Trombetta, A.K. Shalek, S.M. Gillespie, H. Wakimoto, et al., Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma, Science. 344 (2014) 1396–1401. https://doi.org/10.1126/science.1254257.
[22] T. Li, W. Yang, M. Li, S. Zhang, X. Zhou, M. Zuo, et al., Engrailed 2 (EN2) acts as a glioma suppressor by inhibiting tumor proliferation/invasion and enhancing sensitivity to temozolomide, Cancer Cell Int. 20 (2020) 65. https://doi.org/10.1186/s12935-020-1145-y.
[23] W. Lou, B. Ding, L. Xu, W. Fan, Construction of potential glioblastoma multiforme-related miRNA-mRNA regulatory network, Front Mol Neurosci. 12 (2019) 66. https://doi.org/10.3389/fnmol.2019.00066.
[24] T.R. Mercer, I.A. Qureshi, S. Gokhan, M.E. Dinger, G. Li, J.S. Mattick, et al., Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation, BMC Neurosci. 11 (2010) 14. https://doi.org/10.1186/1471-2202-11-14.
[25] J.M. Lemée, A. Clavreul, P. Menei, Intratumoral heterogeneity in glioblastoma: don't forget the peritumoral brain zone, Neuro Oncol. 17 (2015) 1322–1332. https://doi.org/10.1093/neuonc/nov119.
[26] R. Grossman, N. Shimony, D. Shir, T. Gonen, R. Sitt, T.J. Kimchi, et al., Dynamics of FLAIR volume changes in glioblastoma and prediction of survival, Ann Surg Oncol. 24 (2017) 794–800. https://doi.org/10.1245/s10434-016-5635-z.
[27] Y. Liu, X. Xu, L. Yin, X. Zhang, L. Li, H. Lu, Relationship between glioblastoma heterogeneity and survival time: an MR imaging texture analysis, AJNR Am J Neuroradiol. 38 (2017) 1695–1701. https://doi.org/10.3174/ajnr.A5279.
[28] A. Chaddad, C. Desrosiers, M. Toews, Radiomic analysis of multi-contrast brain MRI for the prediction of survival in patients with glioblastoma multiforme, Annu Int Conf IEEE Eng Med Biol Soc. 2016 (2016) 4035–4038. https://doi.org/10.1109/embc.2016.7591612.
[29] Y. Wang, Y. Hou, W. Zhang, A.A. Alvarez, Y. Bai, B. Hu, et al., Lipolytic inhibitor G0S2 modulates glioma stem-like cell radiation response, J Exp Clin Cancer Res. 38 (2019) 147. https://doi.org/10.1186/s13046-019-1151-x.
[30] E. Eskilsson, G.V. Røsland, G. Solecki, Q. Wang, P.N. Harter, G. Graziani, et al., EGFR heterogeneity and implications for therapeutic intervention in glioblastoma, Neuro Oncol. 20 (2018) 743–752. https://doi.org/10.1093/neuonc/nox191.
[31] G. Guo, K. Gong, N. Beckley, Y. Zhang, X. Yang, R. Chkheidze, et al., EGFR ligand shifts the role of EGFR from oncogene to tumour suppressor in EGFR-amplified glioblastoma by suppressing invasion through BIN3 upregulation, Nat Cell Biol. 24 (2022) 1291–1305. https://doi.org/10.1038/s41556-022-00962-4.
[32] X.X. Guo, S. An, Y. Yang, Y. Liu, Q. Hao, T.R. Xu, Rap-interacting proteins are Key players in the Rap symphony orchestra, Cell Physiol Biochem. 39 (2016) 137–156. https://doi.org/10.1159/000445612.
[33] J. Volovetz, A.D. Berezovsky, T. Alban, Y. Chen, A. Lauko, G.F. Aranjuez, et al., Identifying conserved molecular targets required for cell migration of glioblastoma cancer stem cells, Cell Death Dis. 11 (2020) 152. https://doi.org/10.1038/s41419-020-2342-2.
[34] H. Gittleman, D. Lim, M.W. Kattan, A. Chakravarti, M.R. Gilbert, A.B. Lassman, et al., An independently validated nomogram for individualized estimation of survival among patients with newly diagnosed glioblastoma: NRG Oncology RTOG 0525 and 0825, Neuro Oncol. 19 (2017) 669–677. https://doi.org/10.1093/neuonc/now208.
[35] M. Zhou, Z. Zhang, H. Zhao, S. Bao, L. Cheng, J. Sun, An immune-related six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme, Mol Neurobiol. 55 (2018) 3684–3697. https://doi.org/10.1007/s12035-017-0572-9.
[36] A. Chaddad, P. Daniel, S. Sabri, C. Desrosiers, B. Abdulkarim, Integration of radiomic and multi-omic analyses predicts survival of newly diagnosed IDH1 wild-type glioblastoma, Cancers (Basel) 11 (2019). https://doi.org/10.3390/cancers11081148.
[37] X. Zhang, H. Lu, Q. Tian, N. Feng, L. Yin, X. Xu, et al., A radiomics nomogram based on multiparametric MRI might stratify glioblastoma patients according to survival, Eur Radiol. 29 (2019) 5528–5538. https://doi.org/10.1007/s00330-019-06069-z.
[38] R. Verma, R. Correa, V.B. Hill, V. Statsevych, K. Bera, N. Beig, et al., Tumor habitat-derived radiomic features at pretreatment MRI that are prognostic for progression-free survival in glioblastoma are associated with Key morphologic attributes at histopathologic examination: a feasibility study, Radiol Artif Intell. 2 (2020) e190168. https://doi.org/10.1148/ryai.2020190168.
[39] Y. Choi, Y. Nam, J. Jang, N.Y. Shin, Y.S. Lee, K.J. Ahn, et al., Radiomics may increase the prognostic value for survival in glioblastoma patients when combined with conventional clinical and genetic prognostic models, Eur Radiol. 31 (2021) 2084–2093. https://doi.org/10.1007/s00330-020-07335-1.
[40] S. Ammari, R. Sallé de Chou, C. Balleyguier, E. Chouzenoux, M. Touat, A. Quillent, et al., A predictive clinical-radiomics nomogram for survival prediction of glioblastoma using MRI, Diagnostics. 11 (2021) 2043. https://doi.org/10.3390/diagnostics11112043.
[41] A. Fathi Kazerooni, S. Saxena, E. Toorens, D. Tu, V. Bashyam, H. Akbari, et al., Clinical measures, radiomics, and genomics offer synergistic value in AI-based prediction of overall survival in patients with glioblastoma, Sci Rep. 12 (2022) 8784. https://doi.org/10.1038/s41598-022-12699-z.
[42] G. Hajianfar, A. Haddadi Avval, S.A. Hosseini, M. Nazari, M. Oveisi, I. Shiri, et al., Time-to-event overall survival prediction in glioblastoma multiforme patients using magnetic resonance imaging radiomics, Radiol Med. 128 (2023) 1521–1534. https://doi.org/10.1007/s11547-023-01725-3.