Biological underpinnings of radiomic magnetic resonance imaging phenotypes for risk stratification in IDH wild-type glioblastoma

Fangzhan Guan[1†], Zilong Wang[1†], Yuning Qiu[1,2†], Yu Guo[1], Dongling Pei[1], Minkai Wang[1], Aoqi Xing[1], Zhongyi Liu[1], Bin Yu[1], Jingliang Cheng[3], Xianzhi Liu[1], Yuchen Ji[1], Dongming Yan[1*], Jing Yan[3*] and Zhenyu Zhang[1*]

Abstract

Background To develop and validate a conventional MRI-based radiomic model for predicting prognosis in patients with IDH wild-type glioblastoma (GBM) and reveal the biological underpinning of the radiomic phenotypes.

Methods A total of 801 adult patients (training set, N = 471; internal validation set, N = 239; external validation set, N = 91) diagnosed with IDH wild-type GBM were included. A 20-feature radiomic risk score (Radscore) was built for overall survival (OS) prediction by univariate prognostic analysis and least absolute shrinkage and selection operator (LASSO) Cox regression in the training set. GSEA and WGCNA were applied to identify the intersectional pathways underlying the prognostic radiomic features in a radiogenomic analysis set with paired MRI and RNA-seq data (N = 132). The biological meaning of the conventional MRI sequences was revealed using a Mantel test.

Results Radscore was demonstrated to be an independent prognostic factor (P < 0.001). Incorporating the Radscore into a clinical model resulted in a radiomic-clinical nomogram predicting survival better than either the Radscore model or the clinical model alone, with better calibration and classification accuracy (a total net reclassification improvement of 0.403, P < 0.001). Three pathway categories (proliferation, DNA damage response, and immune response) were significantly correlated with the prognostic radiomic phenotypes.

Conclusion Our findings indicated that the prognostic radiomic phenotypes derived from conventional MRI are driven by distinct pathways involved in proliferation, DNA damage response, and immunity of IDH wild-type GBM.

Keywords IDH wild-type glioblastoma, Conventional sequences, Machine learning, Biological underpinnings, Survival

†Fangzhan Guan, Zilong Wang and Yuning Qiu contributed equally to this work.

*Correspondence:
Dongming Yan
mrdmyan@163.com
Jing Yan
fccyanj@zzu.edu.cn
Zhenyu Zhang
fcczhangzy1@zzu.edu.cn
Full list of author information is available at the end of the article

Guan *et al. Journal of Translational Medicine*   (2023) 21:841

Page 2 of 14

Introduction

Glioblastoma (GBM) accounts for 50.1% of primary malignant brain tumors, with a median survival rate of 8 months and 5-year survival rate of 6.9% [1]. The current consensus treatment for GBM is maximal resection followed by chemoradiotherapy and chemotherapy; however, the median survival is still less than 15 months [2]. The poor prognosis of patients with GBM is partly attributed to intratumoral heterogeneity, which is reflected in complex mutations in genes and disordered biological pathways [3]. Isocitrate dehydrogenase (IDH) mutation have been established as a key prognostic factor and the pivotal molecular diagnostic marker for adult-type diffuse gliomas [4, 5]. Recently, several studies have shown that patients with IDH wild-type GBM have heterogeneous clinical outcomes [6–8]. Therefore, prognostic markers stratifying patients with IDH wild-type GBM are beneficial for guiding tumor management and informing personalized treatments.

Radiomics can non-invasively quantify tumor phenotypes by converting visual medical images into robust sub-visual digital indicators [9], and is a promising imaging biomarker in several GBM studies [10–12]. Radiogenomics, a rapidly booming field, provides biological interpretability to the data-driven nature of radiomics. Recent radiogenomics studies have uncovered the association between radiomics signatures and biological underpinnings in GBM [13–15]. However, previous radiogenomic studies have revealed the biological pathways behind radiomic features using either gene set enrichment analysis (GSEA) or weighted gene coexpression network analysis (WGCNA) methods [13, 14]. However, the focus of the two methods is different. The goal of GSEA is to determine whether genes in a gene set tend to appear enriched at the top or bottom of a preordered gene list [16], while WGCNA focuses on finding collections (modules) of genes that are synergistically expressed (with consistent trends of variation) in the overall genes [17]. Koyama et al. [18] identified the same immune-related processes using both GSEA and WGCNA methods in their study. However, few studies have investigated the biological pathways underlying radiomic phenotypes using both GSEA and WGCNA in radiogenomic analysis. Leveraging both GSEA and WGCNA to obtain the intersectional pathways for biological interpretation of the radiomics phenotypes will be more convincing with increase reproducibility and robustness.

Conventional MRI sequences have been extensively investigated in radiomics because of their accessibility and widespread use [13, 14, 19]. Different conventional MRI sequences are related to explicit tumor imaging morphologies in gliomas [20]. In particular, radiomic signatures derived from single conventional MRI sequences exhibit an excellent diagnostic value in gliomas. For example, Chen et al. [21] developed a radiomics-based model derived from the T1c sequence to differentiate gliomas from brain metastases, with an area under the ROC curve (AUC) of 0.80. Li et al. [22] confirmed that radiomic features based on the T2 FLAIR sequence can predict the expression of Ki-67, S-100, wave proteins, and CD34. However, the biological basis of each MRI sequence remains elusive, and adequate biological evidence is lacking for single-sequence radiomics applications and promotion.

Therefore, the current study has two objectives. First, a prognostic radiomic risk score (Radscore) was constructed and validated to stratify the patients with IDH wild-type GBM. Second, radiogenomic analysis utilized intersectional pathways enriched by both GSEA and WGCNA to explore the biological underpinnings of prognostic radiomic phenotypes.

Materials and methods

Study design

This study was a part of the ongoing research of the registered clinical trial "MR Based Survival Prediction of Glioma Patients Using Artificial Intelligence" (WHO International Clinical Trial Registry Platform: Clinical-Trials.gov, Trial registration ID: NCT04215211). The Human Scientific Ethics Committee of the First Affiliated Hospital of Zhengzhou University (FAHZZU) and Henan Provincial People's Hospital (HPPH) approved the study, and the requirement for written informed consent was waived due to the retrospective nature of this analysis. For patients providing fresh tumor specimens, informed consent was obtained. The study framework is shown in Fig. 1 includes two parts: radiomics profiling and radiogenomics analysis. First, an MRI-based Radscore was developed for predicting overall survival based on a training set and validated on an internal and external validation set. Then, based on the radiogenomics analysis set with both MRI and RNA-seq, two methods were used to identify the driving pathways underlying the prognostic Radscore: WGCNA and GSEA. Third, the intersectional pathways of the two methods were used to annotate prognostic radiomic phenotypes.

Study cohorts

This study collected information on patients with histologically confirmed IDH wild-type GBM from FAHZZU and HPPH, between January 2011 and December 2021. Our study cohort (n=801) had three sets: (1) a radiomics analysis set (n=710, from FAHZZU) with preoperative conventional MRI sequences including T2-weighted fluid-attenuated inversion recovery, T1-weighted gadolinium contrast-enhanced, T1-weighted, and
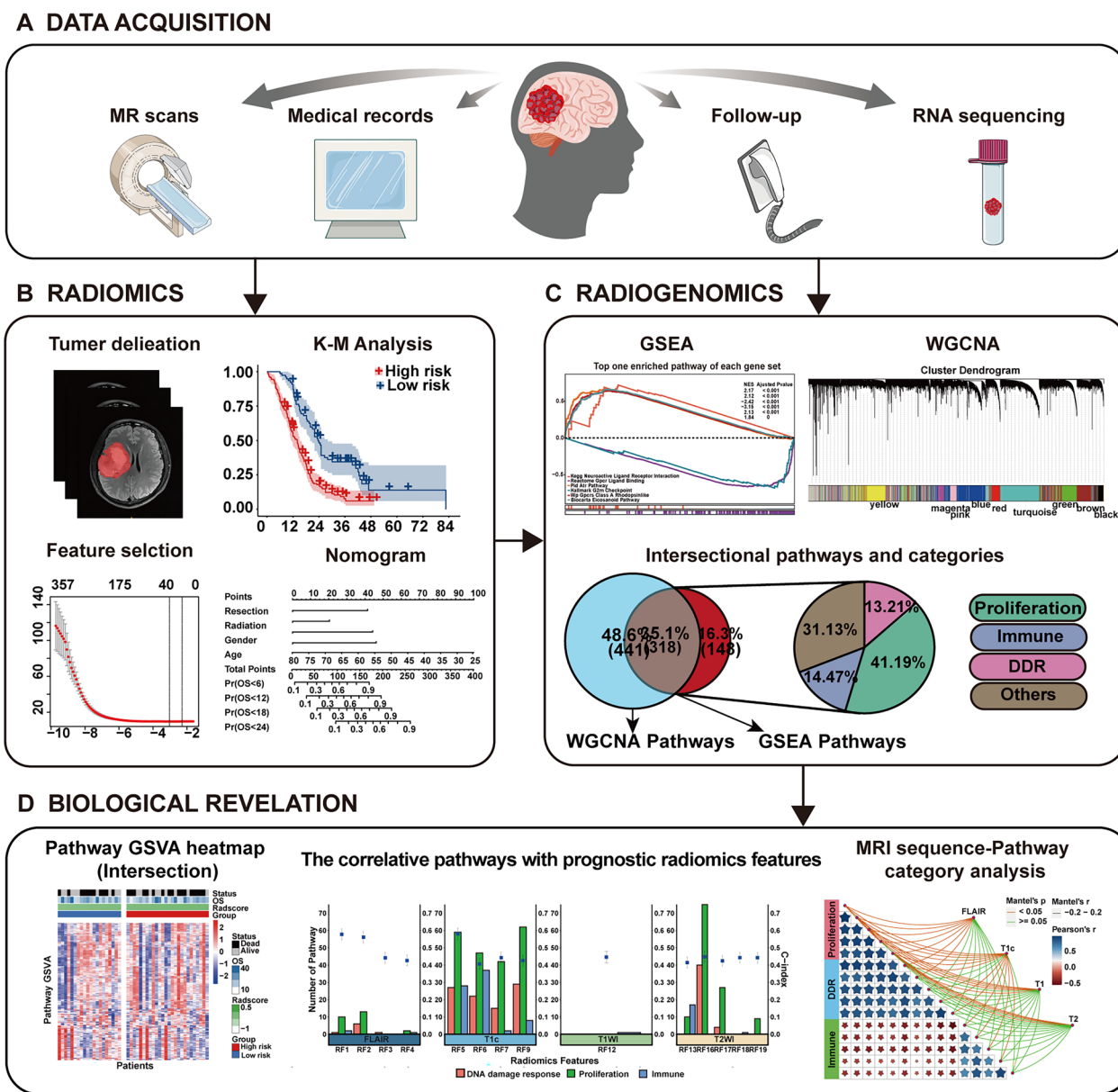
Guan *et al. Journal of Translational Medicine*   (2023) 21:841

Page 3 of 14

Fig. 1 Workflow of this study. **A** Data acquisition. **B** Radiomic model construction and validation. **C** Radiogenomics analysis: using both GSEA and WGCNA methods to reveal the biological basis behind radiomics, and obtained the intersectional pathways. **D** The intersectional pathways of the two approaches was used to explore the biological basis behind individual prognostic features and sequences

T2-weighted images (FLAIR, T1c, T1, and T2) for developing and validating the prognostic Radscore; (2) an external validation set (n = 91, from HPPH) with preoperative MRI (FLAIR, T1c, T1, and T2) to externally validate the reproducibility of the prognostic radiomic model; and (3) a radiogenomics analysis set (n = 132, from FAHZZU) with paired MRI and RNA-seq to identify biological pathways underlying the radiomic features, which is a subset of the radiomics analysis set. Specifically, patients in the radiomics analysis set were randomly divided into

a training data set (n = 471) and an internal validation set (n = 239), where clinical parameters including sex, age, preoperative Karnofsky performance status (KPS) scale, extent of resection, and adjuvant therapies were balanced. The selection procedure is shown in Additional file 1: Fig. S1.

Guan *et al. Journal of Translational Medicine*   (2023) 21:841

Page 4 of 14

Image acquisition, image processing and tumor delineation

This study followed image biomarker standardization initiative (IBSI) guidelines [23, 24]. All the necessary details are presented in Additional file 1: Table S2 to ensure the robustness of the extracted features. The modalities and parameters of the MRI are consistent between the external validation set and the radiomics analysis set and adhere to rules in Additional file 1: A1 and Table S2. First, N4ITK-based bias field distortion correction preprocessing to achieve image standardization. Subsequently, all voxels were isotropically resampled into $1 \times 1 \times 1$ mm^3 voxels using trilinear interpolation. Rigid registration was performed on the MRI for each patient using axial resampled T1c as a template with a mutual information similarity metric, generating the registered images, namely rFLAIR, rT1c, rT1, and rT2. Histogram matching was performed to normalize intensity distributions. The whole tumor area (including enhancing area, non-enhancing area, and necrosis, if any) was delineated as the signal abnormal regions in the white matter on rFLAIR images, whereas rT2w and rT1c images were used to cross-check the extension of the whole tumor areas. A neuroradiologist (J.Y.) with 12 years of experience, who was blinded to the clinical data, manually delineated the tumor contours section by section on transverse sections using open-source software (ITK-SNAP, version 3.8.0; http://www.itk-snap.org). To select robust features against intra-rater and inter-rater variations, the delineation was repeated by the same neuroradiologist (J.Y.) and another neuroradiologist (Z.Y.Z, 12 years of experience) on 100 randomly selected patients, yielding an intra-rater repeatability test dataset and an inter-rater data set, respectively.

Radiomic feature extraction

Based on this delineation, we extracted 4746 IBSI-based features using PyRadiomics (version 3.0) within the quantitative image feature pipeline from all four MRI sequences. The extracted features included shape features, first-order intensity features, and higher-order texture features. All the extracted features are summarized in Additional file 1: Table S1.

Statistical analysis
Radiomic model construction and validation

The radiomics analysis set was divided into a radiomics training subset (n = 471) and an internal validation subset (n = 239), using stratified random sampling at a ratio of 2:1 with balanced patient characteristics. A three-stage feature selection approach was used. First, to enhance the robustness of the features, any feature with an intraclass correlation coefficient (ICC) < 0.85 is discarded. Second,

we selected the features that were highly correlated with OS. The remaining features with univariate concordance index (C-index) ≥ 0.55 (positive correlation) or ≤ 0.45 (negative correlation) were selected as better prognostic variables for further analysis. Third, least absolute shrinkage and selection operator (LASSO) penalized Cox proportional hazards regression [25] was used on the training set to select the optimal feature subset. Finally, a Radscore based on 20 features (denoted as RF1–RF20 in Additional file 1: Fig. S2) was developed. A Radscore-based cutoff value calculated with the R package "survminer" divides the training set into high- and low-risk groups and is subsequently applied to the validation set. The association between Radscore and survival was evaluated using Kaplan–Meier analysis. C-index was calculated to measure the discrimination performance of the model. A calibration curve was used to assess the agreement between the predicted and observed results in the model and a decision curve was used to measure the clinical usefulness of the model. Net Reclassification Improvement (NRI) and Akaike Information Criterion (AIC) were used to assess the improved performance and potential risk of overfitting in the model, respectively.

Radiogenomics analysis

Based on the radiogenomics analysis set with both MRI and RNA-seq, the biological pathways underlying Radscore were identified using GSEA and WGCNA.

GSEA analysis for driving pathways identification

First, differentially expressed genes (DEGs) between the high- and low-risk subgroups stratified by Radscore were identified using the R package DESeq2. Then, the values of Log_2FoldChange for each gene were arranged in reverse order and used to enrich the overrepresented pathways with an R package clusterPro-filer based on six annotated genes: Kyoto Encyclopedia of Genes and Genomes (KEGG), Hallmark, Reactome, BioCarta, Pathway Interaction Database (PID), and WikiPathways (WP). False-discovery rate (FDR)-corrected P < 0.05 was considered as significant enrichment. Thereafter, a gene set variation analysis (GSVA) was used on each enriched pathway to quantify its activity [26]. Pearson correlation was performed to assess if the pathway GSVA score was significantly associated (FDR < 0.01) with the Radscore. Finally, the significantly correlated pathways were selected for further analysis.

WGCNA analysis for driving pathways identification

First, WGCNA was performed on the radiogenomics analysis set to cluster highly interconnected genes into a few gene modules that may be involved in common biological processes [17]. Modules significantly associated

Guan *et al. Journal of Translational Medicine*   (2023) 21:841

Page 5 of 14

with the Radscore were identified based on a sample-based GSVA, where FDR < 0.01 indicated significance. The detailed process of WGCNA is described in Fig. 3A, B. Within each Radscore-associated module, enrichment analyses were performed for KEGG, Hallmark, Reactome, BioCarta, PID, and WP using the R package clusterProfiler [27]. Finally, significantly correlated pathways (FDR < 0.01) were identified.

Biological pathways underlying Radiomic phenotypes

First, the intersectional pathways of GSEA and WGCNA were identified to improve the biological reliability of the enriched pathways. The identified pathways were categorized into several types. A Pearson correlation was performed on prognostic radiomic features and intersection pathways, where FDR < 0.05 indicated significance. The top feature-associated pathways of each pathway category were selected for further exploration. Finally, a Mantel test was used to assess the potential associations between the pathway categories and MRI sequences using 'vegan' R package.

Results

Patients' characteristics

In total, 801 patients were included in this study. The distribution of clinical characteristics was balanced between the training and validation sets (chi-square or Wilcoxon p-value > 0.05). The demographic and clinical information of the study cohort is summarized in Additional file 1: Table S3.

Radiomic model construction and validation

Radiomic model construction

After the three-stage feature selection, 20 features remained. Based on the LASSO COX regression model (Additional file 1: Figs. S3, S4), Radscore is calculated according to the following formula: Radscore = $0.0312141 \cdot RF1 + 0.0862303 \cdot RF2 - 0.0182051 \cdot RF3 - 0.0329717 \cdot RF4 + 0.0145359 \cdot RF5 - 0.0302801 \cdot RF6 - 0.0597537 \cdot RF7 - 0.0195800 \cdot RF8 - 0.0079536 \cdot RF9 - 0.0010230 \cdot RF10 - 0.0620081 \cdot RF11 - 0.0130946 \cdot RF12 - 0.065238 \cdot RF13 + 0.0197421 \cdot RF14 + 0.0004487 \cdot RF15 - 0.0020062 \cdot RF16 - 0.0599923 \cdot RF17 - 0.0440648 \cdot RF18 - 0.0590665 \cdot RF19 - 0.0581274 \cdot RF20$. The optimal Radscore cutoff was −0.1026. Patients were categorized into a high-risk group (Radscore of at least −0.1026) and low-risk group (Radscore less than −0.1026).

Radiomic model validation

Association of Radscore with OS was found in the training set (log-rank P < 0.001; hazard ratio [HR] = 18.55, 95% CI 11.78, 29.23) and further confirmed in the internal validation set (log-rank P < 0.001; hazard ratio [HR] = 15.68, 95% CI 8.01, 30.68) and external validation sets (log-rank P < 0.001; hazard ratio [HR] = 12.84, 95% CI 5.31, 31.04), as shown by the Kaplan–Meier curves in Fig. 2A–C, respectively. The clinical modal (CM) nomogram and radiomic-clinical modal (R-CM) nomograms are shown in Fig. 2G, I. The C-index and AIC values for the three models are summarized in Additional file 1: Table S4. With lower AIC values, the C-index of the R-CM nomogram improved by 0.047, 0.058, and 0.054 for the training, internal, and external validation sets, respectively, compared with the CM nomogram. The calibration curve of the R-CM nomogram demonstrated better agreement between the predicted and observed survival for the probability of 6-, 12-, 18-, and 24-month deaths, as shown in Fig. 2H, J. Similarly, a total NRI of 0.403 (95% CI:0.308,0.473, p < 0.001) on the R-CM nomogram indicated improved classification performance. The decision curves (Fig. 2D–F) indicated that the R-CM nomogram was more beneficial than the CM nomogram alone. Multivariate Cox analysis revealed that the Radscore was an independent risk factor (HR = 39.998; 95% CI 12.803, 124.962; P < 0.001).

GSEA analysis for driving pathways identification

Based on the radiogenomics analysis set, 646 pathways were enriched by GSEA analysis, where FDR < 0.05 was considered as significant enrichment. Then, 466 derived from the 646 pathways were obtained after performing a Pearson correlation between the enriched pathways and the Radscore, where FDR < 0.01 was considered as a significant association. A complete list of radscore-related pathways is provided in Additional file 1: Table S5, and a heatmap is shown in Fig. 3B. The top enriched pathways in each gene set are shown in Fig. 3A, C and D, respectively.

(See figure on next page.)

Fig. 2 Validation of the radiomic signature. **A**–**C** Kaplan–Meier curves for patients stratified by the Radscore (cutoff = −0.1026) in the training set (**A**), internal validation set (**B**) and external validation set (**C**). **D**–**F** Decision curve analysis (DCA) for radiomic-clinical model nomogram and clinical model nomogram to estimate the OS in the training set (**D**), internal validation set (**E**) and external validation set (**F**). The x-axis represents the threshold probability and the y-axis measures the net benefit. **G**–**J** The clinical model nomogram (**G**) and the radiomic-clinical model nomogram (**I**) for predicting the 6-, 12-, 18-, and 24-month OS, along with the calibration curves for assessment of the clinical model nomogram (**H**) and the radiomic-clinical model nomogram (**J**)
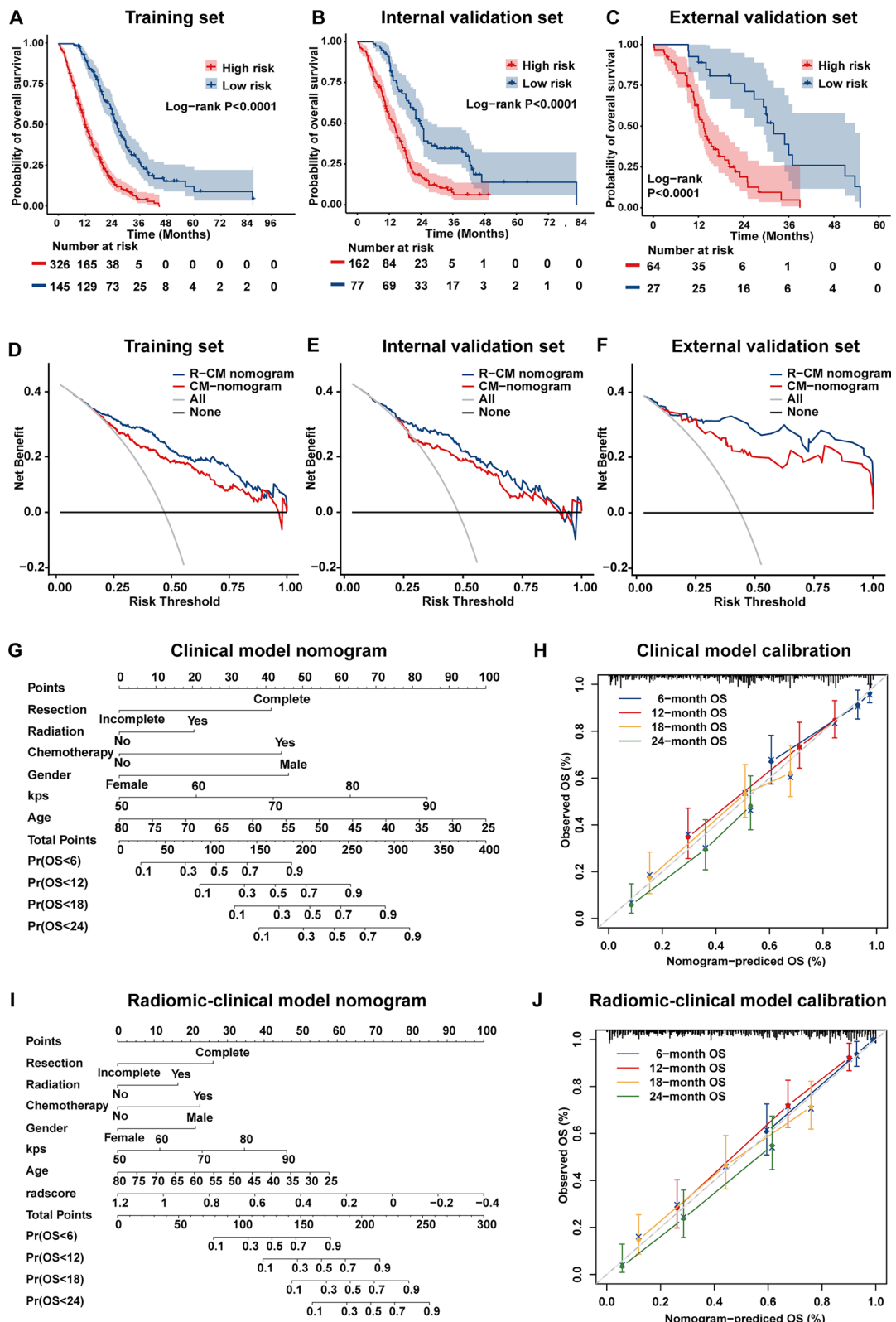
Fig. 2 (See legend on previous page.)

Guan *et al. Journal of Translational Medicine*   (2023) 21:841
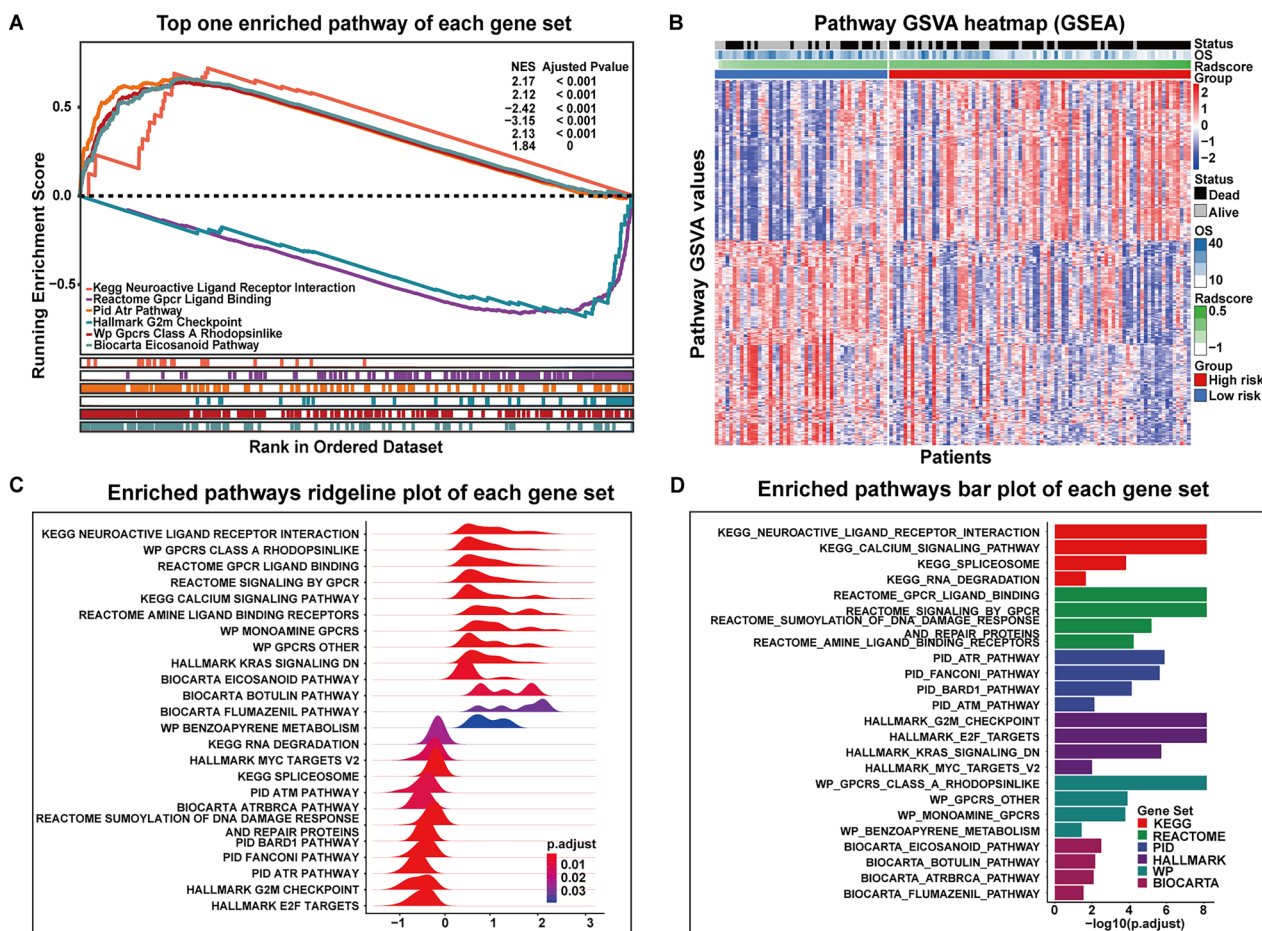
Page 7 of 14

Fig. 3 Results of gene set enrichment analysis. **A** Six representative pathways showing the most significantly enriched one from Kyoto Encyclopedia of the Genome (KEGG), Hallmark, Reactome, BioCarta, Pathway Interaction Database (PID), and WikiPathways sorted by enriched FDRs. **B** A heatmap of the distribution of clinical features and gene set variation analysis (GSVA) scores of GSEA enrichment-related pathways in 132 patients in the radiogenomics analysis set. After performing a pearson correlation, 466 pathways were selected as prognostically relevant (FDR < 0.01) and clustered according to their specific function. Patients were ranked according to their radscore, and pathway activity (GSVA values) demonstrated differences between high and low risk groups. **C** The Ridgeline plot shows 16 representative pathways (according to FDR) originating from 6 datasets were selected to demonstrate the distribution of enriched pathways between different datasets and the distribution of correlations with radscore. **D** Bar plot shows the FDR values for the top four pathways enriched by each gene set

WGCNA analysis for driving pathways identification

The results of WGCNA are summarized in Fig. 4A. Nine gene modules were derived based on the radiogenomic analysis set. In order to obtain MRI-related modules, we calculated the module GSVA values for each patient and subsequently did a pearson correlation with its corresponding Radscore. Five (turquoise module, 3722 genes; blue module, 3186 genes; brown module, 2820 genes;

(See figure on next page.)

Fig. 4 Results of weighted gene coexpression network analysis. **A** Clustering dendrogram of identified coexpression gene modules in the radiogenomics analysis set represented in the color-coded row. **B** Bubble charts show the results of GSVA values and radscore correlation of the 9 modules, FDR less than 0.01 is considered as significant correlation of modules with radscore (modules on the right side of the red line). **C** A heatmap of the distribution of clinical features and gene set variation analysis (GSVA) scores of modules enrichment-related pathways in 132 patients in the radiogenomics analysis set. After performing a Pearson correlation, 759 pathways were selected as prognostically relevant (FDR < 0.01) and clustered according to their specific function. Patients were ranked according to their radscore, and pathway activity (GSVA values) demonstrated differences between high and low risk groups. **D** Bar plot shows the FDR values for the top four pathways enriched by each gene set. **E** Bubble diagram showing enriched pathways count and generation for 16 representative pathways (according to FDR) from 6 gene sets
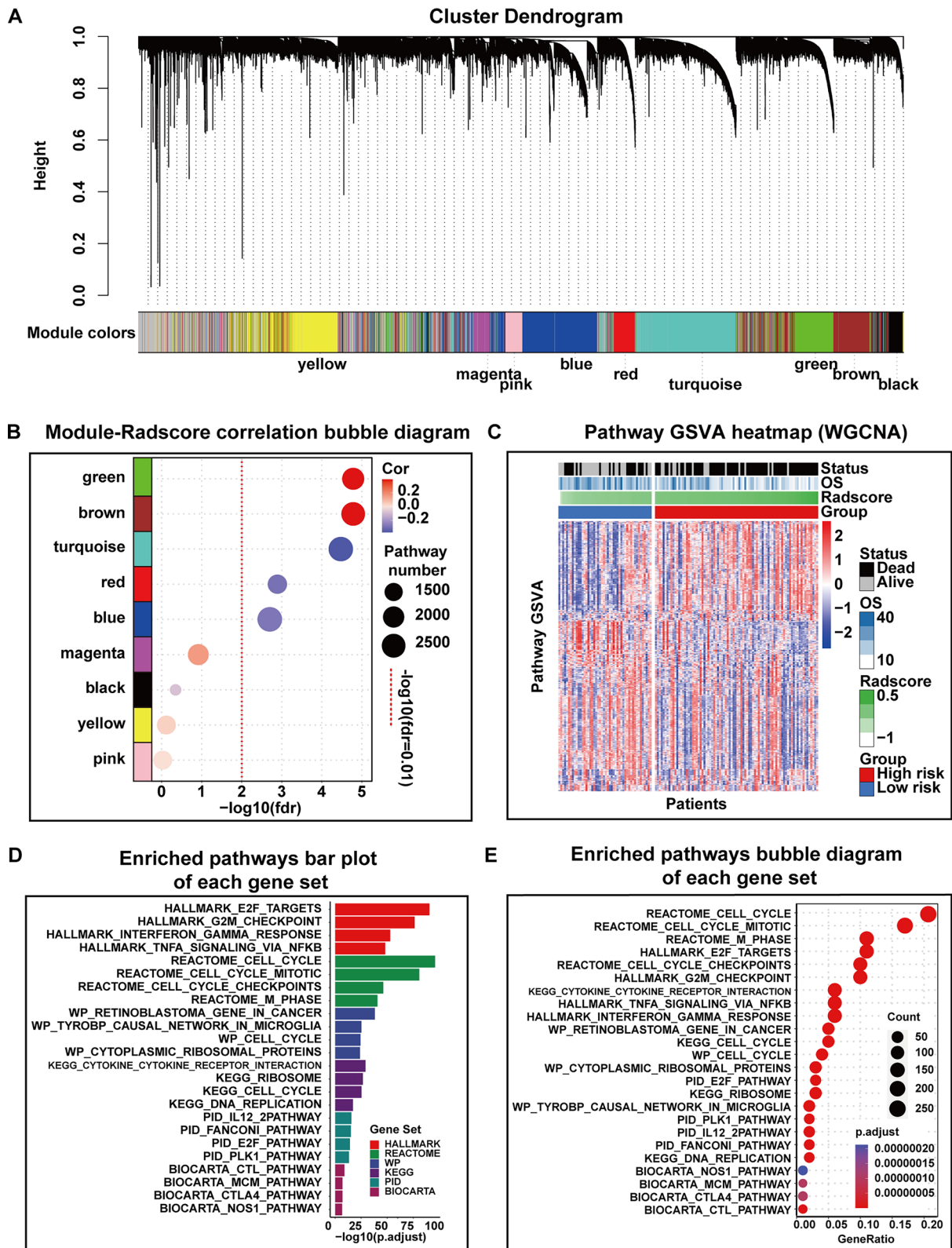
Guan *et al. Journal of Translational Medicine* (2023) 21:841

Page 8 of 14

Fig. 4 (See legend on previous page.)

green module, 2,031 genes; red module, 884 genes; as listed in Additional file 1: Table S6) of the nine modules were correlated with Radscore (Pearson correlation r = −0.37, FDR < 0.01 for turquoise module; Pearson correlation r = −0.28, FDR < 0.01 for blue module; Pearson correlation r = 0.39, FDR < 0.01 for brown module; Pearson correlation r = 0.39, FDR < 0.01 for green module; Pearson correlation r = −0.30, FDR < 0.01 for red module). The module-selection process is illustrated in Fig. 4B. A pathway enrichment analysis was performed using the selected modules. After controlling for FDR less than 0.01, 759 enriched pathways were identified, which were considered to have a significant correlation with the Radscore. A complete list of radscore-related pathways is provided in Additional file 1: Table S7. A GSVA heatmap of the enriched pathways is shown in Fig. 4C. The top enriched pathways in each gene set are shown in Fig. 4D, E.

Biological pathways underlying radiomic phenotypes
Identification of the intersectional pathways
A total of 318 intersectional pathways were derived from GSEA and WGCNA approaches, as illustrated in Fig. 5A. 318 intersectional pathways from 6 common databases: 27 from BIOCARTA, 32 from KEGG, 17 from PID, 182 from REACTOME and 51 from WP (P < 0.01), they are summarized in Additional file 1: Table S8. We annotated the biological functions of each of these pathways at different database levels by reviewing the relevant literature and accessing https://www.gsea-msigdb.org, and subsequently these pathways were classified into several different categories according to their different biological functions, including tumor proliferation, immune regulation, DNA damage response (DDR), and others (including viral infections, ion channel transport, transmitter transport, and complex cellular functions). We focused on revealing the potential biological underpinnings of the radiomic phenotype in the former three distinct pathway categories. The pathways and their categories are summarized in Additional file 1: Table S9 and presented as a heat map in Fig. 5B.

Identification of biological underpinning of the radiomic phenotypes
First, After performing a Pearson correlation with the prognosis-related features, 14 features (FLAIR of 4, T1c of 4, T1 of 1, T2 of 5) and 219 pathways (proliferation of 131, immune of 46, DDR of 42) were obtained, where the Pearson's FDR < 0.05 was considered significant. The detailed results are shown in Fig. 5C. Representative pathways for each pathway category are shown in Fig. 5D. Second, a Mantel test was performed to further explore the radiogenomics link between pathway categories (131 proliferation-associated, 46 immune-associated, and 42 DNA damage-response-associated) and MRI sequences (FLAIR 4 features, T1c 5 features, T1 3 features and T2 8 features). 170 pathways derived from intersectional pathways were identified, which were considered to have a significant correlation with the MRI sequences (P < 0.05), and the number and category of the associated pathways behind each sequence were summarized in the Fig. 6A. A heat map of each sequence with their associated top 10 pathways in the radiogenomics analysis set is shown in Fig. 6B. The top 5 typical pathways associated with each pathway category were selected are showing in Fig. 6C.

Discussion
Our research focused on the following three endeavors. First, we constructed and independently validated a Radscore calculated from preoperative conventional MRI sequences for predicting OS in adult IDH wild-type GBM. Second, in contrast to previous studies in radiogenomics that used pathways enriched by a single approach, our study acquired the intersectional pathways enriched by both GSEA and WGCNA methods to reveal the biological underpinnings behind radiomic features. Third, the biological pathways underlying the radiomic phenotypes were systematically investigated.

As the field of radiomics flourishes, biological validation will become an indispensable assessment criterion in clinical decision-making, and has recently been applied to radiomics studies [28]. For GBM, two radiogenomic studies have explored the biological meaning behind radiomic phenotypes based on conventional MRI sequences [13, 14]. However, both studies investigated patients

(See figure on next page.)
Fig. 5 Radiogenomics linking between 20 radiomic features constituting the Radscore and their significantly associated pathways. **A** Venn diagram and Pie chart of the intersective pathways. **B** A heatmap showing the activation of 318 intersective pathways clustered according to their corresponding biological pathway categories (proliferation, immunity, DNA damage response) in a high- and low-risk group of 132 patients. **C** Bar charts revealed the number of biological pathways behind individual features. Fourteen features derived from different MRI sequences have 219 different pathways (proliferation of 131, immune of 46, DDR of 42) associated behind them. **D** A bubble plots of correlation between prognostic radiomic features and classic biological pathways. Correlation results for the top 5 correlated pathways (FDR < 0.01) and prognostic features in each pathway category
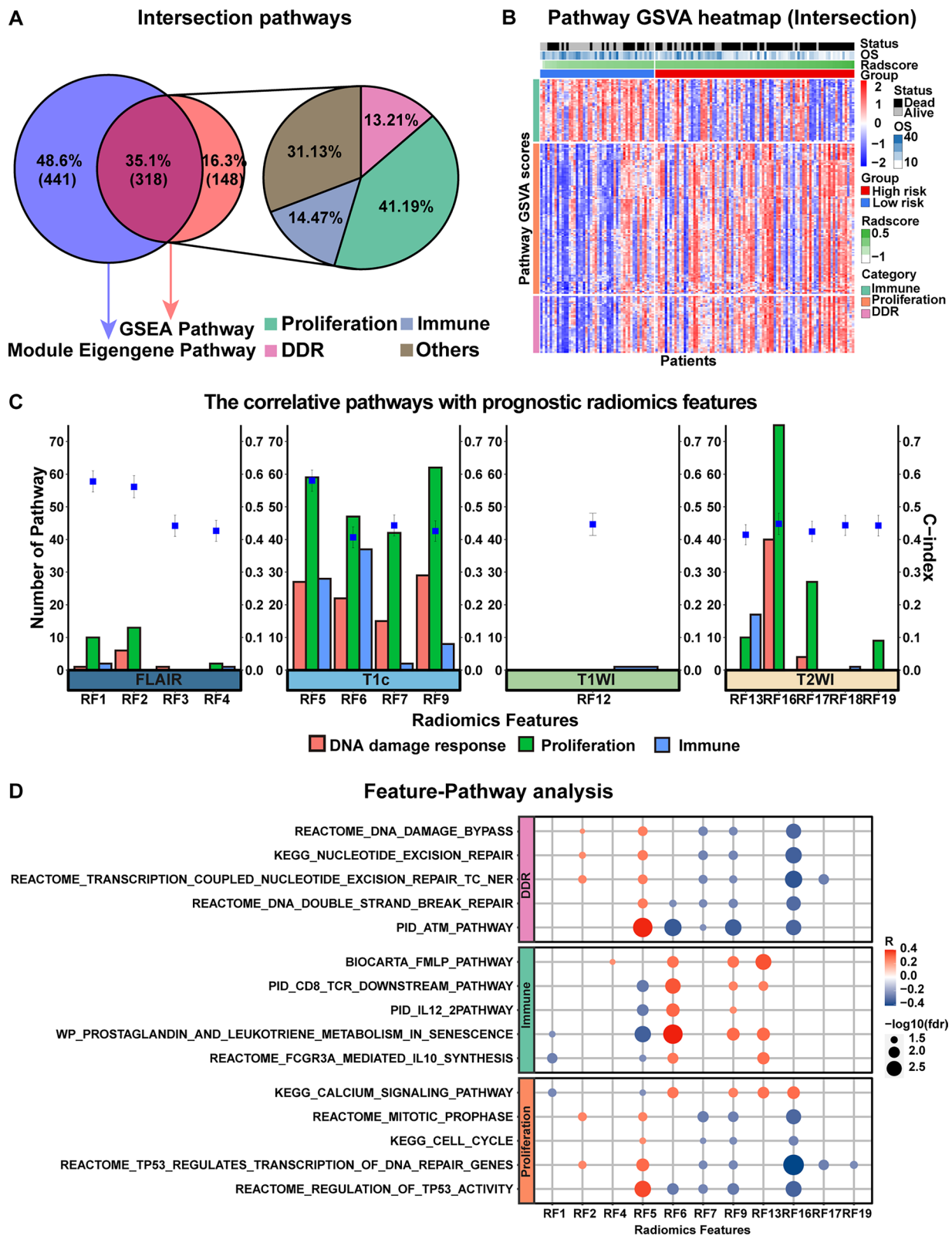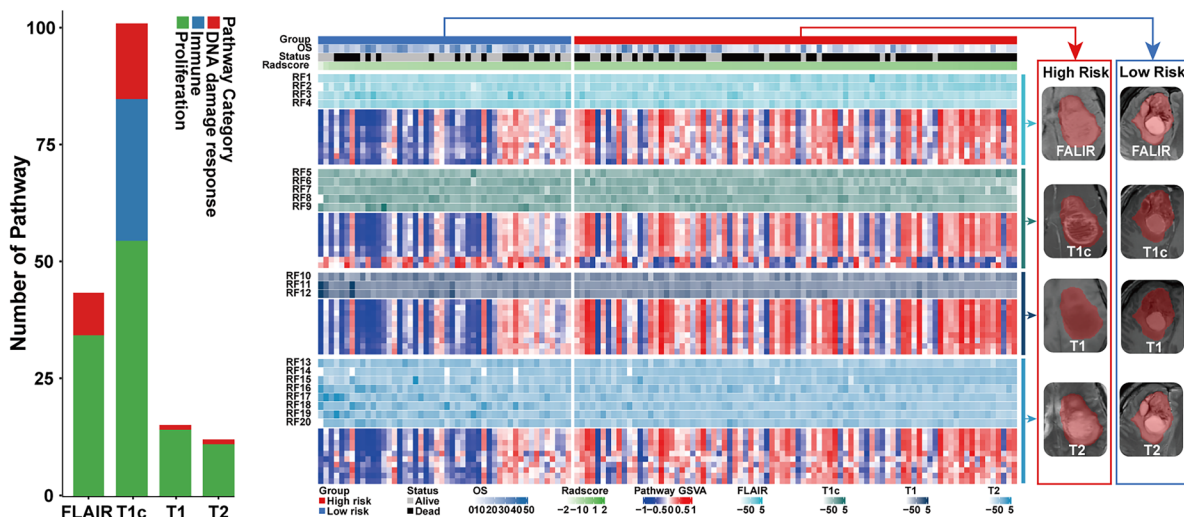
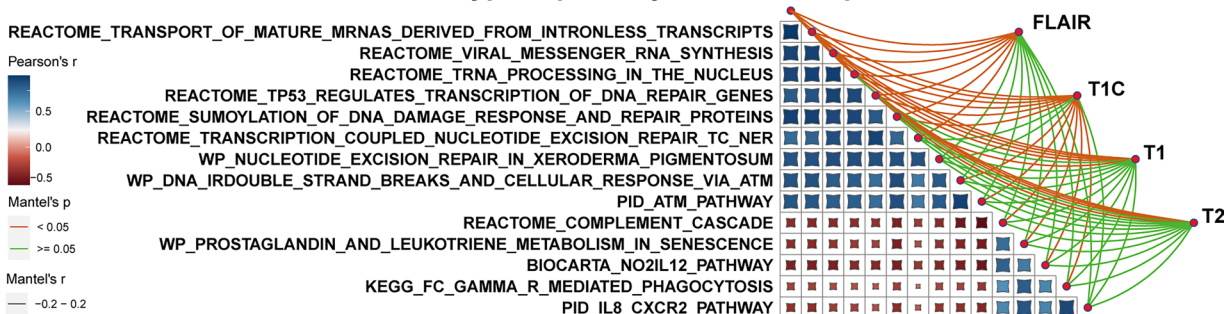**Fig. 5** (See legend on previous page.)

Fig. 6 Radiogenomics linking four MRI sequences consist of 20 prognostic radiomic features and their significantly associated pathways. **A** The mantel test results of the number and category of biological pathways behind each sequence. **B** Heatmap of four sequences along with their top 10 significantly associated pathways represented by their pathway gene set variation analysis (GSVA) score across the radiogenomics analysis set with risk groups, Radscore, overall survival, and survival status. The 10 rows immediately after each sequences represented by these prognostic features indicate the activation level (in gene set variation analysis score) of the top 10 significant pathways. **C** Mantel test results between four MRI sequences and typical pathways derived from three biological pathway categories

with pathologically diagnosed GBM which includes IDH mutant astrocytomas according to the 2021 WHO classifications of the central nervous system tumors (CNS5) [5]. Our study focused on IDH wild-type GBM, which is more in accordance with the definition of GBM according to the WHO CNS5 [5].

Several previous studies have confirmed the association of MRI features with prognosis and molecular subgroups of gliomas [29–31]. In this study, Radscore was shown to be significantly associated with patient survival prognosis (OS and status) with a higher predictive power than existing clinical indicators. Patients stratified by the same radscore in both the training and internal validation groups as well as the external validation group showed a different prognosis (log-rank P < 0.001), with the radscore

< −0.1026 subgroup surviving significantly longer than patients in the radscore ≥ −0.1026 subgroup. Similarly, in radiomic-clinical modal (R-CM) nomograms, patients with smaller radscore scores had longer overall survival. Furthermore, radcore significantly outperformed existing clinical predictors (gender, age, preoperative Karnofsky performance status (KPS) score, extent of resection, radiotherapy and chemotherapy) in the OS predictions at 6, 12, 18 and 24 months.

We elaborated on the biological meaning of individual prognostic features in terms of the categories and number of driving pathways. For the categories of driving pathways, our radiogenomics analysis revealed that five prognostic radiomic features (i.e., RF1, RF5-RF7, RF9) are associated with three major pathway categories

Guan *et al. Journal of Translational Medicine*   (2023) 21:841

Page 12 of 14

(DDR, proliferation, and immune pathways), while the RF12 are only associated with one pathway related to the immune response. These findings demonstrate that multiple biological processes may be involved in different radiomic features. It is worth noting that RF5-RF7 and RF9 all belong to the T1c sequence, whereas RF12 is the only radiomic feature derived from the T1 sequence, which indicates that the radiomic features from the T1c sequence are associated with more biological information than those from other sequences. For the number of driving pathways, we found that each of the top features (RF5, RF6, and RF16) had more than 100 associated pathways. Further radiogenomic analysis revealed that these three features were all derived from imaging textures. Consistent with previous radiomics findings [32, 33], our results also demonstrated the prominence of textural features in the prognostic imaging features. In addition, the most enriched features of DDR and proliferation pathways were RF6 derived from T1c, and RF16 derived from T2 was the most enriched feature for the immune pathways.

When it comes to MRI sequence, radiomic features extracting form T1c and FLAIR were more related to biological pathways of GBM compared to T1 and T2. For T1c, the result shows that all three pathway categories (DDR, proliferation, and immune pathways) and 100 intersectional pathways (58.8%) are associated with radiomic features derived from T1c. This partly explains why imaging features derived from T1c showed substantially incremental value in GBM prognostication [34, 35]. For FLAIR, there were two pathway categories (DDR and proliferation) and 43 intersectional pathways (25.3%) associated with radiomic features derived from FLAIR. These radiogenomic result biologically corroborates the potent performance of FLAIR in the progression of GBM [36, 37].

Combining the genetic characteristics of IDH wildtype GBM and our radiogenomic findings, we propose potential explanations for the biological mechanisms underlying the radiomic model. First, the high-risk group identified by the radiomic model is associated with proliferation and DDR pathway categories that promote GBM progression. More specifically, the positive activity of proliferation and DDR pathways in high-riks group, including MAPK signaling pathway, P53 pathway, STAT3 pathway, and DNA damage response pathways are reported to result in the induction of GBM growth [38−40]. Second, the high-risk group is also assoicated with immune pathways that promote GBM progression under immunosuppressive conditions. Previous studies have confirmed that GBM cells can inhibit the maturation and functioning of immune cells by secreting a variety of cytokines that upregulating immune checkpoint pathways such as programmed cell death protein-1 (PD-1) pathway, which leads to the progression of GBM [41, 42]. Recently, targeted therapies, such as DDR inhibitors and anti-PD-1 immunotherapy, are reported to be promising in elongation of GBM patients' clinical outcomes [43, 44]. Hence, our radiogenomic results may shed light on noninasive identification of key pathways of IDH wild-type GBM at different risk stratification, and further informing individualized treatment strategies.

This study has several limitations. First, the current study is retrospective and needs to be substantiated by prospective multicenter studies. Second, incorporation of advanced MRI sequences, such as perfusion-weighted imaging (PWI), diffusion tensor imaging (DTI), and magnetic resonance spectroscopy (MRS), may provide additional imaging information and enhance prognostication performance of the radiomic model. Third, IDH wild-type astrocytomas with TERT promoter mutations, EGFR amplification, and +7/−10 chromosome copy number changes need to be further considered in future study to fully elucidate the intratumoral heterogeneity of IDH wild-type GBM according to the WHO CNS5.

In conclusion, this study proposed a radiomic model using preoperative conventional MRI sequences to predict the clinical outcomes of patients with IDH wild-type GBM. Remarkably, radiogenomic results demonstrated that prognostic radiomic phenotypes derived from conventional MRI are associated with distinct pathways involved in proliferation, DDR, and immunity in IDH wild-type GBM.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-023-04551-3.

Additional file 1. A1: MRI sequence parameters. A2: Description of radiomic features used in our study. A3: RNA samples preparation and sequencing. A4: Detection of IDH mutation. A5: WGCNA process and module acquisition details. **Table S1.** A summary of the radiomic features extracted. **Table S2.** A summary of the parameters according to Image Biomarker Standardisation Initiative (IBSI). **Table S3.** Characteristics of patients in the training set, internal validation set and external validation set. **Table S4.** A summary of the C-index and AIC values for OS prediction of three models. **Table S5.** A summary of the Radscore-related pathways enriched by GSEA. **Table S6.** A summary of the genes in the five Radscore-related modules. **Table S7.** A summary of the pathways enriched by Radscore-related modules. **Table S8.** A summary of the intersectional pathways enriched by GSEA and WGCNA. **Table S9.** A summary of the pathway categories on proliferation, DDR and Immune. **Figure S1.** The criteria for patients' inclusion and exclusion. **Figure S2.** Forest plot of prognostic radiomic features. **Figure S3.** Radiomic feature selection 1. **Figure S4.** Radiomic feature selection 2. **Figure S5.** Incremental value of radiomic model. **Figure S6.** Radscore for each patient. **Figure S7.** Cluster dendrogram and corresponding trait heat map for each patient in WGCNA. **Figure S8.** The Soft threshold selection process. **Figure S9.** Kaplan–Meier plot of fivefold cross-validation.

Data availability
Data and materials used to support the findings of this study are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
All 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.

Author details
[1]Department of Neurosurgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, China. [2]Academy of Medical Sciences, Zhengzhou University, Zhengzhou 450001, Henan, China. [3]Department of MRI, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, China.

References
1. Ostrom QT, Price M, Neff C, Cioffi G, Waite KA, Kruchko C, Barnholtz-Sloan JS. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2015–2019. Neuro Oncol. 2022;24:v1–95.
2. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, Mirimanoff RO. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med. 2005;352:987–96.
3. Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC, Curtis C, Watts C, Tavare S. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci USA. 2013;110:4009–14.
4. Hartmann C, Hentschel B, Wick W, Capper D, Felsberg J, Simon M, Westphal M, Schackert G, Meyermann R, Pietsch T, Reifenberger G, Weller M, Loeffler M, von Deimling A. Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. Acta Neuropathol. 2010;120:707–18.
5. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW. The 2021 WHO classification of tumors of the central nervous system: a summary. Neuro Oncol. 2021;23:1231–51.
6. Brown NF, Ottaviani D, Tazare J, Gregson J, Kitchen N, Brandner S, Fersht N, Mulholland P. Survival outcomes and prognostic factors in glioblastoma. Cancers. 2022;14:3161.
7. Ramos-Fresnedo A, Pullen MW, Perez-Vega C, Domingo RA, Akinduro OO, Almeida JP, Suarez-Meade P, Marenco-Hillembrand L, Jentoft ME, Bendok BR, Trifiletti DM, Chaichana KL, Porter AB, Quinones-Hinojosa A, Burns TC, Kizilbash SH, Middlebrooks EH, Sherman WJ. The survival outcomes of molecular glioblastoma IDH-wildtype: a multicenter study. J Neurooncol. 2022;157:177–85.
8. Wang Z, Guan F, Duan W, Guo Y, Pei D, Qiu Y, Wang M, Xing A, Liu Z, Yu B, Zheng H, Liu X, Yan D, Ji Y, Cheng J, Yan J, Zhang Z. Diffusion tensor imaging-based machine learning for IDH wild-type glioblastoma stratification to reveal the biological underpinning of radiomic features. CNS Neurosci Ther. 2023. https://doi.org/10.1111/cns.14263.
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278:563–77.
10. Kickingereder P, Neuberger U, Bonekamp D, Piechotta PL, Gotz M, Wick A, Sill M, Kratz A, Shinohara RT, Jones D, Radbruch A, Muschelli J, Unterberg A, Debus J, Schlemmer HP, Herold-Mende C, Pfister S, von Deimling A, Wick W, Capper D, Maier-Hein KH, Bendszus M. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. Neuro Oncol. 2018;20:848–57.
11. Kickingereder P, Burth S, Wick A, Gotz M, Eidel O, Schlemmer HP, Maier-Hein KH, Wick W, Bendszus M, Radbruch A, Bonekamp D. Radiomic profiling of glioblastoma: identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models. Radiology. 2016;280:880–9.
12. Bae S, Choi YS, Ahn SS, Chang JH, Kang SG, Kim EH, Kim SH, Lee SK. Radiomic MRI phenotyping of glioblastoma: improving survival prediction. Radiology. 2018;289:797–806.
13. Beig N, Bera K, Prasanna P, Antunes J, Correa R, Singh S, Saeed BA, Ismail M, Braman N, Verma R, Hill VB, Statsevych V, Ahluwalia MS, Varadan V, Madabhushi A, Tiwari P. Radiogenomic-based survival risk stratification of tumor habitat on Gd-T1w MRI is associated with biological processes in glioblastoma. Clin Cancer Res. 2020;26:1866–76.
14. Sun Q, Chen Y, Liang C, Zhao Y, Lv X, Zou Y, Yan K, Zheng H, Liang D, Li ZC. Biologic pathways underlying prognostic radiomics phenotypes from paired MRI and RNA sequencing in glioblastoma. Radiology. 2021;301:654–63.
15. Yan J, Sun Q, Tan X, Liang C, Bai H, Duan W, Mu T, Guo Y, Qiu Y, Wang W, Yao Q, Pei D, Zhao Y, Liu D, Duan J, Chen S, Sun C, Wang W, Liu Z, Hong X, Wang X, Guo Y, Xu Y, Liu X, Cheng J, Li ZC, Zhang Z. Image-based deep learning identifies glioblastoma risk groups with genomic and transcriptomic heterogeneity: a multi-center study. Eur Radiol. 2022;33:904–14.
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.
17. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9: 559.
18. Koyama Y, Sato Y, Sakamoto M. POS0390 genes of "defense, response to virus" in peripheral blood of anti-MDA5 positive dermatomyositis were upregulated as compare with other forms of dermatomyositis. ~Suppressing RIG-I like receptor signaling or type 1/2 interferon signaling were the keys for survival. Ann Rheum Dis. 2022;81:450–1.
19. Zheng H, Li J, Liu H, Ting G, Yin Q, Li R, Liu M, Zhang Y, Duan S, Li Y, Wang D. MRI radiomics signature of pediatric medulloblastoma improves risk stratification beyond clinical and conventional MR imaging features. J Magn Reson Imaging. 2022;58:236–46.
20. Upadhyay N, Waldman AD. Conventional MRI evaluation of gliomas. Br J Radiol. 2011;84(Spec No 2):S107-11.

Guan *et al. Journal of Translational Medicine*   (2023) 21:841

Page 14 of 14

21. Chen C, Ou X, Wang J, Guo W, Ma X. Radiomics-based machine learning in differentiation between glioblastoma and metastatic brain tumors. Front Oncol. 2019;9: 806.

22. Li J, Liu S, Qin Y, Zhang Y, Wang N, Liu H. High-order radiomics features based on T2 FLAIR MRI predict multiple glioma immunohistochemical features: a more precise and personalized gliomas management. PLoS One. 2020;15:e227703.

23. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook G, Davatzikos C, Depeursinge A, Desseroit MC, Dinapoli N, Dinh CV, Echegaray S, El NI, Fedorov AY, Gatta R, Gillies RJ, Goh V, Gotz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar R, Lenkowicz J, Lippert F, Losnegard A, Maier-Hein KH, Morin O, Muller H, Napel S, Nioche C. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput Image-based phenotyping. Radiology. 2020;295:328–38.

24. van Griethuysen J, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan R, Fillion-Robin JC, Pieper S, Aerts H. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77:e104-7.

25. Tibshirani R. The lasso method for variable selection in the COX model. Stat Med. 1997;16:385–95.

26. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform. 2013;14: 7.

27. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.

28. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. Radiology. 2021;298:505–16.

29. Yan J, Zhang B, Zhang S, Cheng J, Liu X, Wang W, Dong Y, Zhang L, Mo X, Chen Q, Fang J, Wang F, Tian J, Zhang S, Zhang Z. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. NPJ Precis Oncol. 2021;5:72.

30. Pei D, Guan F, Hong X, Liu Z, Wang W, Qiu Y, Duan W, Wang M, Sun C, Wang W, Wang X, Guo Y, Wang Z, Liu Z, Xing A, Guo Z, Luo L, Liu X, Cheng J, Zhang B, Zhang Z, Yan J. Radiomic features from dynamic susceptibility contrast perfusion-weighted imaging improve the three-class prediction of molecular subtypes in patients with adult diffuse gliomas. Eur Radiol. 2023;33:3455–66.

31. Duan J, Zhang Z, Chen Y, Zhao Y, Sun Q, Wang W, Zheng H, Liang D, Cheng J, Yan J, Li ZC. Imaging phenotypes from MRI for the prediction of glioma immune subtypes from RNA sequencing: a multicenter study. Mol Oncol. 2023;17:629–46.

32. Korfiatis P, Kline TL, Coufalova L, Lachance DH, Parney IF, Carter RE, Buckner JC, Erickson BJ. MRI texture features as biomarkers to predict MGMT methylation status in glioblastomas. Med Phys. 2016;43:2835–44.

33. Yang D, Rao G, Martinez J, Veeraraghavan A, Rao A. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. Med Phys. 2015;42:6725–35.

34. Ellingson BM, Kim HJ, Woodworth DC, Pope WB, Cloughesy JN, Harris RJ, Lai A, Nghiemphu PL, Cloughesy TF. Recurrent glioblastoma treated with bevacizumab: contrast-enhanced T1-weighted subtraction maps improve tumor delineation and aid prediction of survival in a multicenter clinical trial. Radiology. 2014;271:200–10.

35. Boxerman JL, Zhang Z, Safriel Y, Larvie M, Snyder BS, Jain R, Chi TL, Sorensen AG, Gilbert MR, Barboriak DP. Early post-bevacizumab progression on contrast-enhanced MRI as a prognostic marker for overall survival in recurrent glioblastoma: results from the ACRIN 6677/RTOG 0625 Central Reader Study. Neuro Oncol. 2013;15:945–54.

36. Li M, Huang W, Chen H, Jiang H, Yang C, Shen S, Cui Y, Dong G, Ren X, Lin S. T2/FLAIR abnormity could be the sign of glioblastoma dissemination. Front Neurol. 2022;13: 819216.

37. Hoff BA, Lemasson B, Chenevert TL, Luker GD, Tsien CI, Amouzandeh G, Johnson TD, Ross BD. Parametric response mapping of FLAIR MRI provides an early indication of progression risk in glioblastoma. Acad Radiol. 2021;28:1711–20.

38. Vitucci M, Karpinich NO, Bash RE, Werneke AM, Schmid RS, White KK, McNeill RS, Huff B, Wang S, Van Dyke T, Miller CR. Cooperativity between MAPK and PI3K signaling activation is required for glioblastoma pathogenesis. Neuro Oncol. 2013;15:1317–29.

39. Yoon J, Grinchuk OV, Tirado-Magallanes R, Ngian ZK, Tay E, Chuah YH, Lee B, Feng J, Crasta KC, Ong CT, Benoukraf T, Ong D. E2F and STAT3 provide transcriptional synergy for histone variant H2AZ activation to sustain glioblastoma chromatin accessibility and tumorigenicity. Cell Death Differ. 2022;29:1379–94.

40. Zhang XN, Yang KD, Chen C, He ZC, Wang QH, Feng H, Lv SQ, Wang Y, Mao M, Liu Q, Tan YY, Wang WY, Li TR, Che LR, Qin ZY, Wu LX, Luo M, Luo CH, Liu YQ, Yin W, Wang C, Guo HT, Li QR, Wang B, Chen W, Wang S, Shi Y, Bian XW, Ping YF. Pericytes augment glioblastoma cell resistance to temozolomide through CCL5-CCR5 paracrine signaling. Cell Res. 2021;31:1072–87.

41. Nduom EK, Weller M, Heimberger AB. Immunosuppressive mechanisms in glioblastoma. Neuro Oncol. 2015;17(Suppl 7):i9–14.

42. Raphael I, Kumar R, McCarl LH, Shoger K, Wang L, Sandlesh P, Sneiderman CT, Allen J, Zhai S, Campagna ML, Foster A, Bruno TC, Agnihotri S, Hu B, Castro BA, Lieberman FS, Broniscer A, Diaz AA, Amankulor NM, Rajasundaram D, Pollack IF, Kohanbash G. TIGIT and PD-1 immune checkpoint pathways are associated with patient outcome and anti-tumor immunity in glioblastoma. Front Immunol. 2021;12: 637146.

43. Bonm A, Kesari S. DNA damage response in glioblastoma: mechanism for treatment resistance and emerging therapeutic strategies. Cancer J. 2021;27:379–85.

44. Zhao J, Chen AX, Gartrell RD, Silverman AM, Aparicio L, Chu T, Bordbar D, Shan D, Samanamud J, Mahajan A, Filip I, Orenbuch R, Goetz M, Yamaguchi JT, Cloney M, Horbinski C, Lukas RV, Raizer J, Rae AI, Yuan J, Canoll P, Bruce JN, Saenger YM, Sims P, Iwamoto FM, Sonabend AM, Rabadan R. Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. Nat Med. 2019;25:462–9.

Publisher's Note