Title

Deep Learning Segmentation of Infiltrative and Enhancing Cellular Tumor at Pre- and Posttreatment Multishell Diffusion MRI of Glioblastoma.

Journal

Radiology Artificial Intelligence, 6(5)

ISSN

2638-6100

Authors

Gagnon, Louis

Gupta, Diviya

Mastorakos, George

et al.

Publication Date

2024-09-01

DOI

10.1148/ryai.230489

Peer reviewed

Radiology: Artificial Intelligence

ORIGINAL RESEARCH

Deep Learning Segmentation of Infiltrative and Enhancing Cellular Tumor at Pre- and Posttreatment Multishell Diffusion MRI of Glioblastoma

Louis Gagnon, MD, PhD • Diviya Gupta, MD • George Mastorakos, MD • Nathan White, PhD • Vanessa Goodwill, MD • Carrie R. McDonald, PhD • Thomas Beaumont, MD, PhD • Christopher Conlin, PhD • Tyler M. Seibert, MD, PhD • Uyen Nguyen, BS • Jona Hattangadi-Gluth, MD • Santosh Kesari, MD, PhD • Jessica D. Schulte, MD, PhD • David Piccioni, MD, PhD • Kathleen M. Schmainda, PhD • Nikdokht Farid, MD • Anders M. Dale, PhD • Jeffrey D. Rudie, MD, PhD

From the Departments of Radiology (L.G., D.G., C.C., T.M.S., U.N., N.F., A.M.D., J.D.R.), Pathology (V.G.), Radiation Medicine and Applied Sciences (C.R.M., T.M.S., J.H.G.), Neurologic Surgery (T.B.), Bioengineering (T.M.S.), and Neurosciences (J.D.S., D.P., A.M.D.), University of California San Diego, 9500 Gillman Dr, La Jolla, CA 92093; Cortechs.ai, San Diego, Calif (G.M., N.W.); Department of Translational Neurosciences, Pacific Neuroscience Institute and Saint John's Cancer Institute at Providence Saint John's Health Center, Santa Monica, Calif (S.K.); and Department of Biophysics, Medical College of Wisconsin, Milwaukee, Wis (K.M.S.). Received October 31, 2023; revision requested December 18; final revision received June 5, 2024; accepted August 8. **Address correspondence to** L.G. (email: *louis.gagnon.9@ulaval.ca*).

Authors declared no funding for this work.

Conflicts of interest are listed at the end of this article.

Radiology: Artificial Intelligence 2024; 6(5):e230489 • https://doi.org/10.1148/ryai.230489 • Content codes: NR MR AI

Purpose: To develop and validate a deep learning (DL) method to detect and segment enhancing and nonenhancing cellular tumor on pre- and posttreatment MRI scans in patients with glioblastoma and to predict overall survival (OS) and progression-free survival (PFS).

Materials and Methods: This retrospective study included 1397 MRI scans in 1297 patients with glioblastoma, including an internal set of 243 MRI scans (January 2010 to June 2022) for model training and cross-validation and four external test cohorts. Cellular tumor maps were segmented by two radiologists on the basis of imaging, clinical history, and pathologic findings. Multimodal MRI data with perfusion and multishell diffusion imaging were inputted into a nnU-Net DL model to segment cellular tumor. Segmentation performance (Dice score) and performance in distinguishing recurrent tumor from posttreatment changes (area under the receiver operating characteristic curve [AUC]) were quantified. Model performance in predicting OS and PFS was assessed using Cox multivariable analysis.

Results: A cohort of 178 patients (mean age, 56 years ± 13 [SD]; 116 male, 62 female) with 243 MRI timepoints, as well as four external datasets with 55, 70, 610, and 419 MRI timepoints, respectively, were evaluated. The median Dice score was 0.79 (IQR, 0.53–0.89), and the AUC for detecting residual or recurrent tumor was 0.84 (95% CI: 0.79, 0.89). In the internal test set, estimated cellular tumor volume was significantly associated with OS (hazard ratio [HR] = 1.04 per milliliter; $P < .001$) and PFS (HR = 1.04 per milliliter; $P < .001$) after adjustment for age, sex, and gross total resection (GTR) status. In the external test sets, estimated cellular tumor volume was significantly associated with OS (HR = 1.01 per milliliter; $P < .001$) after adjustment for age, sex, and GTR status.

Conclusion: A DL model incorporating advanced imaging could accurately segment enhancing and nonenhancing cellular tumor, distinguish recurrent or residual tumor from posttreatment changes, and predict OS and PFS in patients with glioblastoma.

Supplemental material is available for this article.

Glioblastoma is the most common primary malignant brain tumor in the adult population (1) and has a very poor prognosis, with a median overall survival (OS) of 15–21 months and a 5-year OS of 10% (2). Multimodal MRI is used for evaluating tumor burden, surgical planning, and monitoring disease response to treatment over time both before and after surgery. However, delineating the extent of nonenhancing infiltrative tumor and distinguishing recurrent tumor from posttreatment changes remain challenging (3).

While a recent study has shown that supramarginal resection including as much infiltrative disease as feasible improves survival (4) beyond the standard treatment of gross total resection (GTR) of enhancing tumor, radiation therapy, and chemotherapy (5), the widespread application of this surgical technique remains difficult without a clear delineation of the extent of the nonenhancing infiltrative tumor from peritumoral edema. Furthermore, differentiating recurrent tumor from posttreatment changes induced by the combined therapies is crucial, as a delay in the diagnosis of tumor recurrence and progression may prevent prompt treatment, and premature diagnosis of progressive or recurrent tumor may result in unnecessary interventions. In addition, noninvasive prediction of OS and progression-free survival (PFS) could provide valuable guidance for treatment planning, yet remains a challenge (6). Patient-specific features such as age, GTR status, and expression of *MGMT* have been shown to correlate with OS and PFS. However, with the exception of age, these features are not available at presentation to inform goals-of-care conversations and do not change upfront treatment (7,8).

Deep learning (DL) has emerged as a powerful technique to analyze MRI data. Previous research has focused on delineating infiltrative nonenhancing tumor (9–14) and predicting survival (12,15–19). Despite promising prediction performance, most recent advances focus on radiomics (20), which cannot be

Abbreviations

ADC = apparent diffusion coefficient, AUC = area under the receiver operating characteristic curve, DL = deep learning, ECT = enhancing cellular tumor, FLAIR = fluid-attenuated inversion recovery, GTR = gross total resection, HR = hazard ratio, MCW = Medical College of Wisconsin, NECT = nonenhancing cellular tumor, OS = overall survival, PFS = progression-free survival, RSI = restriction spectrum imaging, TCT = total cellular tumor, UCSD = University of California San Diego

Summary

A deep learning model incorporating multishell diffusion MRI could accurately segment enhancing and nonenhancing cellular tumor, distinguish recurrent or residual tumor from posttreatment changes, and predict overall survival and progression-free survival from pre- and postoperative studies of glioblastoma.

Key Points

- A deep learning nnU-Net model trained on 243 multimodal MRI scans including multishell diffusion imaging was able to segment enhancing and nonenhancing cellular tumor, with median Dice scores for total cellular tumor of 0.79 and 0.73 in internal and external test cohorts, respectively.
- The area under the receiver operating characteristic curve for classifying residual or recurrent tumor versus posttreatment changes was 0.84 (95% CI: 0.79, 0.89) in the internal test cohort and 0.91 (95% CI: 0.82, 0.99) in the external test cohort.
- When accounting for age, sex, and gross total resection status, estimated cellular tumor volume was independently associated with overall survival (OS) (hazard ratio [HR] = 1.04 per milliliter [$P <$ .001]) and progression-free survival (HR = 1.04 per milliliter [$P <$ .001]) in the internal test set and with OS in the external postoperative (n = 70; HR = 1.03 per milliliter [P = .03]) and preoperative test cohorts (n = 610 and 419; HR = 1.01 per milliliter [P = .02] and 1.02 per milliliter [$P <$ .001], respectively).

Keywords

Segmentation, Glioblastoma, Multishell Diffusion MRI

interpreted directly by clinicians and lack generalizability across different imaging platforms and institutions (21).

To address these challenges, the current study aimed to develop and evaluate a DL approach that detects and segments enhancing and nonenhancing cellular tumor on pre- and postoperative brain MRI scans in patients with glioblastoma. The method takes advantage of the synergic combination of *(a)* nnU-Net (22), a well-established DL convolutional neural network that has excelled in the segmentation of brain tumors in multimodal MRI volumes, and *(b)* restriction spectrum imaging (RSI) (23), a multishell diffusion-weighted MRI sequence that can measure tumor cellularity. RSI has shown promise in separating nonenhancing infiltrative tumor from peritumoral edema (24) and differentiating posttreatment changes from recurrent tumor in previous studies (25).

Materials and Methods

Study Design and Patients

This retrospective study was approved by the institutional review board of University of California San Diego (UCSD), and the requirement for informed consent was waived. Data were de-identified and collected according to the Health Insurance Portability and Accountability Act standard. Salaries for two authors (G.M. and J.D.R.) were partly funded by a National Institutes of Health Small Business Innovation Research grant (no. R44NS120796) awarded to Cortechs.ai. Only one author (L.G.) had control of inclusion of any data and information that might present a conflict of interest.

A total of five different glioblastoma cohorts were used in this study. Figure 1 provides an overview of the study design. Patients diagnosed with glioblastoma (based on the 4th edition [2016] of the World Health Organization classification) from January 2010 to June 2022 at UCSD were identified from chart review of the institutional medical record (M*Modal Scout; 3M). The inclusion criterion was posttreatment MRI with the multishell diffusion sequence RSI. The pathology database was searched to enrich the sample for cases with pathology-proven progression and treatment-related changes. A total of 244 unique patients (315 timepoints) were initially identified, but 54 timepoints were rejected due to missing MRI sequences and 18 due to inability to distinguish treatment-related changes from residual or recurrent tumor with confidence. The final dataset consisted of 178 patients (243 timepoints). Overall, 192 timepoints had residual or recurrent tumor (45 pathology-proven), while the remaining 51 timepoints had only posttreatment changes (including four with pathology-proven pseudoprogression, nine with clinically confirmed pseudoprogression, 10 with radiation necrosis, and 28 with nonspecific posttreatment changes). Clinically confirmed pseudoprogression was defined as increase in size of an enhancing component in the 0–6 months following the combination of chemotherapy and radiation therapy, which subsequently decreased in size at the follow-up examinations without treatment. Radiation necrosis was defined as increase in size of a ring-enhancing lesion with necrotic core following radiation therapy either proven at pathology or identified by neuroradiologists. Nonspecific posttreatment changes included areas of treatment-related abnormal enhancement or T2 or fluid-attenuated inversion-recovery (FLAIR) hyperintensity that did not fit the criteria for pseudoprogression, radiation necrosis, or recurring disease. The number of timepoints with multifocal disease was 20. A subset of 94 patients with an MRI scan with RSI within 90 days following surgery and with OS and PFS data was included in the survival analysis.

Four external test datasets were included: two postoperative cohorts (Medical College of Wisconsin [MCW], final n = 55 timepoints [20 patients]; Longitudinal Glioblastoma MRI with Expert Response Assessment in Neuro-oncology Criteria Evaluation, or LUMIERE, dataset [26], final n = 70 timepoints [70 patients]) and two preoperative cohorts (University of Pennsylvania Glioblastoma, or UPenn-GBM [27], dataset, final n = 610 patients [610 timepoints]; and University of California San Francisco Preoperative Diffuse Glioma, or UCSF-PDGM [28], dataset, n = 419 patients [419 timepoints]). More details about the MRI sequences and inclusion criteria for these cohorts can be found in Appendix S1.

Data Acquisition and Image Preprocessing

At UCSD and MCW, each patient was imaged using a standard brain tumor protocol containing the following sequences: T1-weighted, T1-weighted contrast-enhanced, T2-weighted, FLAIR, standard diffusion-weighted images with vendor-calculated apparent diffusion coefficient (ADC), RSI multishell diffu-

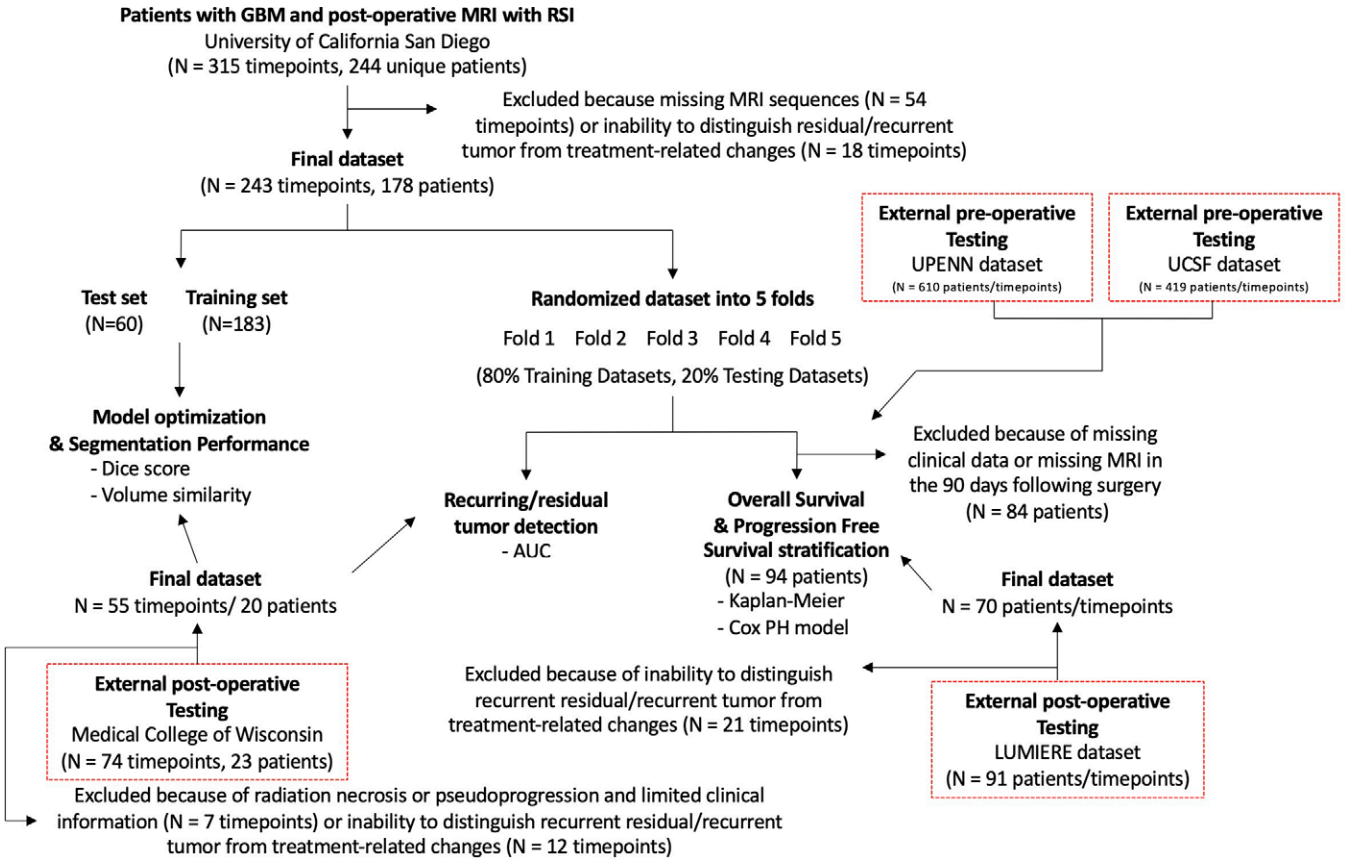

Figure 1: Overview of the experimental design of the study. For model optimization, the dataset was divided into a validation set (60 timepoints) and a training set (183 timepoints). A fivefold analysis was then performed to quantify the performance of the model to detect residual or recurring tumor. A subset of 94 patients with clinical survival data and an MRI scan in the 90-day period following surgery were included in the survival analysis. An external dataset from Medical College of Wisconsin (55 timepoints) was used to test the segmentation performance and the residual or recurring tumor detection. Another external dataset (70 patients and timepoints) from the LUMIERE collaboration was used to test the overall survival (OS) prediction. Two external datasets from the University of Pennsylvania (UPENN) (610 patients and timepoints) and University of California San Francisco (UCSF) (419 patients and timepoints) were used to test OS prediction from preoperative MRI scans. AUC = area under the receiver operating characteristic curve, GBM = glioblastoma, PH = proportional hazards, RSI = restriction spectrum imaging.

sion sequence, and dynamic susceptibility contrast perfusion imaging. The UCSD protocol also included susceptibility-weighted imaging. More details about the specific sequence parameters can be found in Appendix S2. Details about the sequences used in the other datasets are available in Appendix S1.

Image Annotations

For each patient at UCSD and MCW, total cellular tumor (TCT) volume was manually segmented in consensus by a neuroradiology attending physician (J.D.R.) and a neuroradiology fellow (L.G.). Another neuroradiology attending physician (N.F.) reviewed all the timepoints with posttreatment changes. Segmentations were performed based on all available clinical, pathologic, and imaging history, including a thorough review of prior and subsequent MRI scans. Enhancing cellular tumor (ECT) volume was extracted by multiplying the reference standard segmentation with a conventional enhancing tumor mask generated by a DL software (OnQ Neuro; Cortechs.ai). The nonenhancing cellular tumor (NECT) mask was obtained by subtracting ECT from TCT. More details can be found in Appendix S3.

DL Model

We used a nnU-Net (22) with standard 3dfullres configuration. More details can be found in Appendix S4. The final model was trained on a subset (n = 183) of the postoperative data from UCSD (Fig 1).

External Testing

The final model trained from a subset (n = 183) of the postoperative data from UCSD (Fig 1) was used for the segmentation of the four external test datasets. The segmentation performance and accuracy of the model to distinguish recurrent or residual tumor from posttreatment changes were validated on the external dataset obtained from MCW. The ability of the model to predict OS from the cellular tumor volume was validated on the postoperative LUMIERE dataset as well as the preoperative UPenn-GBM and UCSF-PDGM datasets. More details about patient selection and sequences used can be found in Appendix S1.

Statistical Analysis

Performance metrics.—Data from UCSD were randomly allocated into training (183 timepoints) and validation (60 timepoints) sets, with no patients included in both the training and validation set. The model was trained using the training set. The validation set was used to select the best model across a variety of different inputs where the optimal combination

Table 1: Patient Demographic Characteristics and Tumor Statistics

Characteristic	UCSD Dataset				External MCW Dataset	External LUMIERE Dataset	External UCSF-PDGM Dataset	External UPenn-GBM Dataset
	Total	Training Subset	Validation Subset	Survival Analysis Subset				
No. of timepoints	243	183	60	94	55	70	419	610
No. of patients	178	130	48	94	20	70	419	610
Age (y)*	56 ± 13	55 ± 14	58 ± 12	58 ± 13	53 ± 14	63 ± 10	56 ± 15	63 ± 12
Sex								
M	116 (65)	81 (62)	35 (73)	67 (71)	9 (45)	36 (51)	264 (63)	366 (60)
F	62 (35)	49 (38)	13 (27)	27 (29)	11 (55)	34 (49)	155 (37)	244 (40)
GTR	87 (49)	65 (50)	23 (48)	38 (40)	NA	51 (73)	243 (58)	362 (59)

Note.—Unless otherwise specified, data are numbers of patients, with percentages in parentheses. GTR = gross total resection, MCW = Medical College of Wisconsin, NA = not available, UCSD = University of California San Diego.
* Data are means ± SDs.

of inputs to the model was determined by the highest Dice score and volume similarity. The contribution of each individual modality to the optimal model was assessed by training a model with each specific modality alone as an input. While multifocal diseases were segmented as multiple lesions, a global Dice score and volume similarity were computed on all lesions together. A fivefold cross-validation analysis was performed with the best model and two other models to obtain predictions for each timepoint. The performance of the model to detect residual and recurrent tumor versus only treatment changes was assessed by computing the area under the receiver operating characteristic curve (AUC). Comparison between AUCs produced by the different models was assessed with a 10 000 random permutation test to avoid making an assumption about the distribution of the data (29). Cases with either progression or stability of residual disease were considered residual or recurrent. Cases with pseudoprogression, radiation necrosis, or nonspecific posttreatment changes were considered posttreatment changes. To build the receiver operating characteristic curve, the independent variable was the volume of TCT segmented by the model. The receiver operating characteristic curve was constructed by considering different levels of segmented cellular tumor volume as threshold, above which the case was considered recurring or residual. Cases with segmented cellular tumor volume below the threshold were considered posttreatment changes only (negative).

Survival analysis.—The estimated cellular tumor maps (TCT, ECT, and NECT) for each of the patients from the fivefold cross-validation model with the highest performance were used in this analysis. We tested the ability of TCT, ECT, and NECT volumes to stratify OS and PFS with use of Kaplan-Meier analysis. Multiple Cox proportional hazards models were also analyzed using different combinations of regressors, including age, sex, GTR status, *MGMT* status, and one of estimated TCT, ECT, or NECT volume. The threshold for statistically significant difference was $P < .05$. Python (Python Software Foundation) was used for statistical analysis. More details can be found in Appendix S5.

Results

Patient Characteristics

There were 1297 patients across all datasets used in this study (UCSD, $n = 178$; MCW, $n = 20$; LUMIERE, $n = 70$; UCSF-PDGM, $n = 419$; UPenn-GBM, $n = 610$). The internal cohort consisted of 178 patients (243 timepoints), with mean age of 56 years ± 13 (SD) (116 male, 62 female). Table 1 summarizes the patient characteristics of all datasets.

Segmentation Performance and Selection of the Optimal Model

Example output for models with different combinations of inputs are shown in Figure S1, and segmentation performance metrics computed for TCT, ECT, and NECT on the validation set are shown in Table 2. The combination of inputs resulting in the highest Dice score (median, 0.79 [IQR, 0.53–0.89]) and volume similarity (median, 0.93 [IQR, 0.60–0.96]) was T1-weighted, T1-weighted contrast-enhanced, FLAIR, T2-weighted, RSI cellularity, and dynamic susceptibility contrast perfusion imaging–derived cerebral blood volume. Substituting RSI cellularity with an ADC map resulted in a lower Dice score (median, 0.72 [IQR, 0.48–0.83]; $P < .001$) and volume similarity (median, 0.88 [IQR, 0.68–0.97]; $P = .002$). The individual contributions in decreasing order were T1-weighted contrast-enhanced, RSI, T1-weighted, T2-weighted, FLAIR, and dynamic susceptibility contrast perfusion imaging–derived cerebral blood volume on the basis of individual Dice scores and volume similarity (Table 2). Additional examples of TCT segmentation outputs for the optimal model are shown in Figure 2A for cases with residual or recurrent tumor and Figure 2B for cases with posttreatment changes. Individual examples of ECT and NECT segmentation outputs for the optimal model are shown in Figure 3.

Performance for Detecting Residual and Recurrent Tumor Versus Posttreatment Changes

The fivefold analysis included 192 scans with residual or recurrent tumor and 51 scans with posttreatment changes. The

Table 2: Segmentation Performance Metrics

	Dice Score		Volume Similarity	
Model Input	Mean ± SD	Median and IQR	Mean ± SD	Median and IQR
Total cellular tumor				
Optimal inputs				
T1, T1ce, T2, FLAIR, RSI, CBV	0.67 ± 0.26	0.79 (0.53–0.89)	0.77 ± 0.29	0.93 (0.60–0.96)
Other combinations of inputs				
T1, T1ce, T2, FLAIR, RSI	0.64 ± 0.30*	0.76 (0.43–0.87)	0.72 ± 0.33*	0.87 (0.54–0.96)
T1, T1ce, T2, FLAIR, CBV	0.57 ± 0.32*	0.68 (0.39–0.82)	0.70 ± 0.35*	0.87 (0.62–0.95)
T1, T1ce, T2, FLAIR, ADC, CBV	0.64 ± 0.25*	0.72 (0.48–0.83)	0.77 ± 0.27	0.88 (0.68–0.97)
T1, T1ce, T2, FLAIR, DWI, CBV	0.62 ± 0.29*	0.69 (0.44–0.84)	0.75 ± 0.31	0.91 (0.67–0.97)
T1, T1ce, T2, FLAIR	0.57 ± 0.31*	0.65 (0.35–0.82)	0.70 ± 0.32*	0.83 (0.49–0.95)
T1, T1ce, T2, FLAIR, RSI, CBV, SWI	0.66 ± 0.26	0.76 (0.50–0.86)	0.76 ± 0.28	0.90 (0.62–0.97)
T1ce	0.57 ± 0.28*	0.65 (0.39–0.81)	0.70 ± 0.31*	0.87 (0.55–0.91)
RSI	0.49 ± 0.24*	0.57 (0.37–0.71)	0.71 ± 0.30*	0.78 (0.56–0.96)
T1	0.42 ± 0.22*	0.46 (0.27–0.57)	0.62 ± 0.28*	0.72 (0.43–0.84)
T2	0.42 ± 0.22*	0.52 (0.27–0.56)	0.64 ± 0.28*	0.70 (0.47–0.87)
FLAIR	0.34 ± 0.24*	0.27 (0.14–0.56)	0.54 ± 0.33*	0.49 (0.27–0.89)
CBV	0.25 ± 0.22*	0.24 (0.00–0.37)	0.42 ± 0.35*	0.44 (0.02–0.72)
External MCW dataset				
T1, T1ce, T2, FLAIR, RSI, CBV	0.51 ± 0.35	0.73 (0.14–0.81)	0.59 ± 0.41	0.81 (0.16–0.95)
Enhancing cellular tumor				
Optimal inputs				
T1, T1ce, T2, FLAIR, RSI, CBV	0.72 ± 0.32	0.89 (0.54–0.94)	0.77 ± 0.31	0.93 (0.70–0.98)
Other combinations of inputs				
T1, T1ce, T2, FLAIR, RSI	0.68 ± 0.35*	0.89 (0.55–0.95)	0.72 ± 0.35*	0.94 (0.56–0.97)
T1, T1ce, T2, FLAIR, CBV	0.68 ± 0.36*	0.86 (0.52–0.94)	0.74 ± 0.36*	0.91 (0.75–0.98)
T1, T1ce, T2, FLAIR, ADC, CBV	0.73 ± 0.29	0.89 (0.57–0.93)	0.79 ± 0.28	0.92 (0.71–0.97)
T1, T1ce, T2, FLAIR, DWI, CBV	0.70 ± 0.33	0.88 (0.47–0.94)	0.76 ± 0.32	0.93 (0.68–0.98)
T1, T1ce, T2, FLAIR	0.68 ± 0.33*	0.83 (0.51–0.93)	0.74 ± 0.33*	0.90 (0.68–0.98)
External MCW dataset				
T1, T1ce, T2, FLAIR, RSI, CBV	0.62 ± 0.39	0.82 (0.27–0.89)	0.65 ± 0.41	0.87 (0.28–0.98)
Nonenhancing cellular tumor				
Optimal inputs				
T1, T1ce, T2, FLAIR, RSI, CBV	0.44 ± 0.26	0.49 (0.28–0.68)	0.62 ± 0.33	0.71 (0.39–0.92)
Other combinations of inputs				
T1, T1ce, T2, FLAIR, RSI	0.42 ± 0.28*	0.43 (0.15–0.70)	0.59 ± 0.35*	0.71 (0.31–0.88)
T1, T1ce, T2, FLAIR, CBV	0.23 ± 0.20*	0.19 (0.06–0.37)	0.48 ± 0.34*	0.54 (0.08–0.75)
T1, T1ce, T2, FLAIR, ADC, CBV	0.33 ± 0.23*	0.28 (0.12–0.55)	0.55 ± 0.31*	0.61 (0.27–0.83)
T1, T1ce, T2, FLAIR, DWI, CBV	0.33 ± 0.24*	0.29 (0.13–0.56)	0.57 ± 0.35*	0.67 (0.26–0.86)
T1, T1ce, FLAIR, T2	0.24 ± 0.19*	0.19 (0.05–0.38)	0.51 ± 0.32*	0.56 (0.21–0.77)
External MCW dataset				
T1, T1ce, T2, FLAIR, RSI, CBV	0.48 ± 0.32	0.63 (0.14–0.76)	0.57 ± 0.39	0.80 (0.16–0.86)

Note.—Dice scores and volume similarities are shown as means ± SDs and medians with IQRs computed on the validation subset (n = 60), except for the external Medical College of Wisconsin (MCW) dataset, for which the entire dataset was used as test set (n = 55). Cerebral blood volume (CBV) was derived from dynamic susceptibility contrast perfusion imaging. ADC = apparent diffusion coefficient, DWI = diffusion-weighted imaging, FLAIR = fluid-attenuated inversion recovery, RSI = restriction spectrum imaging, SWI = susceptibility-weighted imaging, T1ce = T1-weighted contrast-enhanced.

* Statistically significant difference at the α = .05 level (paired t test against the optimal inputs).

AUC was 0.84 (95% CI: 0.79, 0.89) for the model with both RSI and cerebral blood volume (Fig 4). The AUCs were lower (P = .04; n = 10 000 random permutation test) when RSI was substituted with ADC (AUC, 0.82 [95% CI: 0.77, 0.88]) or when only T1-weighted, T1 contrast-enhanced, T2-weighted, and FLAIR sequences were used as inputs (AUC, 0.79 [95% CI: 0.73, 0.85]). In particular, specificity was 25% (one of four patients) for pathology-proven pseudoprogression, 33% (three of nine patients) for clinically confirmed pseudoprogression, 50% (five of 10 patients) for radiation necrosis, and 86% (24 of 28 patients) for nonspecific posttreatment changes using a threshold of 1 mL of estimated cellular tumor. When

Figure 2: Example cellular tumor segmentation outputs from the optimal combination of MRI inputs. **(A)** Cases with recurrent tumor. **(B)** Cases with posttreatment changes only (ie, no cellular tumor). Top left and bottom left: clinically proven pseudoprogression (PsP); top right: pathology-proven radiation necrosis; bottom-right: pathology-proven pseudoprogression. All images are in the axial view. CBV = cerebral blood volume, Flair = fluid-attenuated inversion recovery, ref. = reference, RSI = restriction spectrum imaging, T1ce = T1-weighted contrast-enhanced.

RSI was substituted for ADC, specificity was 25% (one of four patients) for pathology-proven pseudoprogression, 33% (three of nine patients) for clinically proven pseudoprogression, 70% (seven of 10 patients) for radiation necrosis, and 75% (21 of 28 patients) for nonspecific posttreatment changes, assuming a threshold of 1 mL of tumor volume. When no diffusion was used (T1-weighted, T1-weighted contrast-enhanced, T2-weighted, FLAIR), specificity was 25% (one of four), 33% (three of nine), 30% (three of 10), and 75% (21 of 28).

Predicting Patient Survival

A subset of 94 patients (mean age, 58 years ± 13; 67 male, 27 female) with clinical survival data had an MRI examination with RSI in the 90 days following surgery. OS and PFS were

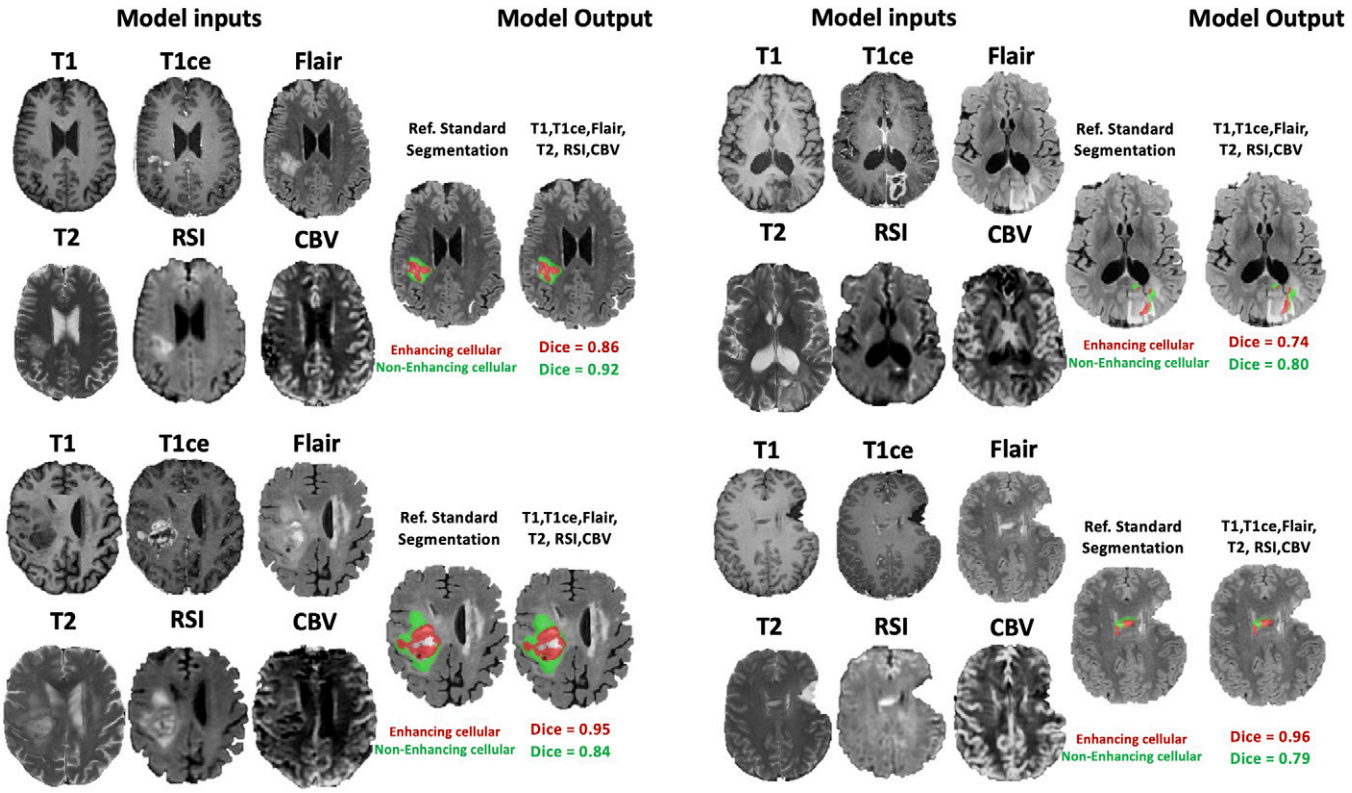

Figure 3: Example enhancing cellular tumor (red) and nonenhancing cellular tumor (green) segmentation outputs from the optimal combination of MRI inputs. All images are in the axial view. CBV = cerebral blood volume, Flair = fluid-attenuated inversion recovery, ref. = reference, RSI = restriction spectrum imaging, T1ce = T1-weighted contrast-enhanced.

stratified based on the GTR status as defined by the clinician and the estimated TCT, ECT, and NECT volumes with use of the respective median estimated tumor volume as a threshold. Kaplan-Meier plots in Figure 5 show that postoperative TCT was a statistically significant predictor of OS (hazard ratio [HR] = 1.65 [95% CI: 1.06, 2.65]; log-rank P = .02) and PFS (HR = 2.04 [95% CI: 1.30, 3.19; log-rank P = .002). HRs between the 25th and 75th percentile of tumor volume were 3.51 (95% CI: 1.80, 6.89) for OS and 4.43 (95% CI: 2.22, 9.25) for PFS. Similar Kaplan-Meier plots are shown for ECT and NECT independently.

A Cox proportional hazards model using age, sex, GTR status, TCT volume, ECT volume, and NECT volume as independent variables was analyzed for both OS and PFS, summarized in Table 3. The estimated TCT, ECT, and NECT volumes were significantly associated with both OS (TCT HR = 1.04 per milliliter [P < .001]; ECT HR = 1.04 per milliliter [P = .002]; NECT HR = 1.07 per milliliter [P = .004]) and PFS (TCT HR = 1.04 per milliliter [P < .001]; ECT HR = 1.05 per milliliter [P < .001]; NECT HR = 1.05 per milliliter [P = .02]) after adjustment for age, sex, and GTR status. The estimated TCT volumes were also significantly associated with OS and PFS after adjustment for age, sex, GTR status, and *MGMT* methylation status (OS HR = 1.03 per milliliter [P = .02]; PFS HR = 1.03 per milliliter [P = .03]). The HRs for continuous variables (such as TCT, ECT, and NECT volumes) indicate the risk per additional milliliter of tumor volume. For example, in our dataset (UCSD), the first interquartile volume for TCT was 3.2 mL, and the third interquartile volume was 24.0 mL. This corresponds to a 20.8-mL difference in TCT, and therefore a 4% per milliliter (multiplied by 20.8 mL = 83.2%) increased risk of death in a given interval between the first and third quartile volume of TCT. As a comparison, in our dataset, *MGMT* methylation reduced the risk of death by 76% (HR = 0.24) in a given interval.

In our dataset, the conventional residual enhancing tumor volume was significantly associated with both OS (HR = 1.02 per milliliter [P = .004]) and PFS (HR = 1.02 per milliliter [P = .006]) after adjustment for age, sex, and GTR status, while the surrounding FLAIR hyperintensity volume was not, either in combination with enhancing tumor (OS HR = 1.00 per milliliter [P = .20]; PFS HR = 1.00 per milliliter [P = .29]) or alone (OS HR = 1.01 per milliliter [P = .06]; PFS HR = 1.01 per milliliter [P = .08]), as shown in Table 3. However, when the estimated TCT volume (cellular) was combined with conventional residual enhancing tumor (regardless of cellularity) together with age, sex, and GTR status in a Cox proportional hazards model, TCT volume was significantly associated with both OS (HR = 1.04 per milliliter [P = .03]) and PFS (HR = 1.04 per milliliter [P = .006]) after adjustment for age, sex, and GTR status, while conventional residual enhancing tumor volume was not (OS HR = 1.00 per milliliter [P = .95]; PFS HR = 1.00 per milliliter [P = .79]).

A Cox proportional hazards model including age, sex, GTR status, and both ECT and NECT volumes as regressors was analyzed, as well as another proportional hazards model, by substituting the estimated cellular tumor volumes with the reference standard cellular tumor volumes manually segmented by the neuroradiologists. As shown in Table 3, the reference standard

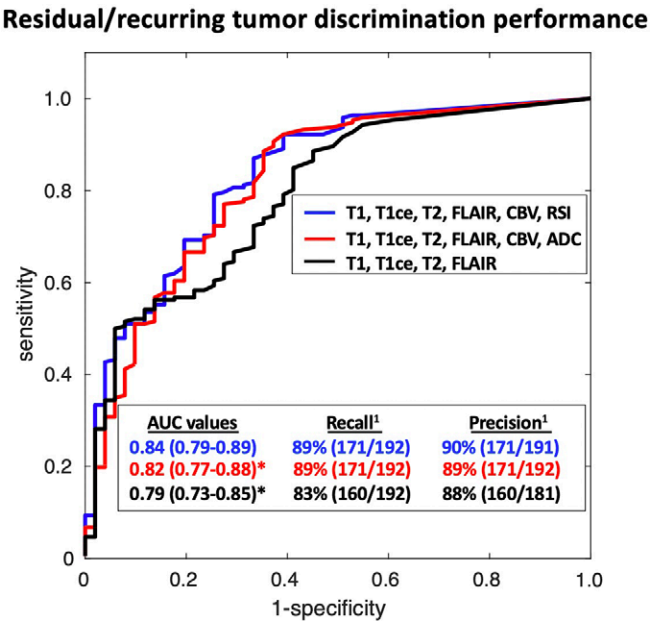

Figure 4: Receiver operating characteristic curves computed from the University of California San Diego dataset (n = 243) show the performance of the optimal combination of MRI inputs (T1-weighted, T1-weighted contrast-enhanced [T1-ce], T2-weighted, fluid-attenuated inversion recovery [FLAIR], cerebral blood volume [CBV], and restriction spectrum imaging [RSI]) to discriminate residual or recurrent tumor from posttreatment changes. The curve for when RSI is substituted for apparent diffusion coefficient (ADC) is also shown for comparison, as well as the curve when only T1-weighted, T1-weighted contrast-enhanced, and FLAIR were used as inputs to the model. Areas under the receiver operating characteristic curves (AUCs) are indicated with 95% CIs in parentheses. * Statistically significant difference at the α = .05 level (n = 10000 random permutation test). [1] Using 1 mL of tumor volume as the threshold.

NECT volumes manually segmented by the neuroradiologists were significantly associated with OS after adjustment for the ECT volumes, age, sex, and GTR status (OS HR = 1.05 per milliliter [P = .004]), while the estimated NECT volumes (from the DL model) were not (OS HR = 1.05 per milliliter [P = .07]).

Postoperative External Testing

The median Dice score computed on the external dataset from MCW (n = 55; 23 scans with residual or recurrent tumor) was 0.73 (IQR, 0.14–0.81), and the median volume similarity was 0.81 (IQR, 0.16–0.86) (Table 2). The AUC was 0.91 (95% CI: 0.82, 0.99). On the external LUMIERE dataset, both estimated TCT volume and ECT volume were independent OS predictors (HR = 1.03 per milliliter [P = .03]; HR = 1.04 per milliliter [P = .046]) (Table 4). NECT volume was not an independent predictor (HR = 1.09 per milliliter [P = .06]). PFS clinical data were not available for the external datasets.

Preoperative Survival Prediction External Testing

The HRs and P values from the Cox proportional hazards model analyzed on two cohorts (UPenn-GBM [610 patients] and UCSF-PDGM [419 patients] datasets) with preoperative MRI scans are shown in Table 4. Each of the estimated TCT, ECT, and NECT volumes were significantly associated with OS after adjustment for age, sex, and GTR status in the combined dataset (HR = 1.01 per milliliter [P < .001]; ECT HR = 1.01 per milliliter [P < .001]; NECT HR = 1.01 per milliliter [P < .001]). Results are also shown in each cohort individually. A Cox proportional hazards model was also analyzed using age, sex, GTR status, and both ECT and NECT volumes as regressors. In this case, ECT and NECT were both significantly associated with OS (HR = 1.01 per milliliter [P < .001]; HR = 1.01 per milliliter [P < .001], respectively).

When the conventional enhancing tumor volume and the surrounding T2 and/or FLAIR hyperintensity were added as additional regressors together with age, sex, and GTR status in a Cox proportional hazards model, ECT and NECT were still significantly associated with OS (HR = 1.01 per milliliter [P = .02]; HR = 1.01 per milliliter [P < .001]), while surrounding FLAIR hyperintensity was not (HR = 1.00 per milliliter [P = .10]). Kaplan-Meier plots obtained from the combined datasets are found in Figure 6 and show that the estimated NECT volume can stratify OS (HR = 1.73 [95% CI: 1.42, 2.12]; HR between the 25th and 75th percentile = 2.31 [95% CI: 1.71, 3.10]; log-rank P < .001) in patients who subsequently undergo GTR (ie, when the entire enhancing component of the tumor was subsequently removed surgically).

Discussion

Identifying nonenhancing, infiltrative tumor and differentiating recurrent enhancing tumor from posttreatment changes in patients with glioblastoma is crucial for clinical management but remains challenging even for experienced clinicians. Herein, we developed a DL method to segment enhancing and nonenhancing cellular tumor, which can assist in detecting infiltrative tumor, distinguishing recurrent or residual tumor from posttreatment changes, and predicting OS and PFS. Estimated cellular tumor volume was able to help distinguish residual or recurrent tumor from posttreatment changes in both internal (AUC, 0.84 [95% CI: 0.79, 0.89]) and external test cohorts (AUC, 0.91 [95% CI: 0.82, 0.99]). Estimated cellular tumor volumes were significantly associated with OS in pre- (HR = 1.01 per milliliter; P < .001) and postoperative cohorts (HR = 1.04 per milliliter; P < .001) and PFS in the postoperative cohort (HR = 1.04 per milliliter; P < .001) after adjustment for age, sex, and GTR status in each case and further stratified survival in preoperative patients with subsequent GTR (HR = 1.54 [95% CI: 1.26, 1.88]; HR between the 25th and 75th percentile = 2.25 [95% CI: 1.66, 3.06]).

Dice scores obtained in this study are lower than those traditionally obtained in pre- and posttreatment Brain Tumor Segmentation (ie, BraTS) challenges (30,31). This is explained by the fact that segmenting the cellular component of the tumor on postoperative studies is a more challenging task than traditional pretreatment tumor segmentation or posttreatment segmentation, which combines tumor with posttreatment changes. Not all areas of abnormal enhancement represent cellular tumor on postoperative scans, as abnormal enhancement can also reflect subacute inflammation, cytotoxic edema, granulation tissue, or pseudoprogression. NECT segmentation is an even more challenging task since infiltrative cellular tumor can have a subtle appearance that is difficult to distinguish from areas of abnormal FLAIR signal intensity. Finally, the lower mean Dice scores obtained in the external MCW test dataset were mainly due to undersegmentation. While substituting RSI for an ADC map or conventional

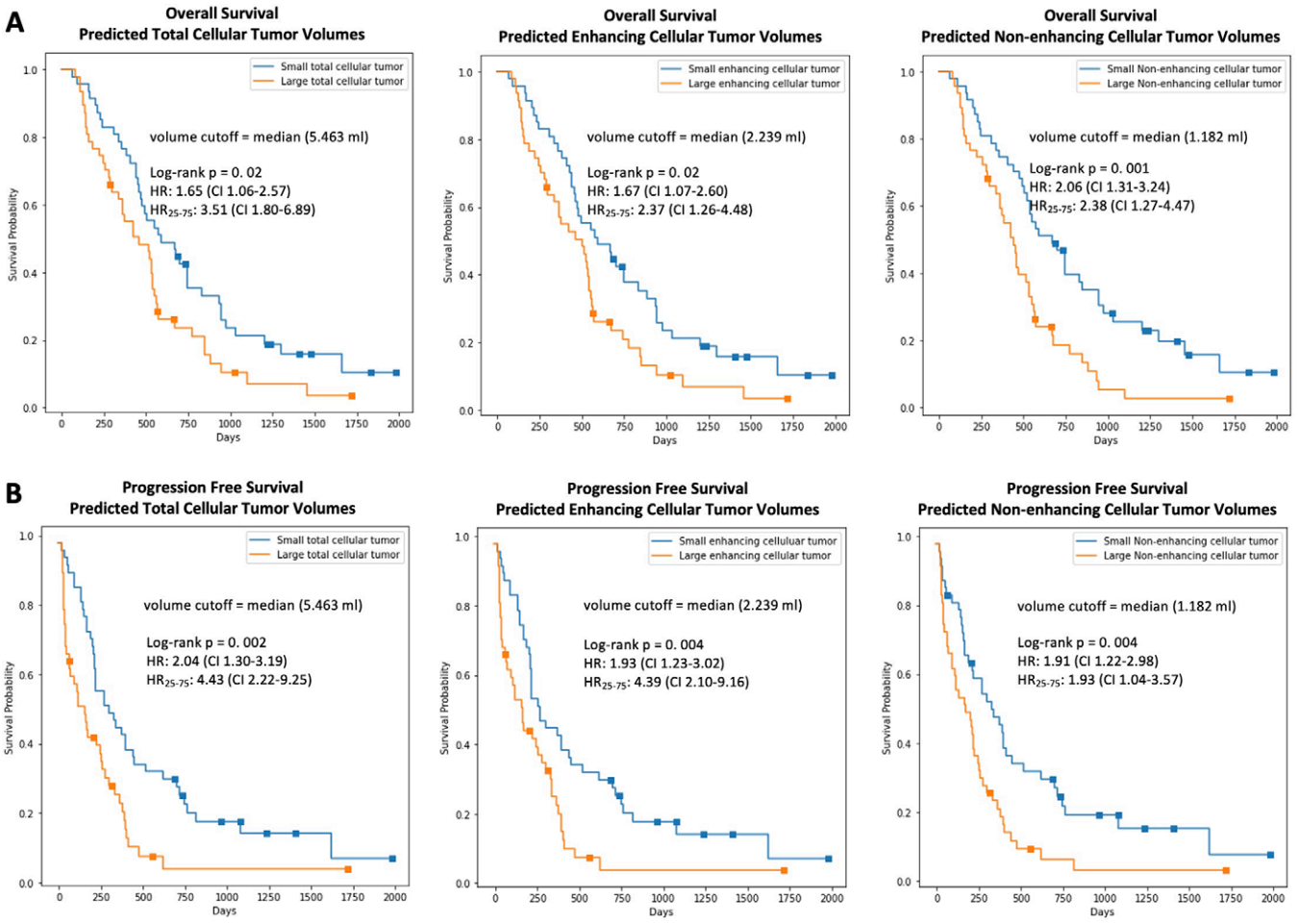

Figure 5: Kaplan-Meier survival analysis on the University of California San Diego dataset shows the ability to stratify **(A)** overall survival and **(B)** progression-free survival from the total cellular tumor volume estimated by the model on postoperative MRI scans. Log-rank *P* values and Cox hazard ratios (HRs) are also shown in the figure. $HR_{25–75}$ represents the HR between the 25th and 75th percentiles.

diffusion-weighted imaging resulted in minor differences in median Dice scores for ECT (0.89 for RSI, 0.89 for ADC, and 0.88 for diffusion-weighted imaging), larger differences were observed for NECT. For example, the median Dice score for NECT went from 0.49 with RSI to 0.29 with diffusion-weighted imaging and 0.19 when no diffusion sequence was included. This is likely due to the fact that NECT relies more on the RSI cellularity map compared with ECT, as other sequences are not as suited to separating infiltrative tumor from edema and gliosis. These results emphasize the utility of a multishell diffusion sequence to specifically improve the segmentation of NECT.

Compared with previous studies attempting to delineate infiltrative nonenhancing tumor (9–11,13,14) and to distinguish recurrent disease from posttreatment changes (DL studies [18,19] and traditional radiologist assessment [32–34]), our study used more MRI sequences (including RSI), had larger sample sizes, and included four external test datasets (two pre- and two postoperative). Interestingly, our model retained its ability to distinguish recurrent or residual tumor from posttreatment changes when RSI (AUC, 0.84) was substituted by an ADC map (AUC, 0.82) or when only conventional MRI sequences were used (AUC, 0.79), making the model usable at a larger scale across different imaging protocols. Importantly, although these last two models did not use RSI in predicting cellular tumor maps, they were trained using the reference standard segmentations performed by the radiologists on the basis of all available clinical, pathologic, and imaging data, including the RSI cellularity maps. The individual specificity for pseudoprogression, radiation necrosis, and nonspecific posttreatment changes align with the challenge of correctly identifying these entities on routine MRI follow-up studies; nonspecific posttreatment changes and radiation necrosis are easier to identify compared with pseudoprogression. The lower specificity for the four pathology-proven cases of pseudoprogression is not a surprise, as their unusual appearance prompted biopsy or resection.

Most recent imaging methods to predict survival have focused on radiomics (12,15,16,35), which cannot be interpreted directly by clinicians and is difficult to generalize across different imaging platforms and institutions. Conversely, our automated cellular tumor segmentation maps are easier for clinicians to interpret since they overlay with conventional MRI scans, with a simple prediction of cellular tumor at each voxel. Our segmentation model was validated in an external dataset with RSI, and our survival analysis was validated in the external LUMIERE dataset using a standard acquisition protocol without RSI, demonstrating the model's potential to be applied on a larger scale across institutions. Our Kaplan-Meier analysis showed that higher volumes of cellular tumor are associated

Table 3: Cox Proportional Hazards Models for Overall Survival and Progression-Free Survival Analysis in the Internal Dataset

	Overall Survival		Progression-Free Survival	
Variable	Hazard Ratio	*P* Value	Hazard Ratio	*P* Value
Standard regressor*				
Age	1.02 (1.01, 1.04)	.01†	1.02 (0.99, 1.04)	.06
Sex	0.88 (0.54, 1.44)	.61	0.66 (0.40, 1.08)	.10
GTR status	0.89 (0.55, 1.46)	.65	0.77 (0.47, 1.27)	.30
TCT	1.04 (1.02, 1.06)	<.001†	1.04 (1.02, 1.06)	<.001†
ECT	1.04 (1.02, 1.07)	.002†	1.05 (1.03, 1.08)	<.001†
NECT	1.07 (1.02, 1.11)	.004†	1.05 (1.01, 1.10)	.02†
ECT and NECT combined‡				
ECT	1.03 (1.00, 1.06)	.02†	1.05 (1.02, 1.07)	<.001†
NECT	1.05 (1.00, 1.10)	.07	1.02 (0.98, 1.08)	.33
Reference standard segmentations together‡§				
ECT	1.04 (1.01, 1.06)	.002†	1.05 (1.02, 1.07)	<.001†
NECT	1.05 (1.02, 1.09)	.004†	1.03 (0.99, 1.06)	.08
ET and TCT‖				
TCT	1.04 (1.00, 1.07)	.03†	1.04 (1.01, 1.07)	.006†
ET	1.00 (0.98, 1.03)	.95	1.00 (0.98, 1.03)	.79
ET (alone)	1.02 (1.00, 1.04)	.004†	1.02 (1.01, 1.04)	.006†
ET and ED combined‖				
ET	1.02 (1.00, 1.03)	.01†	1.02 (1.00, 1.04)	.02†
ED	1.00 (0.99, 1.01)	.20	1.00 (0.99, 1.01)	.29
ED (alone)	1.01 (0.99, 1.01)	.06	1.01 (0.99, 1.01)	.08
MGMT#				
MGMT	0.24 (0.12, 0.50)	<.001†	0.24 (0.12, 0.52)	<.001†
TCT	1.03 (1.00, 1.05)	.02†	1.03 (1.00, 1.05)	.03†
ECT	1.04 (1.01, 1.07)	.01†	1.03 (1.00, 1.06)	.02†
NECT	1.04 (0.97, 1.12)	.30	1.04 (0.97, 1.13)	.27

Note.—Data in parentheses are 95% CIs. ECT = enhancing cellular tumor, ED = surrounding fluid-attenuated inversion recovery hyperintensity, ET = enhancing tumor, GTR = gross total resection, NECT = nonenhancing cellular tumor, TCT = total cellular tumor.
* Regressors included age, sex, GTR status, and one of TCT, ECT, or NECT.
† Statistically significant difference at the α = .05 level. Hazard ratio indicates risk per additional milliliter of tumor volume.
‡ Regressors included age, sex, GTR status, ECT, and NECT.
§ Reference standard refers to the volume manually segmented by the neuroradiologists.
‖ Regressors also included age, sex, and GTR status.
Regressors included age, sex, GTR status, *MGMT* methylation status, and one of TCT, ECT, or NECT.

with worse outcomes, which is consistent with previous studies (12,15,16,35) that used DL to stratify OS and PFS. While the Kaplan-Meier analyses do not incorporate standard clinical information into the survival analysis, the Cox proportional hazards models analyzed demonstrated that the volume of estimated TCT is significantly associated with OS and PFS when adjusted for conventional clinical data such as age, sex, and GTR status, but also when adjusted for the conventional volume of residual enhancing tumor (36) and more biomarkers such as *MGMT* methylation status. Importantly, the volume of surrounding T2 or FLAIR hyperintensity was not significantly associated with OS and PFS, underscoring the need for cellular tumor maps to predict survival accurately. On the preoperative external test cohorts, estimated NECT volume allowed stratification of survival within patients who had undergone GTR. This result correlates well with the fact that supramarginal resection improves survival. In the future, such NECT segmentations may provide guidance for supramarginal resection, facilitating surgical planning by providing improved delineation of infiltrative tumor.

Our study showed that incorporating cellularity information from multishell diffusion MRI improves the segmentation of NECT and allows one to train models that are able to separate residual and recurrent tumor from posttreatment changes and edema better than conventional imaging protocols. While the specific RSI sequence used in this study is not widely available, multishell diffusion acquisitions are commonly found across scanner manufacturers and types, which allow for cellularity information to be extracted using a linear model (23,37,38). We believe that our study will motivate the need to include high-*b*-value or multishell diffusion sequences in standard neuro-oncologic MRI protocols across institutions.

Several limitations merit consideration in our study. First, the current implementation with nnU-Net does not incorporate clinical information (age, sex, *MGMT* methylation status, treatment

Table 4: Cox Proportional Hazards Models for Overall Survival in External Datasets

	Overall Survival	
Dataset and Variable	Hazard Ratio	*P* Value
External LUMIERE dataset (postoperative)*		
TCT	1.03 (1.00, 1.06)	.03†
ECT	1.04 (1.00, 1.08)	.046†
NECT	1.09 (0.99, 1.19)	.06
External UCSF and UPenn datasets (preoperative)*		
TCT	1.01 (1.01, 1.02)	<.001†
ECT	1.01 (1.01, 1.02)	<.001†
NECT	1.02 (1.01, 1.02)	<.001†
ECT and NECT combined‡		
ECT	1.01 (1.00, 1.02)	<.001†
NECT	1.01 (1.01, 1.02)	<.001†
ECT, NECT, ET, and ED combined§		
ECT	1.01 (1.00, 1.01)	.02†
NECT	1.01 (1.01, 1.02)	<.001†
ET	1.01 (1.00, 1.01)	.01†
ED	1.00 (1.00, 1.00)	.10
ET and ED combined§		
ET	1.01 (1.01, 1.02)	<.001†
ED	1.00 (1.00, 1.00)	.08
ED (alone)	1.00 (1.00, 1.00)	.90
External UCSF dataset (preoperative)*		
TCT	1.02 (1.01, 1.02)	<.001†
ECT	1.02 (1.01, 1.03)	<.001†
NECT	1.01 (1.00, 1.03)	.03†
External UPenn dataset (preoperative)*		
TCT	1.01 (1.00, 1.01)	.02†
ECT	1.01 (0.99, 1.01)	.10
NECT	1.01 (0.99, 1.01)	.11

Note.—Data in parentheses are 95% CIs. ECT = enhancing cellular tumor, ED = surrounding fluid-attenuated inversion recovery hyperintensity, ET = enhancing tumor, GTR = gross total resection, NECT = nonenhancing cellular tumor, TCT = total cellular tumor.
* Regressors included age, sex, GTR status, and one of TCT, ECT, or NECT.
† Statistically significant difference at the α = .05 level. Hazard ratio indicates risk per additional milliliter of tumor volume.
‡ Regressors included age, sex, GTR status, ECT, and NECT.
§ Regressors also included age, sex, and GTR status.

history, etc) to guide the DL model. A second limitation is the small size of our external postoperative test dataset. This stems from the limited publicly available postoperative datasets with survival information. A third limitation is the lack of histologic validation of our cellular tumor maps, which will be addressed in future work.

In conclusion, we have developed a DL method for automated segmentation of enhancing and nonenhancing cellular tumor in patients with glioblastoma. We have shown that the segmented cellular tumor maps can distinguish residual or recurrent tumor from posttreatment changes and that the segmented cellular tumor volume is associated with OS and PFS from a postoperative MRI scan and OS from a preoperative MRI scan, with validation on four external datasets. The interpretability of the segmented cellular tumor maps makes our method directly applicable in clinical workflows, highlighting its potential to guide clinical decision-making and therefore benefit both patients and health care providers. Future work will focus on histopathologic validation to further improve the interpretability of cellular maps—which may open the door for their use in surgical planning—as well as radiomics-based differentiation of recurring disease from posttreatment changes.

Author contributions: Guarantors of integrity of entire study, **L.G., J.D.R.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **L.G., T.B., J.H.G., J.D.R.**; clinical studies, **L.G., C.R.M., T.B., U.N., J.H.G., S.K., D.P., K.M.S., A.M.D., J.D.R.**; experimental studies, **L.G., G.M., T.B., D.P., K.M.S., A.M.D., J.D.R.**; statistical analysis, **L.G., G.M., A.M.D., J.D.R.**; and manuscript editing, **L.G., V.G., C.R.M., T.B., C.C., T.M.S., J.H.G., S.K., J.D.S., D.P., N.F., J.D.R.**

Data sharing: Data generated or analyzed during the study are available from the corresponding author by request.

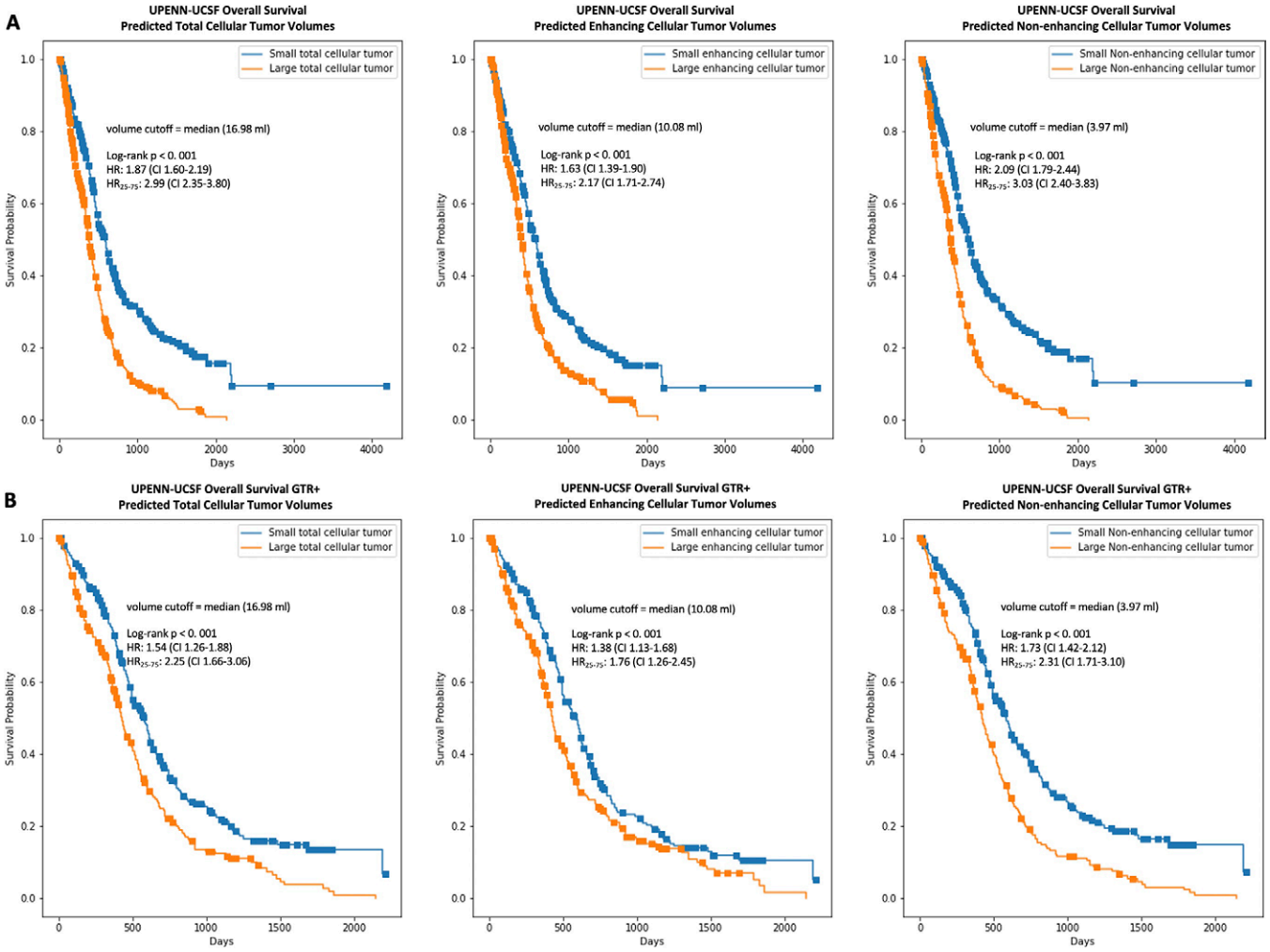

Figure 6: Kaplan-Meier survival analysis on the combined UPenn-GBM and UCSF-PDGM datasets illustrating the ability to stratify overall survival from the total, enhancing, and nonenhancing cellular tumor volumes estimated by the model on preoperative MRI scans. **(A)** All patients. **(B)** Patients who were positive for gross total resection (GTR+). Log-rank *P* values and Cox hazard ratios (HRs) are also shown on the figure. $HR_{25–75}$ represents the HR between the 25th and 75th percentiles.

Disclosures of conflicts of interest: L.G. No relevant relationships. **D.G.** No relevant relationships. **G.M.** No relevant relationships. **N.W.** National Institutes of Health Small Business Innovation Research grant no. 4R44NS120796. **V.G.** No relevant relationships. **C.R.M.** Grants from the National Institutes of Health/National Institute of Neurological Disorders and Stroke; consulting fees from Neurona. **T.B.** No relevant relationships. **C.C.** No relevant relationships. **T.M.S.** Grant to institution from GE HealthCare; consulting fees from GE HealthCare and CorTechs Labs; honoraria for educational events from GE HealthCare, Janssen, and Varian Medical Systems; support for travel/attending meetings from Janssen and Varian Medical Systems; scientific advisory board member for CorTechs Labs; stock options in CorTechs Labs. **U.N.** No relevant relationships. **J.H.G.** No relevant relationships. **S.K.** Stock in CorTechs. **J.D.S.** No relevant relationships. **D.P.** No relevant relationships. **K.M.S.** Grants to institution from the National Institutes of Health/National Cancer Institute (R01CA082500, R01CA255123, and U01CA176110); board membership for Prism Clinical Imaging (nonpaid); stock ownership interest in IQ AI and Prism Clinical Imaging; salary for author and author's spouse from Imaging Biometrics. **N.F.** Participation on a data safety monitoring board or advisory board for CorTechs. **A.M.D.** Grants from the National Institutes of Health (U24DA041123, R01AG076838, U24DA055330, and OT2HL161847); funding through a research agreement with GE HealthCare; member of the scientific advisory board for CorTechs Labs, Human Longevity, and the Mohn Medical Imaging and Visualization Centre; founder of and holds equity in CorTechs Labs. **J.D.R.** Small Business Innovation Research/Small Business Technology Transfer grant 1R44NS120796-01A1 awarded to Cortechs.ai; American Society of Neuroradiology Research Grant in Artificial Intelligence; consulting fees from Cortechs.ai; medical advisory board member for Cortechs.ai and Subtle Medical; associate editor at *Radiology: Artificial Intelligence*; stock/stock options in Cortechs.ai and Subtle Medical.

References

1. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol 2021;23(8):1231–1251.
2. Gutman DA, Cooper LAD, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. Radiology 2013;267(2):560–569.
3. Taylor C, Ekert JO, Sefcikova V, Fersht N, Samandouras G. Discriminators of pseudoprogression and true progression in high-grade gliomas: a systematic review and meta-analysis. Sci Rep 2022;12(1):13258.
4. Vivas-Buitrago T, Domingo RA, Tripathi S, et al. Influence of supramarginal resection on survival outcomes after gross-total resection of IDH-wild-type glioblastoma. J Neurosurg 2021;136(1):1–8.
5. Wick W, Gorlia T, Bendszus M, et al. Lomustine and bevacizumab in progressive glioblastoma. N Engl J Med 2017;377(20):1954–1963.
6. Smits M. MRI biomarkers in neuro-oncology. Nat Rev Neurol 2021;17(8):486–500.
7. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 2005;352(10):997–1003.
8. Poulsen SH, Urup T, Grunnet K, et al. The prognostic value of FET PET at radiotherapy planning in newly diagnosed glioblastoma. Eur J Nucl Med Mol Imaging 2017;44(3):373–381.
9. Akbari H, Macyszyn L, Da X, et al. Imaging surrogates of infiltration obtained via multiparametric imaging pattern analysis predict subsequent location of recurrence of glioblastoma. Neurosurgery 2016;78(4):572–580.
10. Rathore S, Akbari H, Doshi J, et al. Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma: implications for personalized radiotherapy planning. J Med Imaging (Bellingham) 2018;5(2):021219.
11. Yan JL, Li C, van der Hoorn A, Boonzaier NR, Matys T, Price SJ. A neural network approach to identify the peritumoral invasive areas in glioblastoma patients by using MR radiomics. Sci Rep 2020;10(1):9748. [Published correction appears in Sci Rep 2020;10(1):13808.]
12. Yan J, Zhang B, Zhang S, et al. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. NPJ Precis Oncol 2021;5(1):72.

13. Dasgupta A, Geraghty B, Maralani PJ, et al. Quantitative mapping of individual voxels in the peritumoral region of IDH-wildtype glioblastoma to distinguish between tumor infiltration and edema. J Neurooncol 2021;153(2):251–261.
14. Riahi Samani Z, Parker D, Akbari H, et al. Artificial intelligence-based locoregional markers of brain peritumoral microenvironment. Sci Rep 2023;13(1):963.
15. Bae S, Choi YS, Ahn SS, et al. Radiomic MRI phenotyping of glioblastoma: improving survival prediction. Radiology 2018;289(3):797–806.
16. Pálsson S, Cerri S, Poulsen HS, Urup T, Law I, Van Leemput K. Predicting survival of glioblastoma from automatic whole-brain and tumor segmentation of MR images. Sci Rep 2022;12(1):19744.
17. Ismail M, Prasanna P, Bera K, et al. Radiomic deformation and textural heterogeneity (R-DepTH) descriptor to characterize tumor field effect: application to survival prediction in glioblastoma. IEEE Trans Med Imaging 2022;41(7):1764–1777.
18. Park JE, Kim HS, Jo Y, et al. Radiomics prognostication model in glioblastoma using diffusion- and perfusion-weighted MRI. Sci Rep 2020;10(1):4250.
19. Shim KY, Chung SW, Jeong JH, et al. Radiomics-based neural network predicts recurrence patterns in glioblastoma using dynamic susceptibility contrast-enhanced MRI. Sci Rep 2021;11(1):9974.
20. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48(4):441–446.
21. Booth TC, Williams M, Luis A, Cardoso J, Ashkan K, Shuaib H. Machine learning and glioma imaging biomarkers. Clin Radiol 2020;75(1):20–32.
22. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18(2):203–211.
23. White NS, Leergaard TB, D'Arceuil H, Bjaalie JG, Dale AM. Probing tissue microstructure with restriction spectrum imaging: histological and theoretical validation. Hum Brain Mapp 2013;34(2):327–346.
24. White NS, McDonald CR, Farid N, Kuperman JM, Kesari S, Dale AM. Improved conspicuity and delineation of high-grade primary and metastatic brain tumors using "restriction spectrum imaging": quantitative comparison with high B-value DWI and ADC. AJNR Am J Neuroradiol 2013;34(5):958–964, S1.
25. Khan UA, Rennert RC, White NS, et al. Diagnostic utility of restriction spectrum imaging (RSI) in glioblastoma patients after concurrent radiation-temozolomide treatment: a pilot study. J Clin Neurosci 2018;58:136–141.
26. Suter Y, Knecht U, Valenzuela W, et al. The LUMIERE dataset: Longitudinal Glioblastoma MRI with expert RANO evaluation. Sci Data 2022;9(1):768.
27. Bakas S, Sako C, Akbari H, et al. The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics. Sci Data 2022;9(1):453.
28. Calabrese E, Villanueva-Meyer JE, Rudie JD, et al. The University of California San Francisco Preoperative Diffuse Glioma MRI Dataset. Radiol Artif Intell 2022;4(6):e220058.
29. Good P. Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses. Springer Series in Statistics. Springer, 2000.
30. Baid U, Ghodasara S, Mohan S, et al. The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification. arXiv 2107.02314 [preprint] https://arxiv.org/abs/2107.02314. Posted July 5, 2021. Updated September 12, 2021. Accessed June 4, 2024.
31. de Verdier MC, Saluja R, Gagnon L, et al. The 2024 Brain Tumor Segmentation (BraTS) Challenge: Glioma Segmentation on Post-treatment MRI. arXiv 2405.18368 [preprint] https://arxiv.org/abs/2405.18368. Posted May 28, 2024. Accessed June 4, 2024.
32. Wang S, Martinez-Lage M, Sakai Y, et al. Differentiating tumor progression from pseudoprogression in patients with glioblastomas using diffusion tensor imaging and dynamic susceptibility contrast MRI. AJNR Am J Neuroradiol 2016;37(1):28–36.
33. Manning P, Daghighi S, Rajaratnam MK, et al. Differentiation of progressive disease from pseudoprogression using 3D PCASL and DSC perfusion MRI in patients with glioblastoma. J Neurooncol 2020;147(3):681–690.
34. Yu Y, Ma Y, Sun M, Jiang W, Yuan T, Tong D. Meta-analysis of the diagnostic performance of diffusion magnetic resonance imaging with apparent diffusion coefficient measurements for differentiating glioma recurrence from pseudoprogression. Medicine (Baltimore) 2020;99(23):e20270.
35. Sun L, Zhang S, Chen H, Luo L. Brain tumor segmentation and survival prediction using multimodal MRI scans with deep learning. Front Neurosci 2019;13:810.
36. Ellingson BM, Abrey LE, Nelson SJ, et al. Validation of postoperative residual contrast-enhancing tumor volume as an independent prognostic factor for overall survival in newly diagnosed glioblastoma. Neuro Oncol 2018;20(9):1240–1250.
37. Jeurissen B, Tournier JD, Dhollander T, Connelly A, Sijbers J. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. Neuroimage 2014;103:411–426.
38. Li CX, Patel S, Zhang X. Evaluation of multi-shell diffusion MRI acquisition strategy on quantitative analysis using multi-compartment models. Quant Imaging Med Surg 2020;10(4):824–834.